

THE SCOVILLE SPLICE LIBRARY · CRAFT & SCIENCE

Precision & Consistency

One spec. One flavor. *Batch after batch.*

How a designed flavor becomes a specification a formulator can build a product on — the honest mathematics of drift, the four doors it walks through, and the quality system that holds every axis of every organism to a batch variation at or under seven percent.

SCOVILLE SPLICE · GROWN BY OUR FAMILY. ENGINEERED FROM SCRATCH.

Contents

01 Opening essay · Why a number has to hold

02 Chapter 1 · The coefficient of variation — the honest measure of drift.

03 Chapter 2 · Seven percent — the number, and the mouth that sets it.

04 Chapter 3 · Four doors — where drift comes from, and why it adds in quadrature.

05 Chapter 4 · The dose–response curve — EC_{50}, and what we actually measure.

06 Chapter 5 · The genetic anchor — why one genotype starts tighter than a crowd.

07 Chapter 6 · The room — controlled environment as a variance tool.

08 Chapter 7 · The process — the compounds that survive the journey.

09 Chapter 8 · The assay, and the ruler's own wobble — gauge R&R.

10 Chapter 9 · The reference standard — the fixed point the whole scale hangs on.

11 Chapter 10 · The variance budget — spending your tolerance on purpose.

12 Chapter 11 · Stacking — how variation shrinks when you combine.

13 Chapter 12 · Process capability — Cp, Cpk, and the sigma levels.

14 Chapter 13 · The control chart — statistical process control, and the two kinds of cause.

15 Chapter 14 · Lot release — acceptance sampling, and the two risks.

16 Chapter 15 · The heat ladder — every rung exactly where it says.

17 Chapter 16 · Holding eight at once — the Phoenix problem.

18 Chapter 17 · Stability — the number through time and heat.

19 Chapter 18 · The certificate of analysis — what ships with every lot.

20 Chapter 19 · Reformulation — what a stable input unlocks.

21 Chapter 20 · Consistency by design — built in, not inspected in.

22 Closing essay · The same good thing, every time

23 Appendix · The spec reference

Opening essay · Why a number has to hold

There is a sentence near the front of this library, and everything we make rests on it being true. A recipe that calls for IFI-H 900, we wrote, is an instruction — one that tastes the same in August and in February, in Montecito and in Seoul, because it is anchored to the receptor and not to a memory. It is a lovely promise. Read it closely and it is not one promise but two, and the second is far harder than the first.

The first promise is that a flavor can be *measured* — that you can set a single honest number against how loudly a plant speaks to a given receptor, and that the number will mean the same thing to anyone who reads it:

> **IFI = 1,000 × EC_{50}(reference) ÷ EC_{50}(candidate)**

That promise the index keeps on the page. EC_{50} — the concentration at which a stimulus drives its receptor to half of full response — is a physical fact about a molecule and a channel, and dividing a fixed reference by a candidate turns loudness into a figure a tomato, a chili, and a cup of tea can be read against on one scale. The second promise hides inside a single small word: *same.* For IFI-H 900 to taste the same every time, the plant has to *be* 900 every time — this harvest and the next, this lot and the one that ships a year from now, across a thousand fruits grown in rooms none of them will ever see. A number anyone can measure once is a measurement. A number you can hit again, and again, on demand, at scale, is a *specification.* This chapter is about the distance between those two words, and the quality system we built to close it.



The research greenhouse at dawn — where a question about the same good thing, set down on the table every time, first had to become a question about variance.

We did not come to this discipline as accountants. We came to it as a family that cooks for the people it loves, and that has always known the deepest form of care is not a single perfect meal — it is the *same* good thing, set down every time, so that no one has to wonder whether tonight will be the night it is off. Consistency, for us, is reverence made repeatable. A specification is a promise kept to a stranger you will never meet, who will taste what you made a year from now, in a city you have never seen, and find it exactly as you meant it. Everything that follows — the dose–response curve, the

ratio that cancels the room, the four doors, the pinned dawn, the certificate that rides with every lot — is only the long, careful explanation of how that promise is kept.

The two promises, and why the second is the hard one

A measurement is a photograph; a specification is a metronome. The first says *this fruit, on this bench, this morning, read 4,200.* The second says *every fruit this line will ever ship will read 4,200, close enough that your mouth cannot tell them apart.* Nature is generous with the first and withholding of the second. The same wild chili bush, picked a week apart, can swing by half; two fruits from the same branch, one shaded and one in sun, can disagree more than two different cultivars do. The plant is not being dishonest. It is simply doing what a living thing does — responding to light, to water, to the hour it was picked — and every one of those responses is a small motion away from the number.

So the real work of a quality system is not measuring. Measuring is the easy promise. The hard promise is *holding* — driving the wander out of a living process until a thousand fruits become one fruit, repeated. That requires a single honest figure for how far a thing strays from itself, and one piece of mathematics gives it to us:

> **CV = σ ÷ μ** — the standard deviation of a measured score, divided by its mean, read as a percent.

A low coefficient of variation is a plant that lands on its number; a high one is a plant you cannot build on, however beautiful its average. Our line — across all eight axes and every one of the forty-two organisms in our design catalog — holds a batch CV *at or under seven percent.* Not because seven is a magic number, but because it sits below the threshold where a mouth, even a trained one, can tell one lot from the last. The rest of this chapter earns that line.

What EC₅₀ actually measures — the shape of a dose–response curve

Before a number can be held, it has to mean something exact, and the meaning lives in a curve. Put a receptor in front of rising concentrations of the molecule that speaks to it and plot the response, and you do not get a straight line. You get an S — flat and silent at low dose, rising steeply through a middle band, then flattening again as every channel that can open has opened. That S is the dose–response curve, and its algebra is the Hill equation:

> $R = R_{max} \times [A]^h \div (EC_{50}^h + [A]^h)$

Here **R** is the response, **R_max** the ceiling when the receptor is saturated, **[A]** the concentration of the agonist, **EC_{50}** the concentration that produces exactly half of R_max, and **h** the Hill coefficient — the steepness, set by how cooperatively the channel's subunits open. For capsaicin at TRPV1, the trigeminal heat channel, h sits near 2: the subunits help each other, so the curve is steep.

Set **[A] = EC_{50}** and the equation collapses to R = R_max ÷ 2, by construction. That is the whole reason we anchor the index at the half-max point and nowhere else. Walk the curve with h = 2 and the logic becomes physical: at a quarter-dose and a half-dose the receptor is nearly mute; at **0.5 × EC_{50}** the response is R_max × 0.25 ÷ 1.25 = **0.20**; at **EC_{50}** it is **0.50**; at **2 × EC_{50}**, R_max × 4 ÷ 5 = **0.80**; at **4 × EC_{50}** it is **0.94** and climbing toward a ceiling it will never quite reach. A fourfold swing in concentration around the midpoint — from half to double — moves the receptor from one-fifth to four-fifths of full voice. That steepness is the gift: the midpoint is where the curve changes fastest, so it is the one point you can locate with the least error. Threshold and saturation are flat, and on a flat stretch a large change in concentration barely moves the response, which means reading concentration *back* from response there carries enormous uncertainty. The half-activation point is the best-conditioned place on the whole curve. We do not anchor at EC_{50} because it is convenient. We anchor there because it is the one spot where the biology answers a question with the smallest shrug. And the steeper the curve — the higher the Hill coefficient — the more sharply that midpoint is defined; capsaicin's cooperative h near 2 is part of why heat is among the cleanest axes we pin.



The receptor assay where a candidate's EC_{50} is read — rising doses, half-maximal response, the single honest concentration that becomes a score.

The ratio that cancels the room

A raw EC_{50} is a fragile thing. It drifts with the cell line's passage number, the day's temperature, the lot of reagent, the particular plate. Report it bare and you are reporting the room as much as the molecule. The index refuses to do that. It never reports an absolute; it reports a *ratio* to a fixed reference run in the same assay, on the same day, beside the candidate:

> **IFI(candidate) = 1,000 × EC_{50}(reference) ÷ EC_{50}(candidate)**

The scale is pinned so that the native wild-type reference reads exactly **1,000** — because when the candidate *is* the reference, EC_{50}(reference) ÷ EC_{50}(candidate) = 1, and 1,000 × 1 = 1,000.

Everything above 1,000 speaks louder than the wild type; everything below, softer. Now watch what the ratio does to a bad day. Suppose the assay carries some multiplicative bias *b* — the room is a degree warm, the cells a passage old — and it scales *both* EC_{50} readings by the same factor. Then:

> $IFI = 1{,}000 \times (b \cdot EC_{50}(ref)) \div (b \cdot EC_{50}(cand)) = 1{,}000 \times EC_{50}(ref) \div EC_{50}(cand)$

The *b* cancels. The bias divides out of the number entirely. This is the deepest reason a score measured in one city in August equals one measured in another in February: the reference rides along in every run and quietly subtracts the room. The index is not an absolute measurement that we hope travels well — it is a same-day ratio engineered so that whatever is wrong with today is wrong with the reference too, and cancels.

The ratio cancels the day's *bias*, but it is honest about the day's *noise*: random scatter in the two readings does not divide out — it compounds. For a quotient of independent quantities, relative variances add, so the score inherits **$CV_{IFI} = \sqrt{(CV_{ref}^2 + CV_{cand}^2)}$** from the bench. Read each EC_{50} once at the assay's 3.5% and the number would carry $\sqrt{(3.5^2 + 3.5^2)}$ = **4.95%**; hold each to 2.5% with replicate wells and the score inherits $\sqrt{(2.5^2 + 2.5^2)}$ = **3.5%** instead — the assay door's whole budget, bought back one well at a time. The reference does not only cancel the room; replicated, it steadies the hand that reads it.

To make the arithmetic concrete, fix the reference. For every worked figure in this chapter we take the wild-type anchor to half-activate TRPV1 near **$EC_{50}(ref) = 0.70\ \mu M = 700\ nM$** — the value is the assay's own, and what matters is never its absolute size but the ratio it forms. Invert the formula, $EC_{50}(candidate) = 1{,}000 \times EC_{50}(ref) \div IFI$, and the heat ladder resolves into real concentrations. Phoenix at IFI-H 4,200 half-activates the channel at 700,000 ÷ 4,200 = **166.7 nM** — it needs barely a sixth of the reference concentration to drive the receptor to half, which is exactly what "4.2 times as loud" means in molecules.

CULTIVAR	SKU	PRED. SHU	IFI-H	SHU / IFI-H PT	EC_{50} (NM)*
Ember	SS-0087	866K	1,570	552	446
Cinder	SS-0088	1.05M	2,290	459	306
Kindle	SS-0089	1.25M	2,000	625	350
Flare	SS-0090	1.48M	—	—	–
Blaze	SS-0091	1.72M	2,630	654	266
Wildfire	SS-0092	1.96M	2,930	669	239
Scorch	SS-0093	2.21M	3,220	686	217
Firestorm	SS-0094	2.47M	3,490	708	201
Solara	SS-0095	2.72M	3,980	683	176
Phoenix	**SS-0096**	**2.95M**	**4,200**	**700**	**167**

*EC_{50} shown against the illustrative 700 nM reference anchor; the index value is the ratio, not the absolute. All SHU values are computational predictions from our model, not certified laboratory measurements. Flare has a predicted SHU but its receptor loudness has not been modeled on this axis.

SHU is a chemistry number; IFI is a biology number

Run your eye down the two middle columns of that ladder and something quietly important appears. The Scoville figure and the IFI-H figure climb together but they do not climb in step. The ratio of SHU to IFI-H point drifts from **552** at Ember up through **654** at Blaze to **708** at Firestorm — a spread of better than a quarter across the same family. And the cleanest tell sits near the top of the softer end: our model predicts Kindle carries **more** Scoville heat than Cinder (predicted 1.25M against 1.05M SHU) yet registers a *lower* IFI-H (2,000 against 2,290). By predicted capsaicinoid mass it is hotter; by loudness to the receptor it is quieter.

That is not an error. It is the difference between the two numbers made visible. SHU is a chemistry measurement — a count, by dilution, of total capsaicinoid present. IFI-H is a biology measurement — how forcefully that particular blend engages TRPV1. What the tongue reports is never the milligrams; it is the receptor's answer. Capsaicin and its near-twin dihydrocapsaicin, and the minor capsaicinoids beside them, do not all speak to the channel with the same voice, and their balance is set at the Pun1/AT3 acyltransferase — the terminal capsaicin synthase — where vanillylamine condenses with a branched-chain fatty acid. The length and saturation of that acyl chain tune the proportion of each capsaicinoid, and so the *shape* of the burn — its onset, its climb, its loudness per milligram. Two fruits can hold the same mass of heat and spend it differently. This is precisely why a quality system that only controlled total capsaicinoid would still let the flavor wander: you have to hold the receptor number, because the receptor number is the one the eater actually hears.

CV = σ ÷ μ — the honest measure of wander

Now to the holding. Give a specification a target mean μ and the question of consistency becomes the question of σ, the standard deviation around it, and the single figure that makes σ legible across every axis and every scale is the coefficient of variation, CV = σ ÷ μ. It is dimensionless and it is honest: a σ of 294 means one thing on a number near 4,000 and something very different on a number near 800, but a CV of seven percent means the same looseness everywhere.

Put Phoenix's heat target through it. With μ = 4,200 held at **CV = 7%**, σ = 0.07 × 4,200 = **294**. One standard deviation is 4,200 ± 294, the band [3,906, 4,494]; two sigma, inside which a well-behaved process spends about nineteen lots in twenty, is ±588, the band [3,612, 4,788]. That is tight enough that a trained panel reaches for the two lots and finds the same pepper.

Now set that honest figure beside the thing we started from. A wild-type line delivering the same mean carries a CV near **42%**. Its σ is 0.42 × 4,200 = 1,764, and a single standard deviation already spans [2,436, 5,964] — a band wider than the entire engineered ladder from Blaze past Phoenix and off the top of it. One wild harvest, read

honestly, would blur across half the named cultivars. The gap between 42% and 7% is the whole distance between a plant you can admire and a plant you can build on.

One more move earns the number real money. Individual fruits wander more than a lot does, because a lot is a *mean* of fruits, and the standard error of a mean of n independent samples falls as the square root of n. Pool and composite twenty-five fruits into a lot and the lot's CV is 7% ÷ √25 = 7 ÷ 5 = **1.4%**; nine fruits gives 7 ÷ 3 = **2.3%**. We do not promise that every single pepper is within seven percent of the target — we promise it of the lot, and the certificate reports the lot, pooled, which is where √n has already done its quiet work.

The four doors

Seven percent is not found; it is budgeted. Variance walks into a living process through a small number of doors, and the discipline is to know every door, assign each a share of the total, and close each to its share. For independent sources, variances add — which means coefficients of variation add in quadrature, as the root of the sum of squares:

```
CV_total² = CV_seed² + CV_room² + CV_clock² + CV_assay²
```

Four doors, four sources, one budget. **The seed** is genetic wander — plant-to-plant differences in genotype. We close it by writing the two designed Pun1 edits — **S39L · L345G** — into a single fixed genotype and propagating isogenic lines, so that every plant in a lot is, as nearly as biology allows, the same plant; the door that gapes widest in a segregating wild population is nailed nearly shut. **The room** is environmental wander — light, heat, water, feed. **The clock** is developmental wander — how ripe a fruit was at the moment it was picked, which moves a capsaicinoid number more than almost anything else; the fruit's heat load climbs steeply through ripening, so a few days' drift in picking shifts the number more than a degree in the room ever will, and we harvest to a fixed physiological age read from the plant, never to a date on a calendar. **The assay** is analytical wander — the repeatability of the measurement itself, held down by the same-day reference ratio, replicate wells, and pooled sampling.



The controlled chamber — the room where dawn is pinned: photoperiod, temperature, vapor-pressure deficit, spectrum and feed all held constant, so the environmental door barely opens.

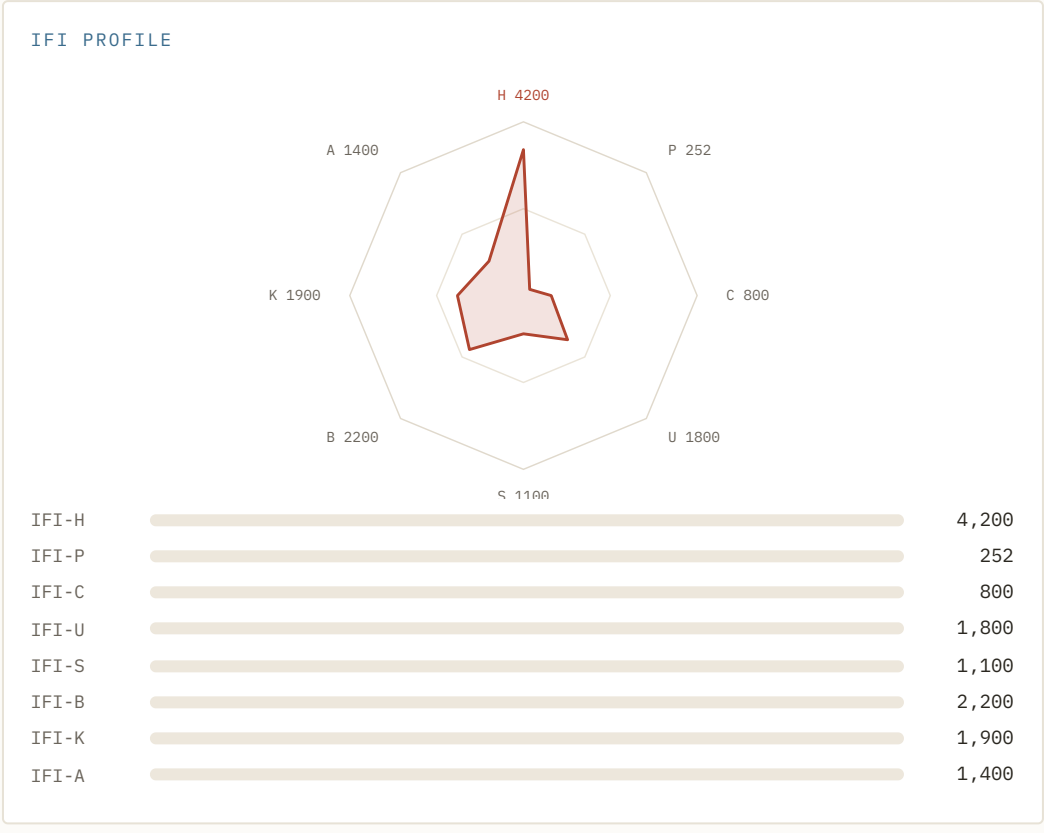
Because the sources add in quadrature, no single door may be allowed to dominate; the largest one sets the floor no matter how clean the other three are. A workable budget looks like the table below, and its quadrature sum lands under the line with margin to spare: $\sqrt{(1.5^2 + 3.0^2 + 4.0^2 + 3.5^2)} = \sqrt{39.5} = \mathbf{6.28\%}$.

DOOR	SOURCE OF VARIANCE	HOW IT IS CLOSED	HELD-TO CV
The seed	genetic (plant-to-plant genotype)	isogenic lines, CRISPR-fixed edits (S39L · L345G)	1.5%
The room	environmental (light · heat · water · feed)	controlled environment; pinned photoperiod and vapor-pressure deficit	3.0%
The clock	developmental (ripeness at harvest)	harvest at a fixed physiological age, not a calendar date	4.0%
The assay	analytical (measurement repeatability)	same-day reference ratio · replicate wells · pooled lot	3.5%
Total	**(independent, in quadrature)**	**$\sqrt{(1.5^2 + 3.0^2 + 4.0^2 + 3.5^2)}$**	**6.28%**

The "pinned dawn" is the room door made literal. In a controlled environment the photoperiod does not drift with the season — dawn arrives at the same hour, at the same spectrum, every day of the plant's life — and temperature, vapor-pressure deficit, and nutrient feed are held on the same discipline. A plant grown under a pinned dawn is answering the same question every morning, and a plant that is always asked the same question tends to give the same answer. That is the room budget, 3.0%, bought with light.

Holding eight numbers at once

The hardest specification we keep is the one that asks for all of it at once. Most organisms we tune speak on two or three axes; Phoenix is the single cultivar that registers on every one, the full-spectrum fruit designed to carry heat and sweetness and savor and the cool sharp edges all together.



Each axis score is an IFI value anchored to its own reference convention — a stated constant that pins IFI = 1,000 on that axis, the same way the Scoville scale is defined by dilution-to-threshold.

Values on different axes reflect different defined reference points and are not directly comparable on a single common scale.

Eight targets, eight receptors, and every one of them must land. Against the scorecard ceiling of 5,000, that is heat at 84% of full scale, zero-bitter at 44%, kokumi at 38%, cool-burn at just 16% — eight numbers at eight different heights, each with its own door-budget, each to be held to CV ≤ 7% at the same time, in the same fruit. Hold each axis to the two-sigma release windows below and you have the real specification Phoenix ships against.

AXIS	RECEPTOR	TARGET IFI	Σ (7%)	RELEASE WINDOW (±2Σ)
H · Heat	TRPV1	4,200	294	3,612 – 4,788
P · Pungency	TRPA1	252	18	216 – 288
C · Cool-Burn	TRPM8	800	56	688 – 912
U · Umami	mGluR1	1,800	126	1,548 – 2,052
S · Sweetness	TAS1R2·R3	1,100	77	946 – 1,254
B · Zero-Bitter	TAS2R38	2,200	154	1,892 – 2,508
K · Kokumi	CaSR	1,900	133	1,634 – 2,166
A · Astringency	Salivary PRPs	1,400	98	1,204 – 1,596

Here the mathematics turns against us, and it is worth naming honestly. A lot passes only if *every* axis lands inside its window, and probabilities of independent events multiply. Suppose each axis independently falls inside its two-sigma window with probability 0.95 — a good single number. The chance that all eight land together is not 0.95 but $\mathbf{0.95^8 = 0.663}$. Hold eight honest axes to a one-in-twenty miss rate each and a third of your lots still fail the joint release. To carry a joint pass rate back up to 0.95, each axis must be held to 0.95^(1/8) = **0.9936** — better than a one-in-a-hundred-fifty miss, per axis, on all eight at once. That is the combinatorial tax on ambition, and it does not go away; it is paid in tighter per-axis CVs and wider, carefully-earned windows. Tighten each axis to a one-in-a-hundred miss and the joint pass climbs to 0.99^8 = **0.923** — and there is no cheaper route up than lowering the per-axis variance itself, which is why the eight-axis fruit consumes more of the quality system than any single-axis line. The full-spectrum fruit is not eight times as hard as a single-axis one. It is harder than that, and the exponent is why.

The certificate, and the honest limits



Lot sampling at the bench — the pooled fruits, the same-day reference, the assay whose result becomes a line on the certificate that travels with the box.

All of it converges on a single sheet of paper that rides with every lot — the certificate of analysis, the document that turns an internal discipline into a promise a stranger can hold us to. It carries the lot's pooled mean on each axis, the sample count n behind that mean, the measured CV, the same-day reference the ratio was taken against, and the release window each axis had to fall inside. A lot that misses a window does not ship; it is not quietly rounded into tolerance. The certificate is how the metronome reaches the formulator: not *trust us*, but *here is the number, here is how tightly it was held, here is the reference it was weighed against.* It is the specification made auditable.

And because numbers are a form of respect, we are exact about what ours cannot do. The index measures receptor loudness in a dish; perception in a mouth integrates many receptors at once, with saliva, temperature, fat, and the food's whole matrix bending the result — so a profile predicts the eating faithfully but never wholly determines it. The ≤7% line is a *lot* specification, earned by pooling, not a vow that every individual fruit sits within seven percent; single fruits wander more, and we say so, which is why we composite and report the lot. The EC_{50} figures worked through this chapter stand against an illustrative reference anchor; the index value is the ratio, and only the ratio, because the absolute is the part that belongs to the room. And seven percent itself is not sacred — we chose it because it sits just under the smallest difference a trained panel can reliably taste, so that holding beneath it makes a thousand fruits read, to the one receptor each axis is tuned to, as one fruit repeated.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

That is the entire argument of a quality system, and it is a quiet one. A plant grown clean, under a pinned dawn, from a fixed seed, picked at a fixed ripeness, measured against a reference that cancels the day, pooled into a lot, and certified before it leaves — so that the stranger a year from now, in a city we will never see, tastes exactly what we meant, and never has to wonder whether tonight is the night it is off. The math is only how we keep the promise. The promise is the point.



PART I OF V

The Measure of Drift

CV = σ/μ · the seven-percent line and the mouth that sets it · four doors and the quadrature rule · the dose-response curve

Chapter 1 · The coefficient of variation — *the honest measure of drift.*

Taste the same dish twice and you will never meet two identical spoonfuls. The second is always a hair warmer or a shade sweeter than the first, because the world is granular and a tongue is a generous instrument. The only real question a maker ever faces is not whether two lots differ — they always do — but *how far* they are allowed to drift before the difference crosses from physics into perception, from a fact about molecules into something a person actually notices and minds. The coefficient of variation answers that question without ceremony, and it is the quietest, most useful number in this whole book.

We did not start loving it in a spreadsheet. We started loving it in the chamber, standing in front of a tray of fruit that had been grown to the same specification, one seed's worth of instruction written into every cell, and watching the readings come back not identical but *close* — close in a way we could finally put a number on. That number is what this chapter is about, and everything we build on top of it in these pages is only ever an effort to make it smaller.



The controlled chamber, where drift is fought one variable at a time — light, water, nutrient, and the honest record of what the fruit actually did.

Two pieces: the mean and the spread

You build the coefficient of variation from two pieces, and it is worth being slow about both, because a number inherited without its derivation is a number you cannot defend.

Measure a score across many fruit and you have a cloud of values — call them $x_1, x_2, \ldots x_n$. Their center is the **mean, μ**: add them and divide by how many there are,

$$\mu = (x_1 + x_2 + \cdots + x_n) / n.$$

This is the figure you would quote — the one number that stands in for the whole lot. But the mean is a liar if it travels alone, because two lots can share a center and share nothing else: one tight as a drumhead, one scattered across the room. So we measure the width.

For each fruit, take its distance from the center, $(x_i - \mu)$ — its deviation. Some are positive, some negative, and if you simply averaged them they would cancel to zero by construction; the mean is, after all, the balance point. So we square each deviation first, which makes every one positive and, not incidentally, punishes a far-flung fruit far more than a near one. Average those squares and you have the **variance**, σ^2. Take the square root to return to the original units — points on the Index, not points-squared — and you have the **standard deviation, σ**:

$$\sigma = \sqrt{\left[\Sigma(x_i - \mu)^2 / n \right]}.$$

That is why σ is called the *root-mean-square* distance of the cloud from its own center: root, of the mean, of the squares. It is the honest width of the scatter, reported in the same currency as the score itself.

One subtlety separates the careful lab from the hopeful one. Dividing the sum of squares by *n* gives the width of *this* lot, exactly — a complete, closed description of the fruit in front of you. But most of the time we are not asking about this lot; we are asking what the *process* will do next, estimating the spread of all the lots this cultivar will ever throw. For that inference, honesty requires dividing by $n - 1$ rather than n — the Bessel correction — because the sample has already spent one degree of freedom locating its own mean, and pretending otherwise flatters the process by making it look tighter than it is. The two forms are close for a well-sampled lot and diverge for a thin one. We report the *n* form when we describe a lot and the $n - 1$ form when we promise what the next lot will do, and we never quietly swap one for the other to win an argument.

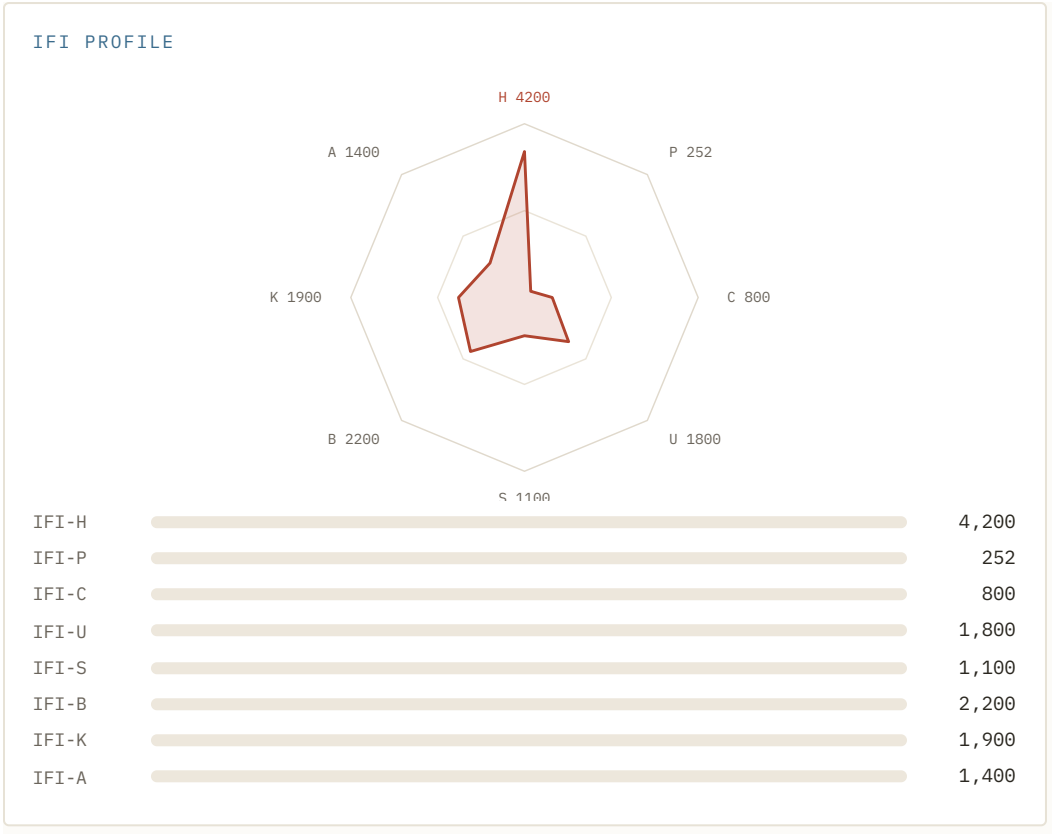
The ratio, and why it must be dimensionless

Now the move that makes the whole quality system possible. Divide the spread by the center and scale to a percent:

> **CV = σ / μ**, reported as a percentage.

Its great virtue — the reason we built the entire discipline on it rather than on any absolute tolerance — is that it is *dimensionless*. It carries no units; it is a pure ratio of a spread to a size. And that is precisely what lets a single standard govern eight wildly different axes at once.

Consider what we are asking one rule to cover. Heat towers near the top of the Index; a delicate cooling finish may sit at a fraction of it; a whisper of astringency somewhere between. The full-spectrum cultivar we designed to register on every axis at once, **Phoenix**, carries all eight at their specified heights simultaneously, which is exactly why it is the hardest thing we make to hold steady — eight independent promises, every lot:



An absolute tolerance — say, "hold every axis to within ±50 points" — would be invisible on the tall axes and ruinous on the short ones, demanding of one a precision many times finer than it asks of another. Watch it fail in a single column:

AXIS	PHOENIX TARGET	±50 POINTS MEANS	EFFECTIVE STRICTNESS
H · Heat	4,200	±1.19 %	loosest
B · Zero-Bitter	2,200	±2.27 %	
K · Kokumi	1,900	±2.63 %	
U · Umami	1,800	±2.78 %	
A · Astringency	1,400	±3.57 %	
S · Sweetness	1,100	±4.55 %	
C · Cool-Burn	800	±6.25 %	
P · Pungency	252	±19.8 %	strictest

The same fifty points is a 1.19 % whisper on heat and a nearly 20 % demand on pungency — a nearly seventeen-fold swing in how strict the "one" rule really is, hidden inside a number that *looks* even-handed. The coefficient of variation refuses the trick. Seven percent of a lot is seven percent of a lot, whether the lot is a wall of capsaicin or a breath of menthol. One line, written once, governs the entire library — which is exactly the kind of single shared standard every surface of this work is built on.

A lot, measured

It pays to see, concretely, what a CV *is* before we spend the rest of this book driving it down. Here is one umami lot — tomato, the fruit we have spent the most love restoring, grown back from the hollow IFI-U 300 the long-haul supply chain left it at toward the savory depth a tomato is supposed to carry. The specification is IFI-U 2,600. We pulled eight representative fruit and read each one.



Tomato, measured. The savory depth the truck took, grown back in — and now held to a number.

FRUIT	IFI-U READING	DEVIATION ($X_i - M$)	SQUARED
1	2,450	−140	19,600
2	2,490	−100	10,000
3	2,495	−95	9,025
4	2,520	−70	4,900
5	2,660	+70	4,900
6	2,685	+95	9,025
7	2,690	+100	10,000
8	2,730	+140	19,600
Σ	**20,720**	**0**	**87,050**

Walk it through, because the arithmetic is the argument. The mean is 20,720 ÷ 8 = **2,590** — ten points under the target, a fact worth noting and, as we will see, well inside the noise. The deviations sum to zero, as they must; the balance point balances. The squared deviations sum to 87,050, and dividing by n = 8 gives a variance of 10,881 points-squared. The square root returns us to the tongue's own units:

> $\sigma = \sqrt{10{,}881} \approx$ **104**, and CV = 104 / 2,590 = **4.0 %**.

(Were we estimating the *process* rather than describing this lot, the $n - 1$ form would give σ ≈ 111 and CV ≈ 4.3 % — the honest, slightly larger number we would quote when promising what the *next* tomato will do.)

Four percent. Read it back into the world and it says: the typical fruit in this lot sits within about a hundred points of the target, and a hundred points at twenty-six hundred is a difference no mouth will ever find. The lot is not *approximately* IFI-U 2,600; for every practical purpose that matters at a table, it *is* IFI-U 2,600. That is the whole aim of what we build here — to make that sentence true not by luck, not once, but on demand, for every axis, in every lot, forever.

From the reading to the score: EC50 and the Index

Where does a "reading" of 2,690 actually come from? Not from how *much* glutamate a tomato holds, but from how *loudly* it speaks to the receptor — and that distinction is the engine under every number on

the Index. The umami axis is anchored to mGluR1, the receptor the tongue uses to hear savory depth. Expose it to a rising concentration of a stimulus and its response climbs along a sigmoid, the dose–response curve, well described by the Hill equation:

> response = $R_{max} \cdot [A]^n / (EC_{50}^n + [A]^n)$.

Here [A] is the concentration offered, R_{max} the receptor's full-throated maximum, *n* the Hill coefficient that sets the steepness of the climb, and **EC_{50}** the single concentration at which the response reaches half of maximum. EC_{50} is the whole point: it is the stimulus's loudness distilled to one number. At [A] = EC_{50} the response is exactly $R_{max}/2$, by definition. With a simple *n* = 1 curve, doubling to [A] = 2·EC_{50} lifts you only to two-thirds of maximum, and halving to [A] = ½·EC_{50} drops you to one-third — a deliberately gentle slope. Steepen it to *n* = 2 and the same two steps swing you to four-fifths and one-fifth instead. That steepness matters for *precision*: the sharper the curve through its midpoint, the more tightly a given spacing of test concentrations pins EC_{50} down, and the smaller the measurement scatter we inherit before a single tomato has even varied.

The Index turns that loudness into a score everyone can read on one page by inverting EC_{50} against a fixed reference:

> **IFI = 1,000 × EC_{50}(reference) / EC_{50}(candidate).**

A lower EC_{50} means the candidate reaches half-max at a lower concentration — it speaks louder — and so scores *higher*. The reference is chosen so that a candidate matching it lands exactly at 1,000. Work it through with round numbers. Take a reference whose half-maximal concentration we fix, for this illustration, at 50 µM; a candidate that also needed 50 µM would score 1,000 × 50 / 50 = 1,000, as designed. Our restored tomato reaches half-max at 19.2 µM. Then

> IFI-U = 1,000 × 50 / 19.2 = 50,000 / 19.2 = **2,604**,

which we report as 2,600. The score is not a metaphor for savoriness; it is arithmetic performed on a concentration.

And because a concentration is a physical quantity, it converts. That 19.2 µM is 19.2 × 10^{-6} moles per litre; multiply by the molar mass of L-glutamic acid, 147.13 g/mol, and the receptor is answering at roughly 19.2 × 10^{-6} × 147.13 ≈ **2.82 mg per litre** at the tongue — a few parts per million doing the work of a Sunday sauce. We find that quietly moving: the depth a tomato gives the body is carried by a vanishingly small, clean amount of a molecule the plant makes itself.

Now the piece that ties this section back to the first. If every reported score is really an EC_{50} measurement run through that formula, how does the *scatter* in the reading relate to the scatter in the measurement? Write the Index as IFI = *k* / EC_{50}, where *k* = 1,000 × EC_{50}(reference) is a fixed constant with no error of its own. Propagate the uncertainty the ordinary way — for a function *y = f(x)*,

the spread $\sigma_y \approx |f'(x)| \cdot \sigma_x$. With $f = k/x$, the derivative is $f' = -k/x^2$, so $\sigma_{IFI} = (k/EC_{50}^2) \cdot \sigma_{EC_{50}}$. Divide by IFI itself to form the coefficient of variation and the constant k cancels, as does one power of EC_{50}:

> $CV(IFI) = \sigma_{IFI} / IFI = \sigma_{EC_{50}} / EC_{50} = \mathbf{CV(EC_{50})}.$

The relative scatter of the score *equals* the relative scatter of the underlying concentration measurement. Inversion against a fixed reference preserves the coefficient of variation exactly, to first order. This is not a convenience; it is a permission. It means every discipline we impose on the wet-bench assay — every steadying of the EC_{50} read — flows through to the published score at one-to-one, and the dimensionless number we track on the tongue is the very same number the pipette is fighting.

The ruler and the plant: two kinds of scatter

Here is the error that quietly undoes most quality programs, and the reason a coefficient of variation is only ever as honest as the measurement beneath it. The scatter you read is never purely the wander of the fruit. It is the fruit's true biological variation *and* the wobble of your own instrument, braided together — and if you blame the plant for the pipette's sins, you will re-engineer a cultivar to fix a fault in your lab.



The stage where the reading is made. Before we trust a number about the plant, we measure what the instrument does to a sample that cannot vary.

The two sources add, but they add in quadrature — as squares, the way independent variances always combine:

> $\sigma^2(observed) = \sigma^2(biological) + \sigma^2(measurement).$

Which means they can be separated, if you are willing to do the work. Take one fruit, homogenize it so there is nothing left to biologically vary, and read that single sample again and again. Whatever spread remains is the instrument confessing its own noise — the measurement standard deviation, $\sigma(measurement)$. Suppose, for our tomato lot, that confession comes back at $\sigma(measurement) = 40$ points. We observed $\sigma = 104$ across the eight fruit. Then the fruit's *true* scatter is not 104 at all:

$$\sigma(\text{biological}) = \sqrt{(104^2 - 40^2)} = \sqrt{(10{,}816 - 1{,}600)} = \sqrt{9{,}216} = \mathbf{96}.$$

The plant is tighter than it looked. Its honest biological CV is 96 / 2,590 = **3.7 %**, while the instrument was contributing 40 / 2,590 = 1.5 % of apparent drift that had nothing to do with any tomato. Fold the two back together in quadrature — $\sqrt{(96^2 + 40^2)} = \sqrt{10{,}816} = 104$ — and the books balance. Report 4.0 % as the biological reality and you have slandered the cultivar by a third of its variance; chase that phantom with new edits and you will spend a season tuning a plant to compensate for a loose seal on a cuvette. We measure the ruler before we measure with it. It is the least romantic discipline in this book and one of the most important.

The square root of n: how many fruit to read

One more relationship earns a formulator's trust, because it governs how hard the question even is. The numbers above describe *individual* fruit. But a lot is rarely sold as a single tomato; it is pooled, pressed, blended — and when we report the *mean* of many fruit, that mean is far steadier than any one of them. Its spread is the standard error,

$$\mathrm{SEM} = \sigma / \sqrt{n},$$

and so the coefficient of variation of the lot's *average* falls as one over the square root of the count. Reading more fruit cannot change the plant, but it sharpens our estimate of where its center truly sits:

FRUIT READ, N	SEM = Σ/√N	CV OF THE LOT MEAN
1	104.0	4.0 %
4	52.0	2.0 %
8	36.8	1.4 %
16	26.0	1.0 %
30	19.0	0.7 %

The curve of diminishing returns is written right into the √n: going from one fruit to four halves the uncertainty, but halving it *again* costs twelve more fruit, not four. This is the arithmetic behind every sampling plan we run — enough fruit to pin the mean to a tolerance we can stand behind, never so many that we are grinding the back half of the orchard to shave a tenth of a percent off a number no palate could resolve. Precision is a budget, and the square root of n tells you the price of each increment.

What ≤ 7 % buys, and what 42 % costs

All of this exists to make a promise keepable. The specification we hold the platform to is a batch-to-batch coefficient of variation of **≤ 7 %**, on every axis a cultivar expresses. To feel what that is worth, set it beside the thing it replaced. Left to the old way — select the phenotype, hope the genotype followed, accept whatever the weather and the open field delivered — capsaicin output wanders at a batch CV near **42 %**. Not a rounding error; a different plant from one harvest to the next.

Carry both numbers onto the heat ladder, where the stakes are loudest, and express the ±1σ window directly in Scoville Heat Units:

CULTIVAR	SKU	PREDICTED PEAK SHU	IFI-H	±1Σ AT 7 % CV
Ember	SS-0087	866,000	1,570	± 60,620
Blaze	SS-0091	1,720,000	2,630	± 120,400
Firestorm	SS-0094	2,470,000	3,490	± 172,900
Phoenix	SS-0096	2,937,948	4,200	± 205,656

At the top of the ladder, Phoenix's engineered window is about ±206,000 SHU around our model's predicted peak of approximately 2.94 million SHU — tight enough that a sauce formulator can dose to a target and trust the next drum to land where the last one did. Now let the discipline lapse and let that same cultivar drift at the wild-type 42 %: the window blows out to ±0.42 × 2,937,948 ≈ **±1,233,938 SHU**. More than a million Scoville of uncertainty on a single promise — a "superhot" that might arrive merely hot or might arrive as something nobody consented to eat. The coefficient of variation is the entire difference between a flavor you can design around and a rumor. Six-fold tighter relative scatter, 42 % down to 7 %, is not a statistic on a slide; it is the reason a number on a label means anything at all.

Below the threshold of notice

Why does a dimensionless measure map so cleanly onto what a person actually tastes? Because perception itself is dimensionless. Weber's law, the oldest result in psychophysics, holds that the smallest change a sense can detect is not a fixed amount but a fixed *fraction* of what is already present: ΔI / I ≈ k, the Weber fraction, roughly constant across the usable range. Add a candle to one already burning and the room leaps brighter; add it to a hundred and no one looks up. The tongue keeps the same law — the just-noticeable difference in taste intensity is commonly found to sit on the order of 15–20 % of the reference.

Lay our numbers against that threshold and the point of the whole discipline arrives. A biological CV of 3.7 %, a measured lot CV of 4.0 %, a platform specification of ≤ 7 % — every one of them falls comfortably *below* the fraction at which a difference becomes perceptible at all. Make the coverage concrete: for a roughly normal lot at μ = 2,590 and σ = 104, about two-thirds of the fruit land within ±104 of target (2,486 to 2,694), and about nineteen in twenty within ±208 (2,382 to 2,798) — and even that outer window is itself a small fraction of any just-noticeable difference. The drift is real; it is simply beneath the floor of notice. This is the quiet triumph of holding a tight CV: not that the fruit stops varying — it never will — but that its variation is driven under the threshold where physics would otherwise become perception, so that the difference between two tomatoes, true in every molecule, never once arrives at a tongue.

Honest limits

We would be breaking faith with the measure if we did not name where it fails, because a tool trusted past its range is as dangerous as no tool at all.

The coefficient of variation needs a mean that is meaningfully far from zero. On an axis a cultivar does not express — a tomato's heat, say, sitting near the floor of the Index — μ approaches zero and CV

= σ/μ bloats toward infinity, reporting catastrophe where there is only silence. For those axes we quote the absolute standard deviation in points and do not embarrass ourselves with a ratio. CV is the language of axes a plant actually speaks.

It also assumes the scatter is roughly symmetric and the scale is a true ratio scale with a real zero — both honored by the Index, both worth checking rather than assuming. A badly skewed lot, or one contaminated by a few wild outliers, will inflate σ and hand you a pessimistic CV that describes the tail rather than the body; there, a robust width is the more honest ruler. And CV is a single number, so it is blind to *shape*. Two lots can post an identical 4.0 % while one is a clean single cloud and the other is secretly two populations — an early picking and a late one, pooled — stacked into a mean that neither sub-lot ever actually visits. A coefficient of variation will not warn you of that bimodality; only looking at the distribution will. We look.

And the first caution is still the last one: a CV is only as truthful as the measurement under it. A beautiful 7 % computed on a drifting instrument is not precision — it is a well-formatted lie, and it fails in the dangerous direction, telling you the fruit is unreliable when in truth your ruler is, or worse, telling you all is well while the assay quietly rots. Separate the ruler from the plant, keep the mean honest, watch the shape, and respect the floor — do those four things, and the coefficient of variation becomes what it is meant to be: the single number that turns "it tastes about right" into something we can engineer against. Everything else is how we drive it down.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

Chapter 2 · Seven percent — *the number, and the mouth that sets it.*

Every quality line has to be drawn somewhere, and an arbitrary line is a weak one — the first hard season, the first expensive lot, someone will ask why not eight, why not ten, and an answer of "it felt right" will not hold. So our line is not drawn by us at all. It is drawn by the mouth. Seven percent is where it is because that is roughly where human taste stops being able to tell the difference, and a consistency you can perceive is the only kind that matters.

A number that honest has to earn itself three times over: in the physics of perception that sets the floor, in the arithmetic of the index that carries the score, and in the statistics of a tasting that decides, cup against cup, whether two lots are one food or two. We walk all three here, because a formulator who holds seven percent should be able to *derive* it, not merely obey it — and because the discipline we ask of a batch is the same discipline we ask of ourselves at our own table: say exactly how much, and mean it.



A single cultivar held in the controlled chamber — identical light, identical air, so the only variable left is the one we mean to measure.

The law beneath the line — *a logarithm hidden in the tongue.*

The physics of the floor is old and well named. In the nineteenth century Ernst Weber noticed that the smallest change a sense can detect is not a fixed amount but a fixed *fraction* of the starting stimulus — lift a weight and you must add proportionally more to a heavy one than a light one before your hand reports a change. Written down, it is almost insultingly simple:

> **ΔI / I = k** — the just-noticeable difference is a constant fraction *k* of the stimulus, the Weber fraction.

That one line has a consequence most people never unpack. If every just-noticeable step is a fixed *fraction* of where you already are, then sensation does not track the stimulus — it tracks the *logarithm* of the stimulus. Sum up the barely-perceptible steps from a threshold intensity I_0 to some intensity I, each step worth $\Delta I = k \cdot I$, and the running total of felt sensation integrates to a logarithm:

> **$S = (1 / k) \cdot \ln(I / I_0)$** — Fechner's law; sensation is the log of the stimulus, scaled by the inverse Weber fraction.

This is why a flavor scale has to be built the way the Index is built. Double the capsaicin and you do not double the fire; you add one fixed increment of felt heat. The tongue is a logarithmic instrument, so a *linear* concentration scale would crowd every interesting distinction into the bottom of the range and waste the top. Anchoring each axis to the half-activation of a receptor — a point defined by a ratio, not an absolute dose — puts the score where the perception actually is: in the land of ratios and logs. A twentieth-century refinement, Stevens' power law, replaces the clean logarithm with $\psi = I^a$ and lets the exponent *a* vary by sense; for most taste intensities *a* is below one (the response is *compressive*, flattening as you climb), while capsaicin heat near its threshold can run *expansive*, climbing faster than the dose. Both pictures agree on the thing that matters to a spec: at ordinary use levels, a change has to clear a stubborn fractional threshold before anyone feels it at all.

Taste obeys Weber, and taste obeys it *loosely.* For most axes the Weber fraction for intensity runs well up into the teens and twenties of percent: a casual taster needs a stimulus to change by something like a fifth before they will reliably call it different. A trained panel, tasting with intent in controlled conditions, sits lower — but still, for intensity judgments at normal use levels, north of our line. The table below is not a claim about any one person's palate; the fractions are representative figures from the psychophysics literature, held deliberately conservative, the set of perceptual tolerances we design *beneath*, axis by axis, so the line holds even for the most discerning mouth we can imagine meeting our food.

AXIS · RECEPTOR	WORKING WEBER FRACTION *K* (DESIGN MARGIN)	FELT CHARACTER OF THE SIGNAL
IFI-H · TRPV1	~ 0.22	loud, robust, forgiving — hard to shift
IFI-S · TAS1R2·R3	~ 0.15	moderate — a rounded, amplified signal
IFI-U · mGluR1	~ 0.15	moderate — depth more than edge
IFI-C · TRPM8	~ 0.14	moderate — a clear, placeable coolness
IFI-P · TRPA1	~ 0.14	moderate — sharp but short-lived
IFI-B · TAS2R38	~ 0.10	narrow — subtraction is easy to notice
IFI-K · CaSR	~ 0.095	narrow — a continuity felt more than tasted
IFI-A · salivary PRPs	~ 0.075	narrowest — the most delicate finish to hold

Set seven percent against those tolerances and the logic falls out cleanly. On the loud, forgiving axes it buys an enormous margin: a lot could wander three times as far as it does and still pass unnoticed. On the delicate axes — astringency, kokumi, the quiet continuity of a broth — the margin is thinner, which is exactly why those axes get the most of our attention and tend, in practice, to land not at the ceiling but comfortably under it, at four, five, six percent. The ceiling is set by the hardest case; most lots clear it with room to spare.

Eight receptors, eight tempers — *why each axis forgives differently.*

The ordering in that table is not arbitrary, and it is not something we chose. It is written into the proteins. Each axis forgives a batch exactly as much as its receptor is built to forgive, and the reasons are mechanical, one channel at a time.

Heat is loud because its channel is loud. Capsaicin binds TRPV1, a non-selective cation channel on the trigeminal endings of the tongue, slipping into an intracellular pocket near the S4–S5 region and lowering the channel's thermal threshold so the body reports a fire that is not there. TRPV1 has a high single-channel conductance and opens cooperatively, so suprathreshold heat saturates into a broad, bright plateau. A plateau is forgiving: once the fire is well lit, a few percent more fuel barely moves the flame, and the Weber fraction stays wide.

Sweetness and umami are moderate because they are amplified and broadly tuned. Both run through class-C G-protein-coupled receptors with a large "Venus-flytrap" extracellular domain — TAS1R2·TAS1R3 for sweet, mGluR1 and the TAS1R1·TAS1R3 heterodimer for umami. A GPCR is an amplifier: one bound molecule triggers a cascade through gustducin, PLCβ2, and the TRPM5 channel, and amplification smooths small input differences into a rounded output. The receptors are also broadly tuned, answering to a family of ligands rather than one key, which blunts their sensitivity to a slight change in any single compound.

Bitterness is narrow because it is a burglar alarm. TAS2R38 is one of a couple dozen bitter receptors, and the whole bitter system evolved to catch potential toxins at the faintest trace — a withdrawal reflex, tuned to be touchy. A system built to notice *the arrival* of something unwelcome is, by the same wiring, quick to notice its *subtraction.* Attenuate bitterness and the alarm registers the quiet; the tolerance has to be narrow because the channel is, by design, suspicious.

Kokumi is narrow because it is not a taste but a continuity. The calcium-sensing receptor, CaSR, answers to extracellular calcium and to γ-glutamyl peptides like glutathione and γ-Glu-Val-Gly, and it does not carry a flavor of its own — it modulates the *persistence and fullness* of the others, the sense that a broth keeps giving. Change a modulator and you do not alter one note; you alter the whole chord's sustain, and a shift in the whole is felt more readily than a shift in a part.

Cool-burn and pungency sit in the middle, for opposite reasons. TRPM8 is the cold-and-menthol channel; menthol shifts its voltage dependence so it opens at ordinary mouth temperature, reporting a cool that places itself cleanly and so is moderately easy to track. TRPA1, the pungency channel, is activated by electrophiles — allicin from garlic, the isothiocyanates of mustard and wasabi, cinnamaldehyde — that *covalently* modify cysteine residues on its cytoplasmic ankyrin domain. Reactive, covalent chemistry is a little unruly and short-lived, which widens the tolerance from the delicate end even as the sharpness itself is vivid.

Astringency is the narrowest because it is barely a receptor at all. The pucker of tea or young wine is not a taste; it is the salivary proline-rich proteins being bound and precipitated by tannins, stripping the mouth of its own lubrication until the tissues report drag. It is a tactile event layered on the shifting baseline of your own saliva, and anything built on a moving baseline is delicate to hold steady. The most physical of the eight sensations is, fittingly, the one that demands the most precision from us.

Read the two sections together and the design table stops being a list of tolerances and becomes a map of receptors. We do not assign the margins; the channels do. Our job is to build *under* whichever one is strictest.

Reading the score off a curve — *where the number actually lives.*

Before we can hold a score to seven percent, we have to say precisely what the score *is* and how a number gets read from a living receptor. The Index anchors every axis to a half-maximal activation:

> **$IFI(c) = 1000 \cdot EC_{50}(ref) / EC_{50}(c)$** — loudness relative to a fixed reference, scaled by a thousand.

The reference EC_{50} for each axis is a stated convention constant that defines IFI = 1,000 as the baseline for that axis — not a literature-derived measurement, but a fixed anchor chosen by convention, in the same spirit as the Scoville scale's dilution-to-threshold definition.

EC_{50} is the concentration of a stimulus that drives its receptor to half of full response. We do not read it off the tongue; we read it off a dose-response curve measured on cells engineered to display the receptor, watching calcium flood in as the dose climbs. Fit that curve and it takes the Hill form:

> **$R = R_{max} \cdot [A]^n / (EC_{50}^n + [A]^n)$** — the sigmoid of dose against response; at $[A] = EC_{50}$, R is exactly $R_{max} / 2$.



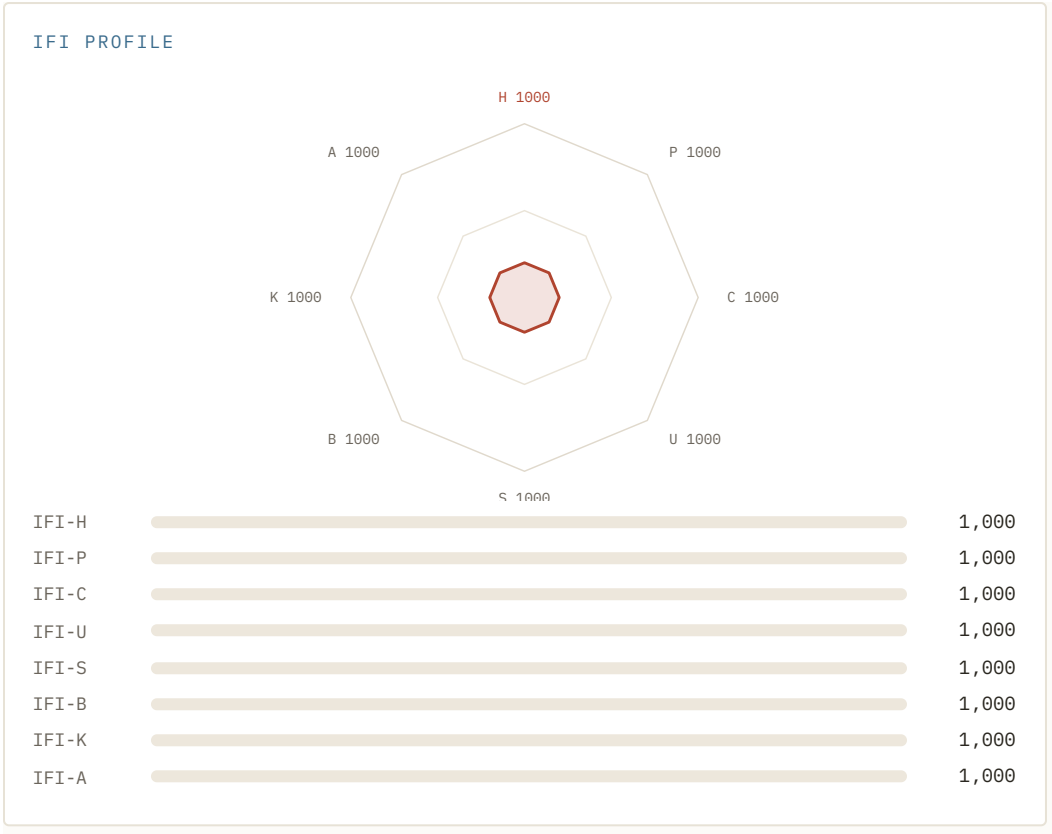
Reading a receptor's answer off the stage — the dose-response curve is where an EC_{50}, and therefore an IFI score, is actually born.

The Hill coefficient *n* is the steepness of that S, and it decides how tightly a measurement can pin the number. Rearrange the curve and the concentration needed to climb from a tenth of full response to nine-tenths is a fixed ratio, 81^(1/n): for a shallow channel (n = 1) that window is an 81-fold spread of concentration; for n = 2 it collapses to 9-fold; for n = 3 to about 4.3-fold; for n = 4 to 3-fold. TRPV1's capsaicin response is cooperative and steep, so its dose-response curve is nearly vertical through the middle — which is exactly why we anchor the score at EC_{50} and nowhere else. The steepest point of the curve is where the assay is most sensitive to a real change and least sensitive to noise, the one place a batch's concentration translates into a sharply defined half-activation. Measure in triplicate, fit each curve, and the scatter of the fitted EC_{50} values *is* the batch's variability — the raw material of the CV we are about to spend the rest of the chapter defending.

Notice what the formula does to the human. The historical heat number, the Scoville Heat Unit, was born in 1912 as an organoleptic dilution test: steep an extract, dilute it in sugar-water until a small panel of tasters could no longer detect the burn, and call the dilution factor the score. That is a *threshold* measurement — a just-noticeable-difference experiment run with people as the instrument — and it carried all the noise of people, swinging by tens of percent between rooms and panels. EC_{50} moves the measurement off the tongue and onto the receptor, which is the first and largest step toward a number you can hold to seven percent. We kept the reverence for the mouth. We just stopped asking it to be a voltmeter.

The arithmetic of the index — *seven percent, worked through the design tiers.*

Now the worked numbers. The reference anchors each axis so that a score of 1,000 means the candidate matches the reference's half-activation exactly — a stated convention constant that defines the baseline for each axis.



Invert the formula and a score becomes a concentration. Because IFI(c) = 1000 · EC_{50}(ref) / EC_{50}(c), the candidate's half-activation, in units of the reference's own, is simply EC_{50}(c) / EC_{50}(ref) = 1000 / IFI. And the fold-potency over the reference — how much *less* of the candidate it takes to reach the same half-activation — is the mirror of that, IFI / 1000. A score, divided by a thousand, is literally how many times louder than the reference the axis speaks. Phoenix's predicted heat score of 4,200 is not a decoration; our model predicts the plant reaches TRPV1's half-activation at 1000/4200 = 0.238 of the reference concentration, which is to say 4.2 times more potently.

Run the flagship heat design series through that inversion and the consistency spec becomes concrete. The last column is each score's ±7% band — the full width a lot is allowed to wander and still count, to the mouth, as the same food.

CULTIVAR · SKU	PREDICTED SHU	IFI-H	EC_{50} (REFERENCE UNITS = 1000 / IFI)	±7% BAND ON THE SCORE
Ember · SS-0087	866K	1,570	0.637	1,460 – 1,680
Blaze · SS-0091	1.72M	2,630	0.380	2,446 – 2,814
Firestorm · SS-0094	2.47M	3,490	0.287	3,246 – 3,734
Phoenix · SS-0096	2.95M	4,200	0.238	3,906 – 4,494

Two things are worth reading off that table. First, the bands do not touch. Firestorm tops out at 3,734 and Phoenix begins at 3,906 — a 172-point no-man's-land between the two closest tiers shown here, and far wider gaps below. The design tiers are spaced deliberately wider than the consistency band, so that any two *batches of one tier* are indistinguishable while the *tiers themselves* stay cleanly, reliably apart. Consistency within a tier and resolution between tiers are the same discipline read in two directions.

Second, the 7% we spend on the score is the same 7% we spend on the concentration beneath it, and that is not an accident of this example but a theorem. Inverting a number preserves its relative

scatter:

> **σ(IFI) / IFI = σ(EC_{50}) / EC_{50}** — to first order, a reciprocal carries its coefficient of variation through unchanged.

The proof is one line of logarithms: ln IFI = ln(1000 · EC_{50}(ref)) − ln EC_{50}(c), so a small wobble d(ln IFI) equals −d(ln EC_{50}(c)), equal in size and opposite in sign. Relative noise in the measured half-activation lands as equal relative noise in the published score. There is a second-order wrinkle an honest formulator should keep in a back pocket: the reciprocal is slightly convex, so averaging many noisy EC_{50} values biases the mean score upward by about the square of the CV — at 7% that is roughly 0.5%, small but real, and we correct for it rather than pretend it is zero. A number offered as a form of respect should not quietly drift half a percent high.

A budget for variance — *where the scatter is allowed to come from.*

Seven percent is a coefficient of variation, the plainest statistic in the drawer: the lot's standard deviation expressed as a fraction of its own target.

> **CV = σ / μ** — the spread of a lot, in units of where it was aiming.

The number earns its weight by comparison. Left to its own devices — select the phenotype, hope the genotype follows, grow it out under whatever the season brings — a wild-type capsaicinoid line scatters at a CV around 42%. Engineered to spec, the same output holds at or below 7%. That is a 6-fold tightening of the standard deviation and, because variance is the square of it, a 36-fold collapse of the variance itself. The mouth was never going to notice a 7% drift; it most certainly noticed a 42% one, which is the whole reason "it tastes different this time" was the oldest complaint in food.

The question a manufacturer actually has to answer is *where those seven points are allowed to come from,* and the arithmetic of combining them is unforgiving in a useful way. Independent, multiplicative sources of variation do not add their percentages; they add their *squared* percentages, and the total is the square root of the sum:

> **CV^2(total) = Σ CV^2(stage)** — independent stages add in quadrature, not in line.

Quadrature is a strict accountant. If the pipeline has several roughly equal, independent stages, each one may only carry 7%/√k before the stack breaches the line. Below is the budget written out — the per-stage ceiling that keeps the finished lot at or under seven percent.

INDEPENDENT STAGES, *K*	MAX CV EACH STAGE MAY CARRY (7% / √K)
1	7.0%
2	4.9%
3	4.0%
4	3.5%
5	3.1%



Replicate trays under one light — a batch is not a plant but a population, and its scatter is the thing the whole program is built to shrink.

A real lot passes through about five such stages — the uniformity of the construct itself, the controlled-environment cultivation, the timing of the harvest window, the extraction and sample prep, and the receptor assay that reads the result — so each stage has to hold near 3.1%. That is the practical meaning of a quality program: not a single heroic measurement at the end, but five quiet disciplines upstream, each kept to roughly three percent, so that quadrature can stack them into seven and no further. It is why the controlled chamber matters, why the harvest is a window and not a guess, why the assay runs in triplicate. Each is a term in a sum of squares, and the sum of squares is what the tongue will eventually audit.

From percent to perception — *the honest experiment, in cups.*

So far we have asserted that seven percent is where perception goes quiet. A sensory scientist would not take that on faith; they would run the experiment, and the experiment has a shape. Sit a trained panel down with two lots and the honest test is a forced choice. In the triangle test, three cups are poured — two from one lot, one from the other — and the taster must pick the odd cup. In the two-alternative test, two cups are poured and the taster must say which is the sweeter, the hotter, the more astringent. If the lots were truly identical the panelist would be guessing: right one time in three on the triangle, one in two on the pair. The whole question of consistency reduces to a single number — how far above chance can the panel climb?

That number runs through a quantity the field calls **d′**, the perceptual distance between two samples measured in units of the noise of the tasting itself.

> **d′ = Δμ / σ(perceptual)** — the gap between two lots, divided by the mouth's own scatter.

Thurstone's century-old insight is that each sip lands as a point drawn from a little cloud, and d′ is just the distance between the two clouds' centers in units of their shared width. When d′ is near zero the clouds sit atop each other and the panel guesses; as d′ grows the odd cup becomes findable and the probability of a correct call climbs its slow S toward certainty. The psychometric function — the exact curve from d′ to proportion-correct — is different for each test, and the difference is instructive. The directional two-alternative test is the more *sensitive* instrument, because telling the taster what to look for is a form of help; the triangle, which primes nothing, is the more conservative. Both are tabulated below from the Thurstonian model.

d′	Triangle: P(correct), chance = 0.333	Directional 2-AFC: P(correct), chance = 0.500	What a panel experiences
0.0	0.333	0.500	pure chance — indistinguishable
0.5	0.360	0.638	a faint pull, unreliable
1.0	0.418	0.760	the occasional lucky find
2.0	0.637	0.921	"I think that one's different"
3.0	0.811	0.983	reliably told apart

Now fold the two halves of the chapter together, because this is where the line is actually set. A lot that sits a fraction f off its target, on an axis with Weber fraction k, presents the panel a perceptual distance of roughly d′ ≈ f / k — the drift measured in units of the axis's own just-noticeable step. Put the 7% ceiling in for f and read off what the most demanding honest test would return, axis by axis.

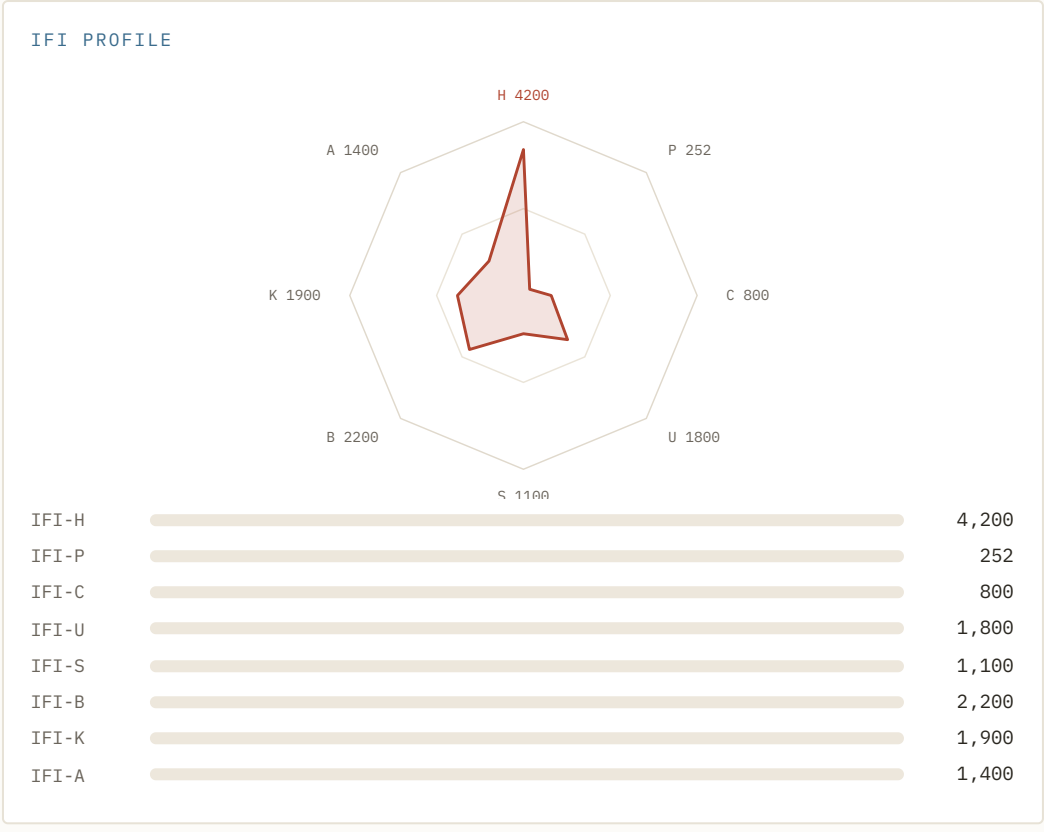
Axis · receptor	Weber *k*	d′ at the 7% ceiling	2-AFC P(correct)	Triangle P(correct)
IFI-H · TRPV1	0.22	0.32	0.59	0.34
IFI-S · TAS1R2·R3	0.15	0.47	0.63	0.36
IFI-U · mGluR1	0.15	0.47	0.63	0.36
IFI-C · TRPM8	0.14	0.50	0.64	0.36
IFI-P · TRPA1	0.14	0.50	0.64	0.36
IFI-B · TAS2R38	0.10	0.70	0.69	0.38
IFI-K · CaSR	0.095	0.74	0.70	0.39
IFI-A · salivary PRPs	0.075	0.93	0.75	0.41

Read it as the design specification it is. On the loud axes, a lot sitting at the very ceiling hands even the sensitive directional test a coin barely weighted — heat at 0.59, a hair above the 0.50 of pure guessing. On the narrowest axis, astringency, the same 7% drift pushes the directional test up toward 0.75: still weak, still far from the reliable 0.92 that a d′ of 2 would bring, but no longer invisible. That single number, 0.75, is the whole argument for why we do not run astringency or kokumi at the ceiling. Drop them to four or five percent and d′ falls back toward 0.5, the directional test sags toward 0.6, and the delicate finish rejoins the rest of the plate in the territory where the honest answer is *I can't tell.* We are not aiming for lots a panel *fails* to distinguish by luck. We are aiming for lots whose d′ is

so small there is almost nothing there to distinguish — so that even a panel built to find differences, trying its hardest, lands near the chance line and reports the truth: these are the same.

Holding eight at once — *the full-spectrum case.*

Every number so far has been one axis at a time, and most foods live mostly on one or two. The hardest case is the one that lives on all eight, because then seven percent is not a line but a box in eight dimensions, and a lot has to land inside it on every face at once. Phoenix is that case — the first cultivar engineered to register on all eight axes together.



Eight scores, and the spec asks each to hold its own ±7% band simultaneously: heat inside 3,906–4,494, pungency inside 234–270, the quiet astringency of 1,400 inside 1,302–1,498, and five more besides. The axes do not measure the same molecule, so to first order their batch variations are independent, which means the probability of a lot clearing *every* band at once is the product of the eight single-axis probabilities — and a program that only *usually* hits each axis will *rarely* hit all eight. This is the real reason the delicate axes are driven well under the ceiling rather than parked on it: an

eight-axis food cannot afford any single axis spending its whole margin, because eight near-misses multiply into a miss. Holding Phoenix is not holding one number seven times easier than holding eight separate foods; it is holding a joint distribution, and the discipline that earns it is the quadrature budget applied on every axis at once. The reward is a plant designed to taste, batch after batch, like the same deliberate chord — not close enough to the last harvest, but, to the one instrument that matters, identical.

What the line cannot do — *the limits we keep in view.*

A number offered this plainly owes the reader its edges. Seven percent is a perceptual-intensity floor, and it is honest only inside the claims it actually makes.

It is a statement about *intensity*, not *identity.* Two lots inside the band are indistinguishable in how loudly each axis speaks; they are not proven molecule-for-molecule identical, and they should not be sold as if chemistry and perception were the same claim. The Index measures loudness, and loudness is what the mouth grades.

Weber's constant is not constant everywhere on an axis. Near threshold, where a signal is just surfacing, and near saturation, where a channel is already full, the fraction *k* swells and the simple d′ ≈ f / k picture loosens — which is why the spec is checked band by band along the design tiers, not assumed uniform from the faintest shishito to the top of the superhots. A line drawn for the middle of a range must be re-checked at its ends.

Intensity is not liking. A change can sit below d′, perfectly imperceptible as a difference, and still shift how much someone *enjoys* the food, because pleasure is built from aroma, texture, memory, and context that no single-axis intensity test touches. We hold the axes to seven percent because that is the part we can measure and promise; we do not pretend it governs the whole of delight.

And the mouth is not one mouth. d′ is a probability, not a verdict, and a program that lands a population of tasters near the chance line will still meet the occasional golden palate, or the occasional lot sitting at the tail of its own distribution, for whom a difference surfaces. The response is not to widen the claim but to tighten the lot — to keep driving σ down until the tail itself falls quiet. That is the direction every discipline in this chapter points: Weber telling us where perception begins, the receptors telling us which axis is strictest, the dose-response curve telling us where to read the number, quadrature telling us where the scatter may come from, and d′ telling us, in cups, when we have finally arrived at a food we can hand the body with intention and call, truthfully, the same. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

Chapter 3 · Four doors — *where drift comes from, and why it adds in quadrature.*

A designed flavor is a promise that a number will hold — that the warmth we built into a fruit this spring will arrive on a stranger's tongue a year from now, in a kitchen we will never stand in, exactly as we meant it. We did not come to that promise as statisticians. We came to it as a family that has always believed the deepest form of care is not one perfect meal but the *same* good thing, set down every time, so that no one has to wonder whether tonight is the night it is off. To keep a promise like that you have to know every way it can be broken, and the good news — the quiet, almost unreasonable good news at the center of this whole chapter — is that in a living process the breakage has a small and knowable set of sources. Drift walks in through exactly four doors.

There is **genetic** variation — whether the plants even carry the same instructions in the first place. There is **environmental** variation — how fully each plant expresses those instructions, under its own light and water and warmth. There is **process** variation — what extraction, drying, milling, and formulation do to the compound after it has left the fruit. And there is **analytical** variation — the wobble in the very act of measuring, which can masquerade as a fault in the plant and send a whole program chasing a ghost. Name the four and you hold a map of every way a number can wander. The rest of quality is shutting them, in order, and the order is not a matter of taste. It is set by a single piece of arithmetic that most makers have backwards.

Why variances add, and coefficients of variation do not

Here is the fact, stated plainly: *coefficients of variation do not add. Variances do.* It sounds like pedantry and it is strategy, and getting it right is the difference between a lab that exhausts itself and a lab that closes the one door actually letting the drift in.

Begin with what a variance is. The standard deviation, σ, is the typical distance a measurement sits from its own mean; the variance, σ^2, is that distance squared. The coefficient of variation normalizes it against the mean so that looseness reads the same everywhere on every axis:

> **CV = σ ÷ μ** — the spread of a score divided by its target, read as a percent.

Now ask what happens when two independent sources of scatter land on the same number at once. Write the measured value as a target plus two small, independent wobbles, X and Y, and the algebra of variance is exact and unforgiving:

```
Var(X + Y) = Var(X) + Var(Y) + 2·Cov(X, Y)
```

The last term is the covariance, the degree to which the two wobbles move together. When the sources are independent — a seed's genotype knows nothing of an assay's reagent lot — that covariance is zero, and it falls away. What survives is clean:

```
σ²(total) = σ²(genetic) + σ²(environment) + σ²(process) + σ²(analysis)
```

The *squares* sum, not the spreads themselves. And because every one of these doors perturbs the same quantity around the same target mean, dividing the whole line by μ² carries the rule intact into the coefficients we actually budget in:

```
CV²(total) = CV²(genetic) + CV²(environment) + CV²(process) + CV²(analysis)
```

This is quadrature — the sources combine like the legs of a right triangle, where the hypotenuse is dominated by the longer leg and barely notices the shorter. Work one example and the gift becomes obvious. Suppose one door stands wide open at six percent and another is ajar at two. Intuition says the total is eight. The arithmetic says $\sqrt{(6^2 + 2^2)} = \sqrt{40} \approx$ **6.3** — the two-percent door added barely a third of a percent to a problem already owned by the six, a five percent lift on a six percent problem. A source a third the size of the largest all but vanishes beneath it.

The reverse is just as instructive, and it is the reason no single door may be left gaping. Two *equal* doors do not double; two sources at five percent each combine to $\sqrt{(5^2 + 5^2)} = \sqrt{50} \approx$ **7.1**, not ten. In general, n equal doors of size c combine to $c \cdot \sqrt{n}$, so the penalty for a second comparable source grows as a square root, gently — while the penalty for one source far larger than the rest is the whole number, because the largest squared term swamps every smaller one. Quadrature rewards balance and punishes a single open door, and that single arithmetic fact writes the entire strategy of a quality system.

The budget, and the order you spend it

Seven percent is not found. It is budgeted — assigned, door by door, and spent in a deliberate order. And the order falls straight out of quadrature: you do not chase all four doors at once, spreading effort thin across sources that cannot, individually, matter. You find the *widest-open* door, you close it, and you watch the total collapse toward whatever the next-widest will allow. Then you find the new widest, and close that.

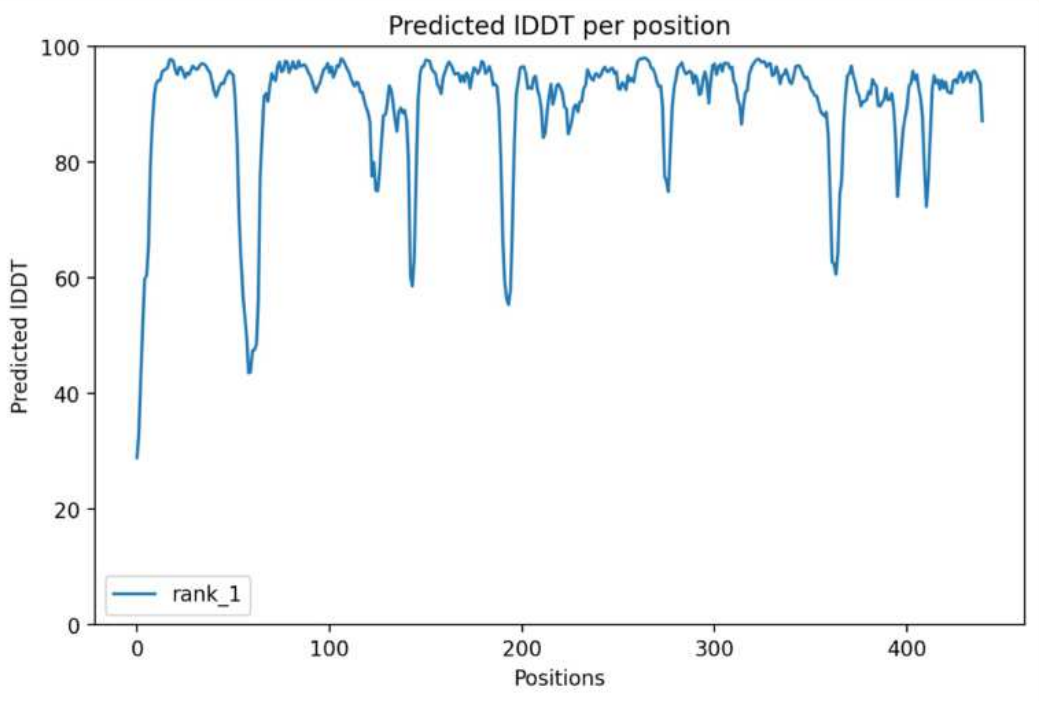
Watch it happen on a real line. A wild-type pepper, grown and read honestly, carries a batch CV near **42%** — and that 42 is not one failure but four doors, apportioned. The engineered line we ship is designed to hold **≤7%**. The distance between them is walked one door at a time, and the staircase looks like this.

STAGE	GENETIC	ENVIRONMENTAL	PROCESS	ANALYTICAL	CV(TOTAL)
0 · wild-type, as found	40%	10%	6%	5%	**42.0%**
1 · seed fixed (isogenic)	3%	10%	6%	5%	**13.0%**
2 · room pinned	3%	4%	6%	5%	**9.3%**
3 · process tightened	3%	4%	3%	5%	**7.7%**
4 · ruler disciplined	3%	4%	3%	3%	**6.6%**

Every total is the quadrature of its row — $\sqrt{(40^2 + 10^2 + 6^2 + 5^2)} = 41.96$ at the top, $\sqrt{(3^2 + 4^2 + 3^2 + 3^2)} = 6.56$ at the bottom. Read the staircase and the two deepest lessons of this chapter are sitting in plain numbers. First: closing the widest door is enormously powerful and yet it does **not** take you to zero. Driving genetics from 40% to 3% — a thirty-seven-point fix, the largest single move in the entire program — only moves the total from 42% to 13%, because the moment the seed is nailed shut the ten-percent room becomes the thing you are now measuring. You did not fail; you simply revealed the next door. Second: the total never falls below its own largest remaining term, so there is no crossing the seven-percent line until the *last* door is shut. Discipline at the exit cannot rescue a gaping door upstream, and a magnificent seed cannot rescue a drifting ruler.

This is why a lab that does not know the quadrature rule works itself to exhaustion polishing a two-percent analytical step while a fifteen-percent environmental door stands wide — spending its best weeks on a term that, squared, contributes almost nothing. A lab that knows the rule spends everything on the one door that is actually open, collects the collapse, and only then looks for the next. The budget takes care of itself, because at every step exactly one term is worth touching.

The first door · the seed, where the plant's instructions live



The Pun1/AT3 capsaicin synthase, colored by model confidence — the active-site cleft where vanillylamine meets its acyl partner, and the fold two designed edits were written to hold still.

The widest door in all of plant variation is genetic, and it is wide for a reason worth understanding at the molecule. Heat in a pepper is not stored; it is *made*, at the end of a long enzymatic relay. The phenylpropanoid branch — **PAL → C4H → 4CL → HCT → COMT**, a backbone that runs in every vascular plant we have ever opened —

delivers a vanillyl group; a parallel branched-chain route through **BCAT** and **KAS** delivers a short, methyl-branched fatty acid; **pAMT** aminates the vanillyl into vanillylamine; and at the final step the capsaicin synthase **Pun1/AT3** condenses the two into capsaicin. The length and saturation of that acyl chain set the balance between capsaicin and its near-twin dihydrocapsaicin, and so the very *shape* of the burn — its onset, its climb, its loudness per milligram.

Every enzyme in that relay is written in the genome, and in a wild or landrace population the genome is not one text but a crowd of close variants. Plants segregate — a fruit may carry one functional copy of Pun1 or two, one allele of pAMT or another, promoters that fire a little hotter or a little cooler — and each of those differences is a small, independent motion in the amount of compound the plant finally makes. Forty-two percent of wander, most of it genetic, is simply a crowd of slightly different factories all labeled the same. You cannot pin a number to a population that is still arguing about its own instructions.

So we close the door at its source: a single fixed genotype, propagated as isogenic lines, so that every plant in a lot is — as nearly as biology permits — the same plant. Into that fixed text we write two substitutions a breeding program would never have stumbled onto, each scored first in a protein model and each earning its place by a computed change in folding free energy, ΔΔG, where a more negative number means a more stable, more faithfully folded enzyme.

EDIT	EFFECT	ΔΔG (KCAL/MOL)
S39L	binding affinity ↑ (N-terminal)	−2.1
L345G	steric relief, active-site cleft	−1.8
C175S	*disulfide-misfold prevention (structural model annotation; not in shipped construct)*	*−0.9*

The two shipped edits sum to a stabilization of **−3.9 kcal/mol**, and that number does real statistical work, not just catalytic work. Folding is an equilibrium, and the fraction of enzyme that reaches its active shape is governed by that free energy through the Boltzmann relation. At body-adjacent temperature, RT ≈ 0.59 kcal/mol, so the ratio by which the folded state is favored shifts by exp(3.9 ÷ 0.59) ≈ exp(6.6) ≈ **740-fold** toward the competent fold. A factory whose key enzyme almost always folds correctly is a factory whose output stops wobbling for one whole class of reasons — misfold, degradation, a bad day for a fragile protein — and that trimmed wobble is a genetic variance component driven toward zero. Fixing the genotype removes the crowd; stabilizing the fold quiets the one enzyme that is left. Between them the door that gapes at forty percent in a wild population closes to a few percent in ours.

We are honest about the edge of this. A summed ΔΔG assumes the two shipped edits do not interfere with one another — that their effects are additive rather than epistatic — and each computed value carries roughly a kcal/mol of its own uncertainty, so the seven-hundred-fold figure is an order-of-magnitude statement, not a

decimal one. What bears it out is not the prediction but the batch CV that comes back from the bench, lot after lot, under the line. The model points; the measurement decides.

The second door · the room, where the instructions are read



The controlled chamber, where the room is held still — one photoperiod, one temperature, one vapor-pressure deficit — so the same genotype is asked the same question every morning.

Fix the seed and, by quadrature, the room becomes the widest door — the ten-percent term that now sets the total. Genetics is the instruction; environment is how loudly the plant reads it aloud. The same fixed genotype, handed two different rooms, will make two different amounts of compound, because capsaicinoid biosynthesis is not a clock that runs at a constant rate but a response that leans on its conditions. Light is the throttle: the daily integral of photons a leaf receives, its daily light integral, feeds the carbon that the whole phenylpropanoid relay spends, and a plant short of light is a factory short of raw material. Warmth sets the tempo of the enzymes. And mild water stress, famously, *lifts* capsaicinoid accumulation — a drought-stressed pepper concentrates its heat — which means that the single most romantic variable in farming, the weather, is also a door swinging on its hinge.

To close it we take the weather away. In a controlled chamber the photoperiod does not drift with the season; dawn arrives at the same hour, at the same spectrum, every day of the plant's life. Temperature is held, the vapor-pressure deficit that governs how fast a leaf transpires is held, the nutrient feed is held, the irrigation is scheduled rather than prayed for. A plant grown under a pinned dawn is answering the same question every morning, and a plant always asked the same question tends to give the same answer. There is a deep reverence in that stillness — it is how we let a plant be exactly, repeatably itself — and there is a hard number under the reverence: the environmental term falling from ten percent toward four, which is the whole of what the staircase needed from this door.

The third door · the process, what we do to the fruit after we pick it

The third door is the one formulators own most completely, because it opens only after the plant has done its work. Everything between the fruit and the finished blend is process, and every step of it can add or subtract from a number the biology already got right. Extraction is the first risk: the ratio of solvent to solids, the

temperature, the dwell time, all set how completely the capsaicinoids come out of the tissue, and an under-extracted lot reads low for a reason that has nothing to do with the plant. Drying is the second, and the sharpest. Capsaicin is thermolabile — hold it too hot for too long and it degrades, so a drying oven that drifts five degrees across a season writes a slow trend into the data that looks exactly like the crop getting weaker. Milling is the third and the most underestimated: particle size governs surface area, surface area governs extraction yield, so an inconsistent grind becomes an inconsistent number downstream. And formulation is the last — the matrix the compound finally sits in, its pH, its exposure to light and oxygen, each of which can quietly carry capsaicin off over shelf life.

None of this requires heroics to close; it requires standard work held to a specification, which is the least glamorous and most reliable form of care there is. A fixed solvent ratio and extraction temperature. A drying profile logged and bounded, not judged by eye. A particle-size window with a sieve behind it. Storage dark, cool, and inert, so the compound we built is the compound that ships. The failure mode to fear here is not a dramatic one; it is a slow instrument drift — an oven, a mill, a storage room warming through a summer — that masquerades as biology and tempts a program to go re-engineer a plant that was never the problem. The process door is closed with logs and discipline, and it is the cheapest of the four to hold once you have decided to.

The fourth door · the ruler, and its own quiet wobble



The assay bench, where a dose–response curve becomes an EC_{50} and an EC_{50} becomes a score — and where the ruler's own small tremor must be measured before it is ever blamed on the plant.

The last door is the subtlest, because it is the one we build ourselves. Before a flavor can be held to a number, the number has to be read, and the reading has a variance of its own that lives in no plant at all. The score begins as a curve. Put a receptor — for heat, the trigeminal channel TRPV1 — in front of rising concentrations of the compound that speaks to it, and the response is not a line but an S: silent at low dose, rising steeply through a middle band, flattening as every channel that can open has opened. The algebra of that S is the Hill equation,

$$R = R_{max} \cdot [C]^n \div (EC_{50}^n + [C]^n),$$

where EC_{50} is the concentration that drives the receptor to exactly half of its maximum, and n, the Hill coefficient, is the steepness set by how cooperatively the channel's subunits open; for capsaicin at TRPV1 it sits near 2. The EC_{50} is pulled from that curve by nonlinear regression, and here is the first edge case a careful bench respects: the midpoint is the best-conditioned point on the whole curve, so the dilution series must *bracket* it. Sample only the flat threshold and the flat ceiling and the fit will still return a number, but a loose one — a poorly determined EC_{50} is the analytical door swinging open while the plant stands innocent.

From that EC_{50} the score is built as a ratio, and the ratio is the quiet genius of the index:

> **IFI = 1,000 × EC_{50}(reference) ÷ EC_{50}(candidate)**

The scale is pinned so the defined reference reads exactly 1,000 on that axis — a stated convention, the way the Scoville scale itself is defined by dilution-to-threshold, not by any single measured compound. Each sensory axis carries its own reference constant, so IFI values on different axes anchor to different defined points and are not directly comparable across axes. The reason the measurement is a ratio and never an absolute is that a raw EC_{50} drifts with the cell line's age, the day's warmth, the reagent lot — and when a single multiplicative bias scales *both* the reference and the candidate on the same plate, it divides straight out of the quotient. The ratio cancels the room's systematic error. What it cannot cancel is the room's *random* tremor, and this is where quadrature returns to govern the ruler exactly as it governs the plant. For a quotient of two independent measured quantities, the relative variances add — the delta-method result is, to first order,

> **CV^2(IFI) = CV^2(EC_{50} reference) + CV^2(EC_{50} candidate),**

which is the quadrature rule again, wearing a different hat. The consequence is sobering and worth the arithmetic. Read each EC_{50} once at a bench holding five percent, and the score inherits $\sqrt{(5^2 + 5^2)} = \sqrt{50} \approx$ **7.1%** — the ruler alone, with no help from any plant, has already spent the entire batch budget and breached the line. The fix is not a better plant; it is a steadier hand. Read each EC_{50} as the mean of three replicate wells and its sampling CV falls by $\sqrt{3}$, from 5% to 2.9%, so the score now inherits $\sqrt{(2.9^2 + 2.9^2)} \approx$ **4.1%** — the assay door bought back under budget, one well at a time. The same-day reference subtracts the room's bias; replication subtracts its noise. Together they are how the fourth term in the staircase falls from five percent to three.

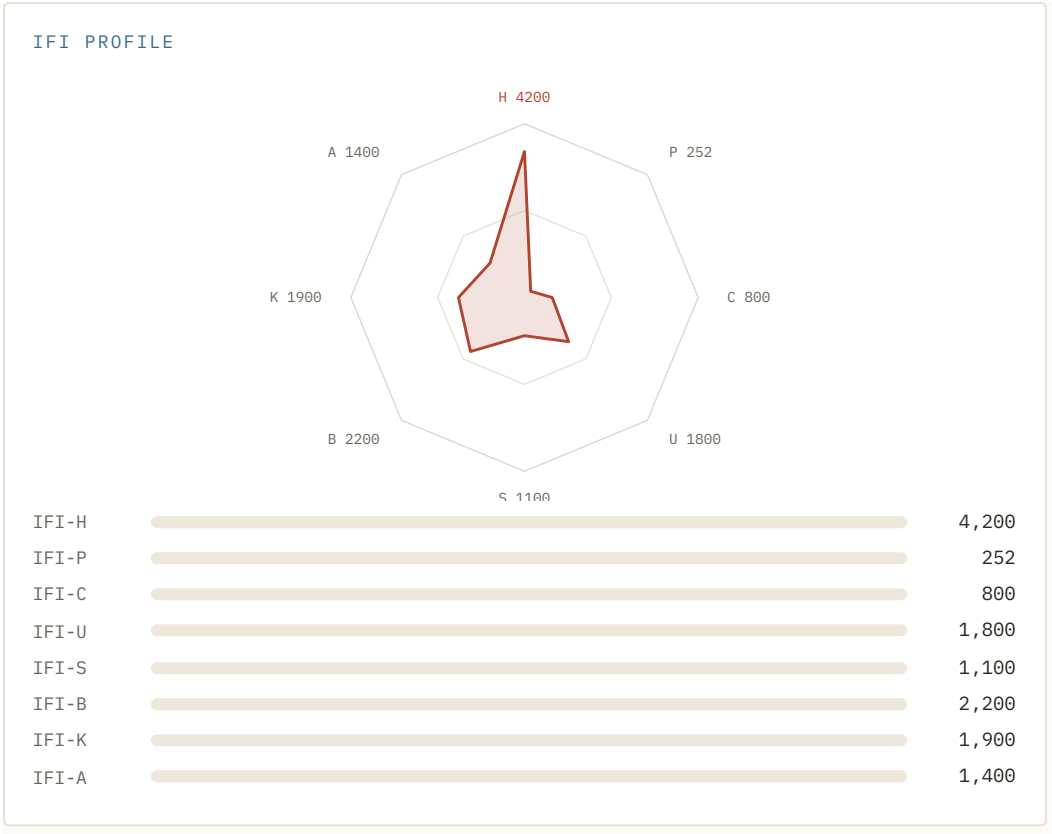
It also matters *which* ruler you pick up, because two very different numbers both call themselves "heat." The Scoville figure is a chemistry number — a count, by dilution, of total capsaicinoid present — and it converts cleanly from a laboratory concentration, since pure capsaicin registers at roughly 16,000,000 SHU and so **SHU ≈ 16 × (capsaicinoids in ppm)**. Our model predicts Phoenix at approximately 2.95 million SHU; back-converting that design figure yields roughly 184,000 ppm, about eighteen percent capsaicinoid by dry mass — an extraordinary computational target, and exactly the

sort of figure our design platform reaches that selection never could. The IFI, by contrast, is a biology number: not how much compound a fruit holds, but how forcefully that blend engages the receptor, anchored to the EC_{50}. The two climb together but not in step.

CULTIVAR · SKU	SHU (PREDICTED)	CAPSAICINOIDS (PREDICTED PPM = SHU ÷ 16)	IFI-H	EC_{50} RATIO (1000 ÷ IFI)
Ember · SS-0087	866K	54,125	1,570	0.637
Cinder · SS-0088	1.05M	65,625	2,290	0.437
Kindle · SS-0089	1.25M	78,125	2,000	0.500
Flare · SS-0090	1.48M	92,500	—	–
Blaze · SS-0091	1.72M	107,500	2,630	0.380
Wildfire · SS-0092	1.96M	122,500	2,930	0.341
Scorch · SS-0093	2.21M	138,125	3,220	0.311
Firestorm · SS-0094	2.47M	154,375	3,490	0.287
Solara · SS-0095	2.72M	170,000	3,980	0.251
Phoenix · SS-0096	**2.95M**	**184,375**	**4,200**	**0.238**

The EC_{50} ratio is just 1,000 ÷ IFI-H — the fraction of the reference concentration a cultivar needs to half-activate the channel — so Phoenix at IFI-H 4,200 reaches the midpoint at 0.238 of the reference, which is precisely what "4.2 times as loud" means in molecules. And now look at the tell buried in the middle of the design progression: our model has Kindle carrying *more* predicted Scoville heat than Cinder (1.25M against 1.05M) yet registering a *lower* IFI-H (2,000 against 2,290). By predicted mass it is hotter; to the receptor it is designed to be quieter. That is not an error; it is the gap between the two rulers made visible, written by the acyl chain that tunes the capsaicinoid blend. The lesson for the fourth door is exact: measure the chemistry number and believe you have characterized the biology, and you will attribute scatter to the wrong door entirely — you will polish a plant whose receptor loudness you never actually read. Hold the number the eater's receptor actually hears, and hold its reference beside it on the same plate.

Phoenix is the hardest ruler of all, because it is the one cultivar designed to land on every axis at once:



Eight receptors, eight numbers, each with its own four-door budget and each held to ≤7% in the same fruit — heat on TRPV1, pungency on TRPA1, cool-burn on TRPM8, umami on mGluR1, sweetness on TAS1R2·R3, zero-bitter on TAS2R38, kokumi on CaSR, astringency on the salivary PRPs. Eight rulers, eight chances for the fourth door to open, which is why the full-spectrum fruit consumes more of the quality system than any single-axis line ever will.

Nesting · telling the plant's scatter from the plate's

There is one more move, and it is what lets you point at the right door rather than guess at it. The four doors are not merely four buckets; they are *nested*, one inside the next. A single reading sits inside a plate, which sits inside a day, which sits inside a lot, which sits inside the cultivar. Scatter introduced at each level is a distinct variance component, and the tool that separates them is the nested analysis of variance — the same quadrature machinery, run in reverse, to take a pile of observed numbers apart into the levels that made them.

Its gift is a single honest figure: the fraction of all the observed variation that is *real* difference between the things you care about, versus *nuisance* born in the levels you do not. That fraction is the intraclass correlation,

> **ICC = σ²(the level you are grading) ÷ σ²(total),**

and it is the difference between an assay that can grade and one that is quietly grading itself. A healthy decomposition of a heat measurement looks like the table below.

LEVEL (NESTED)	WHAT IT IS	Σ² COMPONENT (%²)	Σ (CV %)	SHARE OF TOTAL
Between-cultivar	real design difference — the signal	144	12.0	80.3%
Lot → lot	harvest-to-harvest, one cultivar	16	4.0	8.9%
Day → day	assay day	9	3.0	5.0%
Plate → plate	plate within a day	6.25	2.5	3.5%
Reading (residual)	well-to-well	4	2.0	2.2%
Total		**179.25**	**13.4**	**100%**

Here ICC at the cultivar level is 144 ÷ 179.25 = **0.80**: four-fifths of everything the instrument reports is true difference between the cultivars we are trying to grade, and only a fifth is nuisance from lot, day, plate, and well. That is an assay you can trust to rank a progression. But the number is a warning as easily as a reassurance. Let the plate and day terms balloon — a reagent aging across a week, a reader drifting between mornings — and the signal's share collapses. We have opened programs certain the trouble was biological, run exactly this decomposition, and found **two-thirds** of the scatter living in a drifting assay: an ICC at the biological level of barely 0.33, the plant all but innocent, the ruler quietly guilty. A low intraclass correlation at the fruit level is the instrument telling you, in a number, that you are measuring your own process and calling it your plant.

Decompose first, then spend

That is the whole discipline, and it is the opposite of heroics. The widest door is not always where intuition points, and quadrature punishes guessing: a program that *felt* sure the trouble was in the greenhouse, and acted on the feeling, would have spent a whole season on a growing fix while a drifting reader — a door it could shut in a week — went on writing nuisance into every lot. So you do not act on the feeling. You decompose first; you attribute the scatter to the level that actually owns it; and only then do you spend, on the one door the arithmetic has named. Collect the collapse, re-decompose, find the new widest, and repeat.

We hold two honesties about the method so that no one leans on it past its strength. The first is that quadrature assumes the doors are independent — that genotype and room and bench do not conspire — and when they interact, when a hotter room degrades a fragile compound faster, the covariance term does not vanish and the clean sum of squares becomes an approximation rather than a law; those couplings have to be designed out, held constant, before the decomposition can be believed. The second is that the CV-quadrature form assumes small perturbations around a common mean, and that every variance component in a table is itself an estimate with its own error, sharpened only by measuring enough lots and enough days to pin it. Numbers are a form of respect for the reader, and respect includes saying where a number stops. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

But the shape of the thing is simple and it holds. Four doors, taken in the order the arithmetic assigns — the seed, because it is the widest; then the room; then the process; then the ruler — each a variance-reduction problem with its own tools, each one closed lifting its squared term out of the sum and letting the next come into view. The seven-percent line at the end is not a thing you achieve by a flourish at the exit. It is simply what is left over, what remains and cannot be removed, when four doors have been shut, one at a time, upstream — so that the stranger a year from now, in a city we will never see, tastes exactly the good thing we meant, and never once has to wonder whether tonight is the night it is off.

Chapter 4 · The dose–response curve — *EC50, and what we actually measure.*

Everything in this library hangs from a quantity most people have never been handed a name for, and before we can speak honestly about holding it steady across a thousand harvests we have to look at it directly: the EC_{50}, and the curve it lives on. When a molecule from a fruit our family grew lands on its receptor at the surface of the tongue, the receptor does not fling itself open all at once at some magic threshold. It answers gradually — a little at low concentration, more as the molecule crowds in, and then it saturates and can give no more however hard you press. Plot that answer against concentration and you get a rising S, the sigmoid that governs very nearly all of pharmacology and all of taste. The EC_{50} is simply the concentration at the middle of that S: the dose at which the receptor surrenders exactly half of everything it has. It is the honest, physical measure of *potency* — of loudness. A molecule with a low EC_{50} reaches half-response at a whisper and is therefore loud; one with a high EC_{50} needs a shout to get halfway and is therefore quiet.

We dwell on it because it is the thing we actually measure. We do not measure heat; we measure a half-maximal concentration and call its inverse heat. We do not measure savor or sweetness; we measure where, on a column of rising doses, each receptor gives half of its all. The whole Index is a disciplined way of reporting that one number, axis after axis, and a quality book such as this one is finally a book about keeping that number from wandering. So we begin where the wandering begins: on the curve.



A dose series set across a plate — the half-maximal point we solve for is the single number the whole Index hangs from.

The receptor, one subunit at a time

Take the heat axis, because it is the one the body shouts about. Capsaicin binds TRPV1, a non-selective cation channel standing in the membrane of the trigeminal nerve endings of the tongue — the same channel that opens to temperatures above 43 °C and to a flood

of protons. The molecule warms nothing. It slips into a pocket on the intracellular side of the channel, between the third and fourth transmembrane helices, and from there it lowers the temperature at which the gate would have opened on its own, so the nerve reports a fire that is not in the tissue. This is why the burn is real and the thermometer is calm: the channel's own threshold has been moved, not the mouth's temperature.

The shape of the dose–response curve is written into the architecture of that channel, and this is the part a formulator must feel in the hands. TRPV1 is a homotetramer — four identical subunits arranged around a central pore, each carrying its own vanilloid pocket. Four sites around one gate is the structural recipe for *cooperativity*: the binding of one capsaicin molecule can tilt the whole assembly toward opening, making the next molecule's job easier, so the channel crosses from quiet to loud over a narrower band of concentration than a single lonely site ever would. Cooperativity is not a curiosity here. It is the reason the sigmoid for heat climbs steeply through its middle, and steepness, as the arithmetic below will show, is a tax levied on every imprecision in how much molecule the fruit actually made. The channel's four-fold symmetry and the curve's steep waist are the same fact seen from two sides.

Upstream of all of this sits the chemistry our family actually engineers. Capsaicin is built when vanillylamine condenses with an eight-methyl branched-chain fatty acid at the Pun1/AT3 acyltransferase; the length and saturation of that acyl chain set the balance between capsaicin and its near-twin dihydrocapsaicin, and that balance quietly sets the *shape* of the burn and, as we will see, part of the gap between how much heat a fruit holds and how loudly it speaks. The receptor is where the measurement happens. The enzyme is where the consistency is won or lost.

A rising S, and the number at its middle

The curve has a name and the name has algebra. For a receptor whose maximum response is *R*max, the fractional response to a concentration *c* follows the Hill equation:

```
R(c) = Rmax · cⁿ / (EC₅₀ⁿ + cⁿ)
```

Read it slowly, because the definition we care about falls straight out of it. When *c* equals the EC_{50}, the two terms in the denominator are identical, the fraction collapses to exactly one-half, and the response is half of maximum — the EC_{50} is the half-point not by decree but by arithmetic. Set *c* far below the EC_{50} and c^n is swamped by EC_{50}^n, so the response falls toward zero in proportion to c^n; set *c* far above it and c^n dominates, so the response flattens against its ceiling. The whole S is in that one ratio.

The exponent **n** is the Hill coefficient, and it governs how *steeply* the S rises through its middle. An *n* near one is a gentle, patient climb — a single independent site filling up. An *n* well above one is a cliff: a channel that goes from hush to roar over a hair of concentration because its subunits bind in concert, the four pockets of a receptor like TRPV1 pulling together. We treat the Hill coefficient as a first-

class citizen of every fit, never a footnote, because a steep curve means a small error in concentration becomes a large error in response — and so the axes carried by cooperative, high-*n* receptors are precisely the ones where our process control has to be tightest. The molecule's loudness and the curve's slope are different questions, and a careful library answers both.

Steepness is a tax on carelessness

Here is the derivation worth carrying out of this chapter, because it turns the vague feeling that "steep is dangerous" into a number you can budget against. Write the fractional occupancy of the receptor as $\theta = c^n / (EC_{50}^n + c^n)$, so that response $R = R_{max} \cdot \theta$. Because the sigmoid is natural on a logarithmic concentration axis, differentiate θ with respect to $\ln c$. The result is clean and famous:

$$\frac{d\theta}{d(\ln c)} = n \cdot \theta \cdot (1 - \theta)$$

The slope on the log axis is largest where $\theta(1 - \theta)$ is largest, which is at $\theta = \tfrac{1}{2}$ — exactly the EC_{50}. So the curve is steepest precisely at the half-point we most want to pin, a convenient and slightly cruel coincidence. Now ask the question a formulator actually cares about: if the concentration is off by one small percent, by how much is the response off? That is the elasticity, the ratio of fractional response change to fractional concentration change, and from the expression above it is

> $\frac{dR/R}{dc/c} = n \cdot (1 - \theta)$, which at the midpoint ($\theta = \tfrac{1}{2}$) equals $n / 2$.

A receptor with a Hill coefficient of 2, read near its EC_{50}, converts a 1% error in how much molecule the fruit made into a 1% error in the reported response; a steep cooperative channel with $n = 3$ converts the same 1% slip into 1.5%. The tax scales with the steepness of the very curve the loudest flavors tend to ride.

Hill coefficient n	Elasticity at EC_{50} (n/2)	Response error from a 1% dose error	Receptor character
0.8	0.40	0.40%	blunted / negatively cooperative
1.0	0.50	0.50%	single independent site
1.5	0.75	0.75%	mildly cooperative
2.0	1.00	1.00%	cooperative (TRPV1-like)
3.0	1.50	1.50%	steep, strongly cooperative

This is why we do not treat all eight axes with the same process tolerance. The gentle, near-linear axes forgive a little drift in content; the steep ones punish it, amplifying every wobble in the plant's chemistry into a wobble in the number we print. The breeder who wants a reproducible heat has to grow a more reproducible fruit than the one who wants a reproducible sweetness, and the Hill coefficient is the exchange rate between the two.

How we actually fit the curve

The precision of the whole library is born at the moment the curve is fit, and a careless fit poisons everything downstream of it, so we are deliberate about every step. We never read an EC_{50} off a single dose. We run a *series* of concentrations spaced evenly in the logarithm — each step a fixed *multiple* of the last, never a fixed addition — because the sigmoid is symmetric and well-behaved only on a log axis, and evenly spaced log-doses place the most measurements where the curve is steepest and the EC_{50} actually lives. The spacing is a design choice with arithmetic behind it: a dilution factor f lays down $1 / \log_{10} f$ points per decade of concentration. A half-log series, each dose about 3.16× the last, gives two points per decade — the honest workhorse for scouting a curve. When we already know roughly where the midpoint sits and want to pin it, we tighten to a quarter-log series, each dose about 1.78× the last, four points per decade, and bracket the expected EC_{50} with at least two measured doses on each side so the center is interpolated, never extrapolated.

To the bare Hill form we add two floating parameters, a bottom and a top, giving the four-parameter logistic that real data demand:

$$R(c) = \text{bottom} + (\text{top} - \text{bottom}) \cdot c^{n} / (EC_{50}^{n} + c^{n})$$

The bottom absorbs the receptor's baseline chatter when no ligand is present; the top absorbs the honest ceiling of the assay, which is rarely the textbook maximum. Four parameters — bottom, top, n, and EC_{50} — are fit at once to the whole dose series by nonlinear least squares, and the one we came for, the EC_{50} at the curve's center, comes back not as a bare number but as a number *with an interval*, the span within which the true half-maximal concentration almost certainly lies. We widen that interval honestly, by profile likelihood or by resampling the replicate reads, because a confidence interval quietly understated is the most common way a laboratory lies to itself about its own precision.



A read in progress on the stage — the EC_{50} is interpolated from the whole series, never read off one dose.

One more methodological truth, because it sets the currency the rest of the chapter is paid in. The uncertainty on an EC_{50} is *multiplicative*, not additive: a half-maximal concentration is a position on a log axis, so its error is naturally symmetric in log space and skewed in linear space, which is to say the EC_{50} is approximately log-normal. The consequence is that the right way to speak of its spread is not an

absolute ± in molarity but a *relative* figure — a coefficient of variation, the standard deviation divided by the mean, carried as a percentage. Every door we shut on variance in the rest of this book is measured in that same relative currency, and it is the currency in which the Index score, a ratio, does its arithmetic.

From a half-maximal concentration to a score

Now the step that turns a laboratory quantity into something a tomato, a chili, and a cup of tea can be read against on one page. The Index score on any axis is a ratio of two EC_{50} values — the axis's own fixed reference over the candidate — scaled by a thousand:

> **IFI = 1000 · EC_{50}(reference) / EC_{50}(candidate)**

Each axis carries its own anchor: a reference constant defined to score exactly 1,000 on that axis — a stated convention, like the Scoville scale's dilution-to-threshold definition, not derived from a literature survey. Because each axis is pegged to its own distinct reference compound and receptor system, an IFI score on the heat axis and one on the pungency axis each make a precise statement about their respective receptors, but they cannot be compared with each other directly; they are meaningful within their own axis alone. From there the reading is intuitive. A candidate louder than the reference has the smaller EC_{50}, so the ratio climbs above a thousand; a candidate quieter than its reference scores below it. Work one number all the way through. Our model predicts **IFI-H 4,200** for Phoenix (SS-0096) on heat, so

> 4,200 = 1000 · EC_{50}(ref) / EC_{50}(Phoenix) ⟹ EC_{50}(Phoenix) / EC_{50}(ref) = 1000 / 4,200 = **0.238**.

Our model predicts Phoenix would activate TRPV1 at just under a quarter of the concentration the native reference needs — 4.2× as potent. That single division is the whole translation from a prediction to a score, and it lets us rank the design ladder not as a pile of Scoville numbers but as a column of computed potencies, each one the inverse of a predicted half-maximal concentration.

CULTIVAR	SKU	SHU	IFI-H	EC_{50} (× REFERENCE)	SHU / IFI-H
Ember	SS-0087	866,000	1,570	0.637	551
Cinder	SS-0088	1,050,000	2,290	0.437	459
Kindle	SS-0089	1,250,000	2,000	0.500	625
Blaze	SS-0091	1,720,000	2,630	0.380	654
Wildfire	SS-0092	1,960,000	2,930	0.341	669
Scorch	SS-0093	2,210,000	3,220	0.311	686
Firestorm	SS-0094	2,470,000	3,490	0.287	708
Solara	SS-0095	2,720,000	3,980	0.251	683
Phoenix	SS-0096	2,937,948	4,200	0.238	700

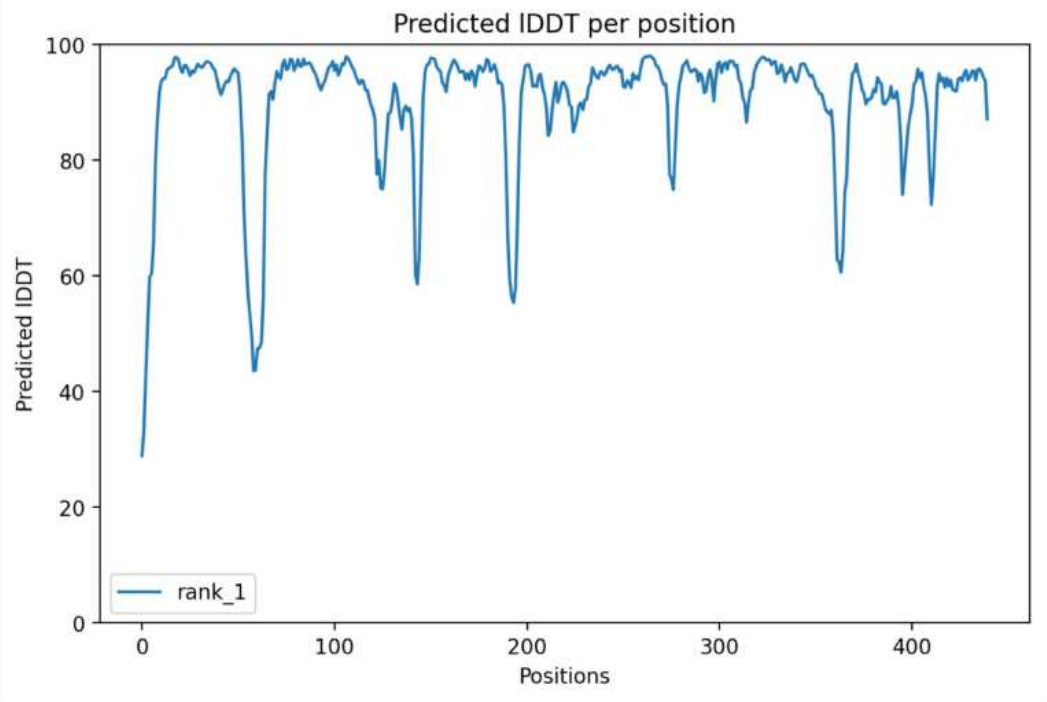
All SHU and IFI-H values in this table are computational predictions for our engineered cultivars — design targets, not certified measurements.

We leave Flare (SS-0090) out of the loudness column on purpose: its computational analysis on the heat axis is not yet complete, and we do not carry a prediction we have not fully computed. That restraint is the whole ethic of the ladder. The EC_{50} column is simply 1000 divided by the IFI-H of each row, and it falls monotonically as the fruit gets louder — from Ember needing 0.637 of a reference dose to Phoenix needing only 0.238. But the last column is where this chapter earns its title, so look at it closely.

What a Scoville number is not

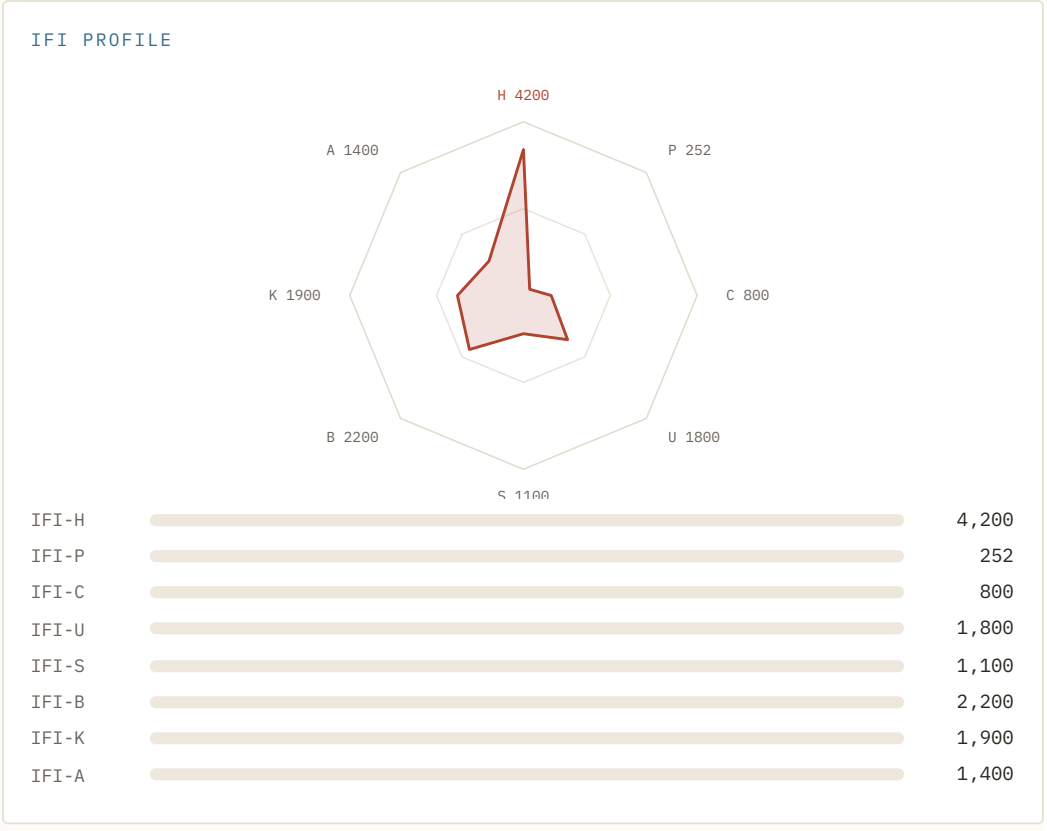
The rightmost column — Scoville Heat Units per point of loudness — should be a constant if SHU and IFI-H were the same thing. It is not. It drifts from 459 to 708 across the ladder, and in one place it does something a content measure should be unable to do: it inverts. Our model predicts Kindle would carry *more* capsaicinoid than Cinder — 1,250,000 SHU against 1,050,000 — and yet it is the *quieter* design, IFI-H 2,000 against Cinder's 2,290. More predicted content, less predicted loudness. That inversion is not an error in the table; it is the clearest possible demonstration that content and loudness measure different things.

The reason is mechanistic, and it lives in the enzyme. A Scoville number, read the modern way, is a *content* measure: the pungent fraction is separated and summed by HPLC, each capsaicinoid weighted by a fixed convention, and the total expressed as a dilution. Loudness is a *response* measure: the actual half-activation of the whole mixture on the channel. These need not track, because the effective potency of a blend is not a fixed-weight sum of masses — capsaicin and dihydrocapsaicin, differing only by one double bond in that eight-methyl acyl chain, carry slightly different affinities for the vanilloid pocket, and a fruit whose Pun1/AT3 acyltransferase pours out a heavier, more saturated analog pool can hold more total capsaicinoid yet speak more softly to TRPV1. The content is what the chromatogram sums. The loudness is what the receptor feels. The Index reports the second because the second is what the mouth actually experiences.



The Pun1/AT3 acyltransferase, coloured by model confidence — the active-site cleft where the capsaicinoid blend, and with it the gap between content and loudness, is set.

Because a formulator still has to move between the two worlds, it helps to work one unit conversion carefully, and to be honest about where it stops being exact. The Scoville convention fixes pure capsaicin at 16,000,000 SHU. Phoenix's predicted SHU of approximately 2.94 million gives 2,937,948 / 16,000,000 = **0.184**, so by the dilution definition its pungent fraction is designed to behave like an 18.4%-of-pure-capsaicin solution: threshold reached at roughly one part in 2.94 million, against one part in sixteen million for the neat molecule. To cross into the receptor's own molar world, carry capsaicin's molar mass of 305.4 g/mol: one part per million by mass is 1 mg/L, and (0.001 g/L) / (305.4 g/mol) = 3.27×10^{-6} mol/L, so **1 ppm capsaicin ≈ 3.27 µM**. These conversions let a content number be placed, approximately, on a concentration axis. But the honest limit is exactly the gap we have been describing: the conversion assumes the pungent fraction is a single analog, and real fruit is a blend, so the molar figure it yields is a content estimate and *not* the mixture's effective EC_{50}. That residue — the difference between what the arithmetic of content predicts and what the curve measures — is the reason we fit the curve at all rather than trusting the chromatogram.



The arithmetic of a ratio

The Index is a ratio of two EC_{50} values, and the instant you divide two numbers their *relative* uncertainties combine — not by taking the better of the two, not by summing them, but in quadrature. The derivation is short enough to carry in the head. Write the score as Q = A / B, where A is the reference EC_{50} and B the candidate's. Take logarithms: ln Q = ln A − ln B. Differentiate: dQ/Q = dA/A − dB/B. The

two error terms are independent, grown from different plates on different days, so when we square and average, the cross term vanishes and only the squares survive:

> **$\mathrm{CV(IFI)^2 \approx CV(EC_{50}, candidate)^2 + CV(EC_{50}, reference)^2}$**

The minus sign in the differential does not matter, because a variance squares it away; what remains is the plain Pythagorean addition of two relative scatters. This single line carries two lessons a formulator should never set down. Work them with numbers so their weight is felt. Suppose a candidate's EC_{50} is pinned to a relative precision of 3% and the reference standard to 2%. The score that rides on both inherits $\sqrt{(3^2 + 2^2)} = \sqrt{13} \approx$ **3.6%** — worse than either input, as a ratio always is. Now lavish every care on the plant and tighten the candidate to 2%: the score still cannot beat $\sqrt{(2^2 + 2^2)} = \sqrt{8} \approx$ **2.8%**, because the reference's own 2% has become a floor no amount of plant husbandry can lower.

CV OF CANDIDATE EC_{50}	CV OF REFERENCE EC_{50}	RESULTING CV OF THE IFI SCORE
1%	2%	2.2%
2%	2%	2.8%
3%	2%	3.6%
5%	2%	5.4%
7%	2%	7.3%
7%	7%	9.9%

The first lesson is that the precision of a score can never beat the precision of the reference it is measured against. If the anchor wobbles at 2%, nothing can be reported more tightly than 2%, however perfectly it was grown — which is why the reference standard is the most jealously kept object in the building, held still so that everything measured against it can be trusted. The second lesson is quieter and matters for how we publish our own headline. The batch coefficient of variation we hold our engineered output to — **≤7%** on capsaicinoid production, against the **42%** that wild-type fruit scatters across — is a CV on *content*. The CV on the reported *score* is a separate, slightly larger quantity, the content CV added in quadrature to the reference's share: two batches each at the 7% spec, read against each other, would scatter at $\sqrt{(7^2 + 7^2)} \approx 9.9\%$. Keeping a printed score inside 7% therefore demands each input tighter than 7% — both near 4.9%, since $\sqrt{(4.9^2 + 4.9^2)} \approx 6.9\%$. We say this plainly because conflating the two is how a careful lab quietly oversells its own reproducibility.

There is one more honest lever, and it is the square root. Random read-noise on a plate averages down with replicates: the CV of a mean of N wells is the single-read CV divided by $\sqrt{N}$. The arithmetic is unforgiving in a useful way — to halve the random contribution you need four times the wells, to quarter it sixteen times. Replication buys precision, but it buys it slowly, which is why we spend at least as much effort upstream, making a fruit whose chemistry does not scatter in the first place, as we spend downstream averaging the scatter away. That upstream work is where the enzyme returns to

the story. The two AI-designed substitutions in the shipped Pun1 construct — **S39L** and **L345G** — are designed to do more than lift peak output; they stabilize the catalysis. (Our structural modelling also identified C175S as a site where a substitution would prevent a disulfide misfold; that finding is a structural annotation from our models and was not carried forward into the final shipped design.) Together, S39L and L345G are predicted to leave fewer enzyme copies folding into variably active forms, so the capsaicinoid output of one plant should resemble the next, narrowing the distribution of EC_{50} values the harvest returns. The projected drop from a 42% wild-type CV to a ≤7% engineered CV is, at its root, a molecular story about a protein designed to fold the same way every time. Consistency is not imposed on the fruit at the loading dock. It is designed into the fruit at the active site.

Eight receptors, eight half-maximal reads

A full-spectrum cultivar's design spans all eight axes at once, and Phoenix is the clearest example we have brought to this stage. Its predicted profile expresses eight computed EC_{50} estimates as eight IFI scores, and every one of them can be turned back into a predicted half-maximal concentration by the same division — 1000 over the score — so the profile is, underneath, a set of eight predicted potencies, each expressed against that axis's own reference constant.

AXIS	RECEPTOR	IFI	EC_{50} (× REFERENCE)
Heat (H)	TRPV1	4,200	0.238
Pungency (P)	TRPA1	252	3.97
Cool-Burn (C)	TRPM8	800	1.250
Umami (U)	mGluR1	1,800	0.556
Sweetness (S)	TAS1R2 / R3	1,100	0.909
Zero-Bitter (B)	TAS2R38	2,200	0.455
Kokumi (K)	CaSR	1,900	0.526
Astringency (A)	Salivary PRPs	1,400	0.714

The table teaches the reading in one glance. Heat towers because our model predicts Phoenix would activate TRPV1 at 0.238 of a reference dose. But look at cool-burn: a score of 800, below a thousand, means its predicted EC_{50} is *1.25×* the reference — our design predicts Phoenix would speak to the cold-sensing TRPM8 channel more softly than the cooling reference does, as a pepper honestly should. We do not inflate the quiet axes to make the profile look complete; a sub-1000 score is a true statement that the design is quieter than the anchor on that axis, and printing it is part of the discipline. Each of these eight predicted scores would, in a live assay, carry its own confidence interval and its own quadrature with its own reference; a full-spectrum predicted profile is not one estimate but eight independent ones, eight curves modelled, eight midpoints computed, eight relative scatters propagated. The radar

that renders them beautifully on the page is, underneath its symmetry, a very patient stack of computed dose–response predictions.

The honest edges

A chapter on what we measure owes a plain account of where the measurement frays, because every one of these edges is a way the EC_{50} can mislead a lab that trusts it too far. The first is extrapolation: an EC_{50} is only as real as the doses that bracket it, and a curve asked for a half-point outside the concentrations actually tested is a guess wearing a decimal. We require measured doses on both shoulders of the midpoint and treat any EC_{50} that drifts to the edge of the tested range as a signal to redesign the series, not to publish. The second is the ceiling: the "top" of the four-parameter fit is the assay's honest maximum, not the receptor's, and a plate that saturates early can flatter an EC_{50} by shrinking the span the midpoint divides — so we watch the top as closely as the center. The third is the assumption of independence behind the quadrature rule. If the candidate and the reference are read on the same plate on the same afternoon, a shared drift — a warm incubator, an aging batch of dye — correlates their errors, and correlated errors do not add in neat squares; they can add almost linearly, inflating the true CV past what $\sqrt{(a^2 + b^2)}$ promises. We break that correlation on purpose, staggering references across days and plates, because the whole arithmetic of the score rests on the two errors being strangers to each other.

Two deeper edges are worth naming because they separate a measurement that is merely *repeatable* from one that is *true*. The coefficient of variation we drive down is a *technical* CV — the scatter of the instrument and the assay — and it is not the same as the *biological* CV, the real plant-to-plant variation in the fruit itself. A library can polish its technical CV to a mirror and still be reporting a fruit that genuinely varies; honesty means carrying both numbers and never letting the tidy one hide the stubborn one. And the Hill coefficient itself can mislead when a receptor's response is not a simple monotone climb — a biphasic or bell-shaped curve, where very high concentrations begin to *suppress* the response, will fool a naïve four-parameter fit into reporting an EC_{50} that no single potency explains. For those receptors we fit the fuller model and say so, rather than forcing an S onto a shape that is not one.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

None of these edges is a reason for less reverence; they are the reason for more. The sigmoid is not a formula we imposed on taste — it is the shape taste has, the honest geometry of a receptor filling up, and the EC_{50} is the one point on it where the channel gives exactly half of itself. Every fruit our family grows is, in the end, an argument made in that geometry: that a tomato can speak umami more loudly, that a pepper can carry its heat at a lower dose and a tighter scatter, that the nourishment we mean to pass to the body can be measured and therefore designed rather than merely hoped for. We spend this whole book narrowing the cloud of EC_{50} values

the harvest hands us, and we do it because a number held steady is a kind of care — the same care that chooses to feed a body well, now made exact enough to keep across a thousand harvests and still taste like the thing we loved.



PART II OF V

Shutting the Doors

the genetic anchor · the room · the process · the assay and the ruler's own wobble · the reference standard

Chapter 5 · The genetic anchor — *why one genotype starts tighter than a crowd.*

Split an heirloom pepper open across a field at harvest and you are not holding one variety. You are holding a crowd. A landrace is a population — a soft cloud of genotypes, each plant a little different in how it builds capsaicin, each answering the same summer in its own dialect. One plant leans on a slightly more active allele of the acyltransferase; its neighbor carries a quieter copy and makes it up with a warmer microclimate; a third sits in shade and splits the difference. That is the romance of the old seed, and it is also its ruin: measure the heat across such a field and the coefficient of variation can run thirty, forty percent before a single thing has gone wrong in the growing. Our own measurements put a wild-type population near **forty-two percent** — the plants were never trying to hit a number. They were only trying, each in its own way, to be a kind of plant. You cannot write a specification on a crowd.

Our cultivar is one genotype, written once and carried identically into every seed. The **two substitutions** we placed in the *Pun1* acyltransferase of Phoenix — **S39L · L345G** — are not *typical* of the plant; they are *in* the plant, every plant, letter for letter, placed by CRISPR after a protein language model and a genome-scale flux-balance search had agreed, in silico, that they would lift catalytic efficiency and fold stability without buckling the structure. Because the edits are fixed homozygous — the same allele on both copies of the chromosome — they do not segregate, do not reshuffle, do not reappear in one fruit and vanish in the next. The widest door in all of plant variation — the question of whether this fruit and that one even share a set of instructions — we close before the first seed is sown. The chapters of this book are each about holding one door shut: the growing room, the ruler, the clean line of a process. This is the door that comes first, and the only one we can shut all the way.

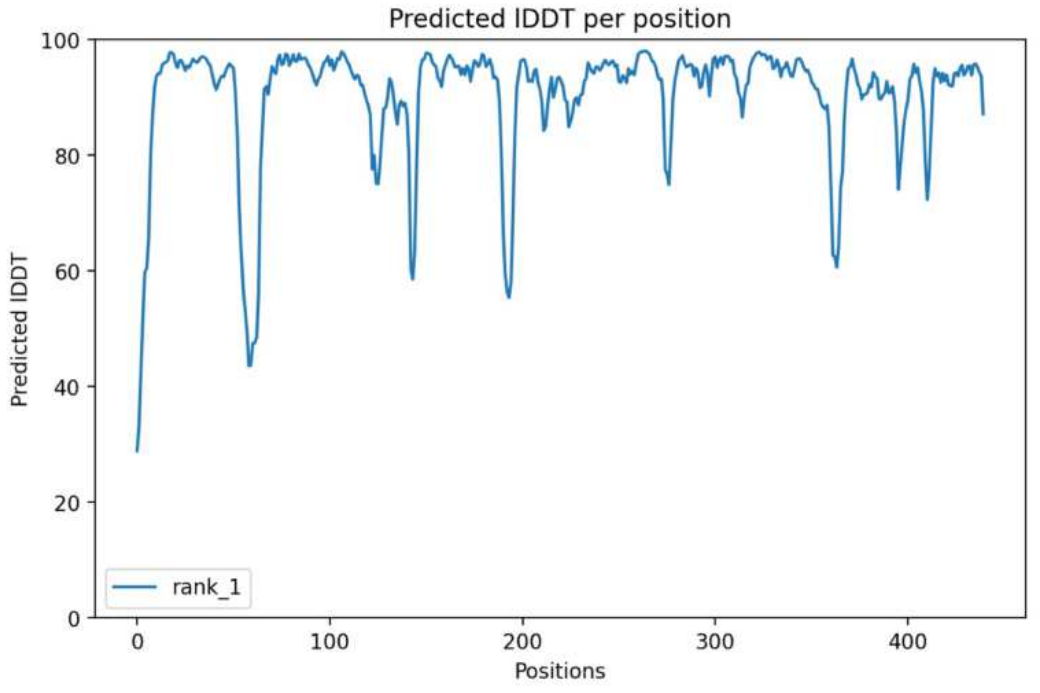


FIG. 5.1 — the *Pun1* / AT3 acyltransferase model, coloured by pLDDT confidence; the binding domain and the active-site cleft, where our two edits sit, resolve at high confidence.

The variance budget — why the genetic term is the deepest lever

To see why the genotype is the deepest lever and not merely the first one, you have to write the heat of a fruit as a sum of where its scatter comes from. A measured heat value is a true genetic set-point plus what the growing did to it plus what the measuring added, and when those three sources are independent they do not add as themselves — they add as their squares. Variance is additive; the standard deviation is not. In the language of the coefficient of variation, which is the standard deviation expressed as a percentage of the mean, that quadrature reads:

> `CV²(total) = CV²(genetic) + CV²(environmental) + CV²(measurement)`

Fill it with what a wild-type field actually gives us. Call the genotype-to-genotype scatter forty percent, the growing twelve, the analytical measurement five. Then the total is $\sqrt{(40^2 + 12^2 + 5^2)} = \sqrt{(1600 + 144 + 25)} = \sqrt{1769} =$ **42.1%** — which is the forty-two percent we measure, arrived at honestly from its parts. Now look at the parts as variance rather than as CV, because variance is the thing that decomposes. The genetic term is 1600 of 1769, which is **90.4 percent of all the scatter in the field.** Not forty of forty-two. Ninety of a hundred. The squaring is the whole lesson: a source that is merely the largest in CV is overwhelmingly the largest in variance, and variance is the currency a specification is written in.

That is what it means to call the genotype the deepest lever. Drive the genetic term to zero and the field does not get a little tighter — it loses nine-tenths of its variance in one move, and the floor drops from forty-two percent to $\sqrt{(0 + 144 + 25)} = \sqrt{169} =$ **13.0%** with nothing else touched. The growing is the same; the ruler is the same; only the crowd is gone. From thirteen, ordinary discipline finishes the job: hold the room so the environmental term falls from twelve to four, hold the assay so the measurement term falls from five to three, and the total lands at $\sqrt{(0 + 16 + 9)} = \sqrt{25} =$ **5.0%** — inside the ≤7%

we promise a formulator. The genetic anchor does not do that last work. It makes that last work *possible*, by removing the term that no growing room and no instrument could ever have reached.

SOURCE OF SCATTER	WILD-TYPE FIELD (CV)	ISOGENIC, SAME GROWING	ISOGENIC + TIGHTENED PROCESS
Genetic — genotype to genotype	40%	**0%**	**0%**
Environmental — the growing	12%	12%	4%
Measurement — the assay	5%	5%	3%
Total (added in quadrature)	**42.1%**	**13.0%**	**5.0%**
Genetic share of total *variance*	90.4%	0%	0%

Read the second column as the thesis of the chapter stated as arithmetic. We did not reduce the largest term. We deleted it, and let the smaller ones stand exposed where process control can reach them. Averaging cannot do this. Measure ten plants from a variable population and the *mean's* uncertainty falls by $\sqrt{10} \approx 3.2$, but the next batch's mean still sits wherever its particular draw of genotypes put it — the genetic term re-enters as a systematic shift from batch to batch that no amount of within-batch replication removes. We are not averaging a crowd and hoping the average holds steady. We are repeating one fruit, and a repeated fruit has no crowd to average.

Reading the enzyme that writes the heat

The genotype earns its place at the head of the budget because of what a single gene product does, molecule by molecule, in the placenta of the fruit where capsaicin is made. Two supply lines meet there. The first is the phenylpropanoid branch that runs in every vascular plant — phenylalanine handed down through PAL, C4H, 4CL, HCT, and COMT to a vanilla-scented aldehyde, then aminated by pAMT to **vanillylamine**, the nitrogen-bearing head of the molecule. The second is a branched-chain fatty-acyl line that builds **8-methyl-6-nonenoyl-CoA**, the forked tail. The two are strangers until one enzyme introduces them: the *Pun1* locus encodes **AT3, the capsaicin synthase**, an acyltransferase that clamps the acyl-CoA and the amine in a single cleft and forms the amide bond that is capsaicin. It is the terminal step — the step that names the molecule. Our flux-balance analysis predicts, however, that the rate is governed upstream, at 4CL and HCT/CCoAOMT in the phenylpropanoid pathway, well before the amine arrives at AT3. Pun1 certifies the product; the upstream supply lines set the pace. The engineered edits in AT3 are therefore not an attempt to open a throttle but to stabilise the terminal synthase — ensuring the condensation runs cleanly and reproducibly each time the substrates arrive.

Capsaicin then leaves the fruit and finds the eater at **TRPV1**, a non-selective cation channel on the trigeminal nerve endings of the tongue — the same channel opened by temperatures above 43 °C and by acid. The molecule warms nothing. Its vanillyl head slips into a pocket between the channel's third and fourth transmembrane

helices and lowers the thermal threshold, so a channel that should have waited for real heat opens at body temperature and reports a fire that is not there. The length and saturation of that forked acyl tail set the balance between capsaicin and its saturated sibling dihydrocapsaicin, and with it the *shape* of the burn — how fast it arrives, how long it stays. This is why two fruits can carry the identical total capsaicinoid load and still speak at different volumes, and why a fruit's Scoville number and its Index score need never move in lockstep: the chemistry counts molecules; the receptor counts how loudly they knock.

Our **two edits** live in AT3, and not one of them was chosen to make the burn louder for its own sake. Each was chosen because, among the substitutions a 650-million-parameter protein language model scored across every residue, these two raised catalytic stability *and* left the fold more stable than they found it — and stability, not speed, is the quiet currency of consistency. The measure is **ΔΔG**, the change the edit makes to the free energy of folding, in kilocalories per mole. The sign convention is worth stating plainly because the whole argument turns on it: a *negative* ΔΔG means the edited protein folds to a *lower*, more favorable energy than the original — it is more stable, not merely faster. Both of ours are negative.

EDIT	RESIDUE CHANGE	POSITION	STRUCTURAL LOCALE	WHAT IT DOES TO THE ENZYME	ΔΔG (KCAL·MOL⁻¹)	FOLDING EQUILIBRIUM SHIFT *(EXP −ΔΔG/RT, 298 K)*
S39L	Ser → Leu	39	N-terminal domain	a hydrophobic side chain tightens substrate binding	−2.1	≈ 35×
L345G	Leu → Gly	345	active-site cleft	a smaller residue relieves steric clash, cleaner turnover	−1.8	≈ 21×
C175S (structural annotation — not in shipped construct)	Cys → Ser	175	core	removes a reactive thiol that could form a non-native disulfide	−0.9	≈ 4.6×
Combined (S39L + L345G)	—	—	whole fold	the two stabilizations compound	**−3.9**	**≈ 730×**

The last column is the chapter's hinge, and it is a short worked example in thermodynamics rather than a flourish. The fraction of enzyme that is correctly folded at equilibrium is governed by a Boltzmann factor: the ratio of the folding equilibrium constant between edited and original protein is $\exp(-\Delta\Delta G / RT)$, where RT at 298 K (25 °C) is 1.987 cal·mol^{-1}·K^{-1} × 298 K = 0.592 kcal·mol^{-1}. Feed each edit through it. S39L: $\exp(2.1 / 0.592) = \exp(3.55) \approx 35$. L345G: $\exp(1.8 / 0.592) \approx 21$. Because free energies add, the two multiply:

exp(3.9 / 0.592) = exp(6.59) ≈ **730**. The edited AT3 sits in its folded state roughly seven hundred times more decisively than the enzyme we started from.

Why does that buy consistency? Because a marginally stable enzyme is a weathervane. If the wild-type protein were, say, ninety percent folded under the ordinary stress of a warm afternoon — a tenth of it slumped and idle — then the amount of capsaicin a cell makes would swing with every wobble in temperature, pH, and crowding, and those swings are exactly the plant-to-plant scatter a formulator cannot have. Push the folding equilibrium seven hundred-fold toward the folded state and that same protein sits past 99.99 percent folded, and, far more to the point, *stays* there when the cell's weather turns. The active fraction stops reporting the afternoon. We did not design for a headline number at the level of the fruit. We designed for an enzyme that folds the same way in a cold room and a warm one, in a vigorous plant and a struggling one — which is to say we designed for consistency at the level of a single chemical bond, long before any fruit existed to be measured.

From a bond to a batch — carrying the fold's steadiness onto the tongue

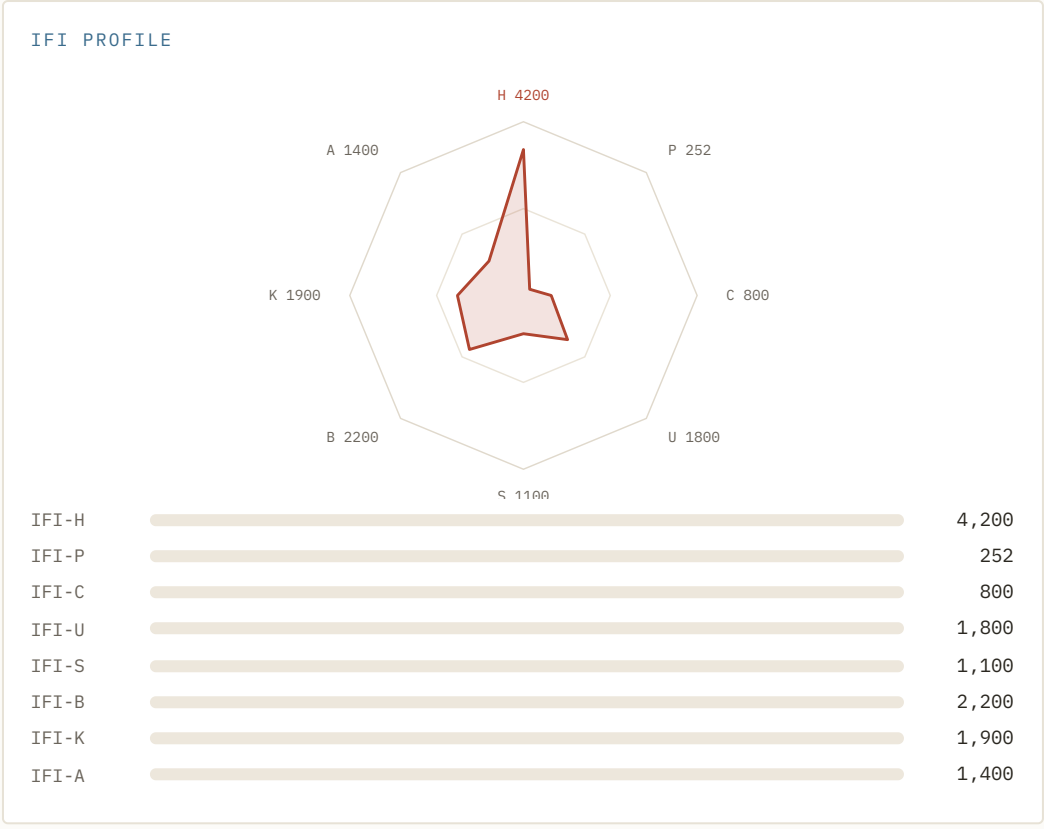
A stable enzyme is a promise about the fruit; the Iyengar Flavor Index is where that promise becomes a number you can read. The Index is a stated convention — defined, like the Scoville scale's dilution-to-threshold definition, by a fixed reference point on each axis, not derived from a literature survey of published EC_{50} values. It does not ask how much of a compound a food holds — it asks how *loudly* that compound speaks at its receptor, and it anchors each axis to the half-maximal activation of its own defined reference constant:

> $IFI(candidate) = 1{,}000 \times EC_{50}(reference) \div EC_{50}(candidate)$

EC_{50} is the concentration at which a stimulus drives its receptor to half of full response. A smaller EC_{50} means a louder stimulus — less is needed to be heard — so the reference sits in the denominator and loudness rises as a candidate's EC_{50} falls. Work it through with the heat axis. Fix the reference at an EC_{50} of **10.0 µM** of capsaicin-equivalent at TRPV1 — the convention constant chosen so the reference scores exactly 1,000, since IFI(ref) = 1,000 × 10.0 ÷ 10.0 = 1,000 by definition. Our model places Phoenix at **IFI-H 4,200**. Invert the formula to recover the physiology it implies: the model implies an effective EC_{50}(Phoenix) = 1,000 × 10.0 ÷ 4,200 = **2.38 µM**. In the model, Phoenix's extract drives the channel to half-maximal firing at 2.38 µM where the reference needs 10.0 — the dose-response curve has slid left by a factor of 4.2, and the whole burn arrives for a quarter of the material. A faint shishito near IFI-H 40 runs the arithmetic the other way: EC_{50} = 1,000 × 10.0 ÷ 40 = 250 µM, a channel that will not half-open until you flood it.

Those are the units a lab actually pipettes, and the conversion is worth showing because it is the kind of quiet step where errors hide. Capsaicin's molar mass is 305.4 g·mol^{-1}, so 1 µM is 0.3054 µg·mL^{-1}. The reference's 10.0 µM is therefore 3.05 µg·mL^{-1}; Phoenix's model-

implied 2.38 µM is 0.727 µg·mL^{-1}. The response itself follows a Hill curve, $R = R_{max} \times [C]^n \div (EC_{50}^n + [C]^n)$, with a capsaicin–TRPV1 cooperativity near n = 2; at $[C] = EC_{50}$ the second term's numerator and denominator meet and the response is exactly half of maximum, which is all "EC_{50}" ever means. The Index takes that one honest point on each curve and places each compound on its own axis, comparable within that axis — a hot sauce on the heat scale, a fresh tomato on the umami scale — against the same defined reference point for that axis. Because each axis is anchored to its own convention constant, the axes do not collapse to a single number; a formulator reads them side by side, not as a single rank.



Phoenix's computed profile is read that way across all eight axes at once — the first cultivar we designed to register on every one of them. But heat is where the Index and the older chemistry pull apart, and a formulator should see exactly where. The Scoville scale measures total capsaicinoid *concentration*; IFI-H measures receptor *loudness*. They are proportional only if every fruit carries the same mix of capsaicin and dihydrocapsaicin, and no two of ours do. Line the flagship ladder up and divide one by the other, and the ratio refuses to hold still.

CULTIVAR	SKU	PREDICTED SHU	IFI-H	AXES	PREDICTED SHU PER IFI-H POINT
Ember	SS-0087	866,000	1,570	2/8	552
Blaze	SS-0091	1,720,000	2,630	4/8	654
Wildfire	SS-0092	1,960,000	2,930	5/8	669
Scorch	SS-0093	2,210,000	3,220	5/8	686
Firestorm	SS-0094	2,470,000	3,490	6/8	708
Solara	SS-0095	2,720,000	3,980	6/8	683
Phoenix	**SS-0096**	**~2,950,000**	**4,200**	**8/8**	**~700**

All SHU values are computational predictions from our genome-scale model, not certified measurements.

The right-hand column is the fingerprint of the acyl tail doing its work: a fruit can carry more total capsaicinoid per unit of loudness, or less, depending on how its AT3 balanced the saturated and unsaturated chains. That is a feature, not a flaw in the ruler — SHU answers the chemist's question and IFI-H answers the eater's, and keeping both lets us tune the *shape* of a burn, not only its size. It is also one more reason to fix the genotype: both measures ride on the same enzyme, so a crowd of AT3 variants would scatter the predicted SHU and the IFI-H *and* the ratio between them. One genotype steadies all three at once.

Fixed homozygous — welding the door, not leaning on it

The phrase *fixed homozygous* carries more weight than it looks, and it is the mechanism by which the genetic door is not merely narrowed but welded shut. A plant carries two copies of each gene, one from each parent. A trait held on only one copy is a trait that gambles. Self-pollinate a plant heterozygous at a locus — one edited copy, one not — and classical segregation deals the allele out to the next generation as a lottery: one quarter of the offspring receive two edited copies, one half receive a single copy, one quarter receive none. The dosage across that progeny is 0, 1, or 2 with probabilities ¼, ½, ¼, a mean of one copy and a **variance of exactly ½** — a half-unit of scatter injected into the population by heterozygosity alone, reborn in full every single generation, upstream of any growing room and untouchable by it. That is variance you inherit whether the summer is kind or not.

An edit fixed on *both* copies ends the lottery. There is nothing left to segregate; every seed, every generation, carries the identical pair, and the trait stops being a thing the plant might inherit and becomes a thing the plant simply *is*. The difficulty of reaching that state by breeding is itself arithmetic, and it explains why the old methods could circle a target for decades without landing on it. To co-fix even our **two edits** by selfing a double heterozygote, the chance that a single offspring draws both as homozygotes at once is $(¼)^2 = 1/16$ — fewer than seven plants in a hundred, and that is the generous case where the two sit on different chromosomes and are already present to be drawn. A breeder is always chasing an allele through a population, always leaving some heterozygosity behind, always re-opening the door a crack with each cross. We close it once, write the edit homozygous at the target in a single engineered step, and then *prove* it: every advanced line is sequenced to confirm both

copies carry the edit before it is ever propagated, so the genetic term cannot drift back over generations the way a selected trait silently does.



FIG. 5.2 — the sequencer bench, where zygosity is confirmed before a line advances; an edit is not "fixed" in our records until both chromosomal copies have been read.

Set the three kinds of starting material side by side and the floor each begins from is plain — and the floor is the whole point, because consistency work can only ever tighten what the genetics first hands it.

STARTING MATERIAL	WHAT IT IS, GENETICALLY	TYPICAL HEAT CV BEFORE ANY GROWING CONTROL
Landrace / open-pollinated	a population; a cloud of genotypes	~42%
Bred elite line	narrowed, but residually heterozygous	often 15–25%
Isogenic engineered line (fixed homozygous)	one genotype, verified identical	the genetic term ≈ 0

Read the bottom row as the thesis of the chapter. The number is not "small." It is, to the limits of our ability to sequence and confirm it, *absent* — the one term in the sum of squares we can drive out entirely rather than merely reduce. Every other source of scatter in the system is a matter of degree: a room held tighter, a ruler held steadier, a process held cleaner. This one alone is a matter of kind.

What the anchor does not do — the honest floor

A book for formulators owes the limits as plainly as the wins, and the genetic anchor has four honest ones. The first is written into the variance budget itself: fixing the genotype drives CV(genetic) to zero and touches *nothing else*. The environmental term and the measurement term are exactly where they were. An isogenic line grown carelessly, in a room whose light and water and heat wander, will still scatter — thirteen percent in our worked field, and worse in a bad one. The anchor is necessary for CV ≤7%; it is nowhere near sufficient. It buys the floor and not the ceiling, and every other chapter of this book exists because the floor is only the beginning.

The second limit is that *isogenic is not identical in expression*. A single genotype still carries a reaction norm — one set of instructions that reads the environment and answers it — so two genetically identical plants in genuinely different conditions will make genuinely different amounts of capsaicin. What fixing the genotype removes is the genotype-by-environment *interaction* as a population term: with only one genotype there is no second genotype for the environment to treat differently, so the scatter collapses onto a single response curve instead of a fan of them. Epigenetic state — methylation, chromatin, the memory a plant keeps of its own stresses — can still vary along that one curve. It is bounded by the environment now, not multiplied by the genetics, and that is a smaller and more honest thing to manage.

The third limit is that identity has to be *maintained*, not assumed. Over enough cell divisions a somatic mutation can arise in a branch; an outcross in a careless seed increase can slip a foreign allele back in; a CRISPR edit can leave an off-target change elsewhere in the genome. None of these is hypothetical, and each is answered by the same discipline rather than by faith: whole-genome sequencing to clear off-targets before a line is named, controlled self-pollination under isolation to keep the seed true, and re-sequencing of the zygosity at the target before propagation. "Verified identical" is a claim we re-earn, not a state we reach once.

The fourth limit is the quietest and the most absolute. Even a perfect, genetically silent material cannot be measured below the reproducibility of the instrument reading it — the analytical CV is a hard floor under everything, which is why the measurement term never fell past three in our budget and never will fall to zero. You can anchor the genotype and hold the room and clean the process, and the ruler still has a tremor in its hand. We state it because a number without its error bar is a kind of dishonesty, and this book is built on refusing that one.



FIG. 5.3 — Phoenix in the controlled chamber: one genotype, carried identically into every plant, so that what the chamber measures is the growing and the ruler, never the crowd.

None of those four limits is an argument against the anchor; they are the shape of what the anchor leaves for the rest of the work to do, and they are small precisely because the genotype was fixed first. We are a family that has spent its life paying close attention to what goes into a body — growing things we trust, measuring them until the trust is earned, and setting a table we would feed to our own. That care is not sentiment here; it is the reason the math is shown

rather than waved, because a number you can check is the most respectful thing a grower can hand a cook. These are computational designs — the fold stabilities, the IFI scores, the SHU estimates are all predictions from our models; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory. We do not manage the genetic variance of our cultivars. Having written it out, letter by verified letter, sequenced on both copies before a single seed is sown, we no longer have any to manage. The field is no longer a crowd. It is one fruit, repeated — and everything after this chapter is the steady, answerable work of keeping it that way.

Chapter 6 · The room — *controlled environment as a variance tool.*

There is an hour before the lights come up in a growing room when the air is still and faintly green and nothing moves, and it is the most honest hour we know. The field gave our grandparents weather — a gift and a gamble, a kind August and a cruel one, written into every fruit whether they wanted it there or not. A vintage, the winemakers call it, and they are right to: it is the signature of a year that will never come again. The room gives us something the field never could. It gives us a specification, because in the room there are no years. There is only the same measured dawn, repeated.



Dew at the measured dawn — the room's gift is that this morning is exactly last morning, so the fruit's fullness is a standing fact and not a lucky season.

A flavor you can taste is a phenotype — an outcome, the visible end of a long private argument between a plant's instructions and its circumstances. And a phenotype wanders. Pick the same cultivar from the same row on two August mornings a year apart and the two fruits will not score the same, because the thing you are tasting is never only the recipe. It is the recipe as executed under one particular sky. This chapter is about the circumstances half of that argument, and about how you win it — how you take a flavor that happens to arrive and turn it into a flavor that is guaranteed to.

Two doors between a seed and a flavor

A century of quantitative genetics gives us one plain equation for how much a phenotype drifts from fruit to fruit, and everything in this room is built to act on it. The total variance you taste is a sum of independent variances:

```
σ²(total) = σ²(genetic) + σ²(environment) + σ²(G×E) + σ²(noise)
```

Read it term by term and the whole strategy falls out. σ²(genetic) is the spread that comes from the plants not being genetically identical — the lottery of a seed population. σ²(environment) is the spread that comes from no two plants, or no two cycles, ever having been grown exactly alike. σ²(G×E) is the interaction, the maddening term

where a genotype that wins under one sky loses under another. And σ^2(noise) is the irreducible biological hum — the stochastic flutter of gene expression and development that remains even when everything you can name is held still.

Two doors stand between a seed and the number on the tongue. Behind the first door is **genetics**, and it fixes *what* the plant will build — the recipe, the set of enzymes, the ceiling on how loud each axis can ever become. We shut that door with design: an engineered line is effectively one genotype, copied, so σ^2(genetic) collapses toward zero and the plants stop disagreeing with each other about the recipe. Behind the second door is the **environment**, and it fixes *how much* of that recipe actually arrives — how fully the fixed instructions are expressed. A tomato written to IFI-U 2,600 will reach 2,600 under generous light and fall short under mean light, not because its instructions changed but because it was handed less to work with. This chapter is the second door, and shutting it is pure variance work.

We report the spread not as a bare standard deviation but as a coefficient of variation, because the units of flavor span three orders of magnitude and a raw σ tells you nothing without its mean beside it:

```
CV = σ ÷ μ (written as a percent)
```

A ±50-unit wobble is catastrophic on a shishito near 40 and invisible on a superhot near 4,200; dividing by the mean is what lets a single tolerance govern the whole ladder. A wild *Capsicum*, left to the old arrangement of genetics-plus-weather, carries a batch-to-batch CV near **42%** — nearly half the signal is noise, and no formulator can build on a thing that moves that much. The tolerance we hold our own fruit to is **≤7%**. Closing the distance from 42 to 7 is the labor of both doors together; the first door is sequence, and the second door is this room.

The light budget

Of every lever in the room the daily light integral is the one that moves the number most, and the one a grower most often lets wander. It is worth its own worked line, because most of holding a flavor flat is simply holding this quantity flat. Light is the carbon budget. Every molecule of capsaicin, every glutamate that will one day flood a receptor, every tannin and every sugar is built from carbon the plant fixed from the air using photons as the currency — so the total photons delivered in a day sets the size of the purse the plant has to spend on flavor.



Golden-hour in the research greenhouse — each photon landing on this canopy is one unit of the day's building budget, and the budget is the thing we hold flat from cycle to cycle.

DLI is only a multiplication. Take the photon flux arriving at the canopy — the photosynthetic photon flux density, PPFD, in micromoles of photons per square metre per second — multiply by the seconds the lamps are lit, and divide by a million to carry micromoles up into moles:

```
DLI (mol·m⁻²·d⁻¹) = PPFD (µmol·m⁻²·s⁻¹) × photoperiod (s) ÷ 1,000,000
```

A canopy held at 400 µmol·m^{-2}·s^{-1} across an eighteen-hour day collects 400 × (18 × 3,600) ÷ 1,000,000 = **25.9 mol·m^{-2}·d^{-1}**. Now watch what a little drift costs. Let the intensity sag to 350 and the day's budget falls to 350 × 64,800 ÷ 1,000,000 = 22.7 mol·m^{-2}·d^{-1}. Let the photoperiod also slip to sixteen hours and it drops to 350 × 57,600 ÷ 1,000,000 = **20.2 mol·m^{-2}·d^{-1}** — a purse 22% smaller than the one we meant to spend. The plant was never told to build a lesser fruit. It was simply handed four-fifths of the carbon and asked to execute the same recipe, and it did, four-fifths as loudly.

This is why the single discipline of pinning DLI carries most of the consistency load. Not because light is the only thing that matters, but because in the light-limited regime where a dense canopy lives, accumulated compound rides very nearly in proportion to the integrated light. Hold the budget flat cycle after cycle and you have held flat the one input that most of the flavor is made from. Let it breathe ±5 mol·m^{-2}·d^{-1} and you have built a σ²(environment) into every lot before a single other variable has moved.

From photons to the number

Here the budget becomes the score, and it is worth following the whole chain, because the woven truth of the Index is that a number on the tongue is a receptor event you can trace all the way back to a photon. Read a food's umami the way we read every axis: take an extract, dilute it step by step, and find the dilution at which it drives the receptor — for savor, the metabotropic glutamate receptor mGluR1 — to exactly half of its maximum response. That half-activation is the anchor of the whole system. A receptor does not respond linearly; it follows a sigmoid, the Hill relation,

```
response = R(max) × [L]ⁿ ÷ (EC₅₀ⁿ + [L]ⁿ),
```

where [L] is the stimulus concentration, n is the Hill slope that sets how steeply the curve turns, and EC_{50} is the concentration at the curve's half-height. Set [L] = EC_{50} and the fraction becomes $EC_{50}^{n} \div (EC_{50}^{n} + EC_{50}^{n}) = ½$, which is the definition made visible: EC_{50} is simply where the receptor is half-answered. The smaller a food's EC_{50} — the less of it needed to reach that half-answer — the louder it speaks. The Index is that number inverted against a fixed reference and scaled by a thousand so a tomato, a chili, and a cup of tea can be read on one page:

```
IFI(candidate) = 1,000 × EC50(reference) ÷ EC50(candidate)
```

Run the arithmetic on a restored tomato written to **IFI-U 2,600**. Rearranging, its half-activating concentration is EC_{50}(tomato) = 1,000 × EC_{50}(ref) ÷ 2,600 = EC_{50}(ref) ÷ 2.6 — it reaches the receptor's half-answer at one 2.6th of the reference concentration, a fruit 2.6 times louder in savor than the anchor it is scored against. The wild tomato the long supply chains and the yield-first breeders left us sits near **IFI-U 300**: EC_{50}(ref) ÷ 0.3, more than three times the concentration for the same half-activation, a tomato you have to eat a great deal of to taste at all. The edits we wrote are designed to restore the plant's own glutamate route — glutamate dehydrogenase, GDH1, and glutamate decarboxylase, GAD, rebuilding the pool that decades of shipping and shelf-life selection bred out — and in our model lift 300 to 2,600, a food roughly **nine times louder** in umami than the one in the grocery aisle.

But the sequence sets the *target*, not the delivery, and this is the whole reason the room exists. The measured score is read off the fruit you actually grew, and that fruit carries only the glutamate it had the carbon to build. Umami traces a clean chain backward into the light: photons fix carbon, the TCA cycle hands off the five-carbon skeleton α-ketoglutarate, GDH1 aminates it to glutamate, glutamate floods mGluR1, and the receptor reports savor. Every link downstream of the first is written in the genes and runs identically in every seed. The first link — how much carbon there was to spend — belongs to the environment. Dilute a fruit's accumulated glutamate by a fifth and you must concentrate its extract by a fifth to reach the same half-activation, so EC_{50} rises by the factor 1 ÷ 0.8 = 1.25 and the delivered Index falls by 0.8. A tomato written to 2,600 but grown under four-fifths of its light does not taste like a different cultivar. It tastes like the same cultivar, quieter — **2,600 × 0.8 = 2,080** — a fifth of the savor we designed, lost not at the bench but at the lamp. Pin the budget and that loss never happens; the number designed is the number delivered.

Heat at the root

Temperature does not feed the plant; it sets the speed of every reaction inside it, and so it governs not how large the budget is but how fast and how faithfully the budget is spent. Enzymes are catalysts with a temperature sensitivity that chemists summarize in a single ratio, Q_{10} — the factor by which a reaction rate changes for every ten-degree shift:

```
rate(T₂) ÷ rate(T₁) = Q₁₀^((T₂ − T₁) ÷ 10),
```

with Q_{10} sitting near 2 for most of the enzymes that matter here. That exponent is unforgiving. A ten-degree rise doubles a rate; a five-degree rise multiplies it by 2^0.5 = 1.41, a 41% swing; even a two-degree drift moves it by 2^0.2 = 1.15, fifteen percent. This is why the root-zone is held within roughly a single degree: at ±1 °C the kinetic span is 2^(±0.1) = 0.93 to 1.07, about seven percent peak to center, and seven percent is already the entire tolerance we are allowed for the finished fruit. Spend it on a wandering thermostat and nothing is left for every other source of drift.

The molecular stakes are concrete. Capsaicin is assembled at a single acyltransferase, Pun1/AT3, where vanillylamine condenses with a branched-chain fatty acid; the sequence edits we wrote are designed to improve that enzyme's catalytic efficiency and its fold stability, but even a well-designed enzyme still runs at the rate its temperature dictates. Hold the root-zone flat and the condensation proceeds at the same pace through every cycle, so the ratio of capsaicin to dihydrocapsaicin — the very *shape* of the burn, its onset and its tail — is laid down the same each time. Let the root-zone wander and you have a correctly designed enzyme producing a variable product, which is the most expensive kind of inconsistency: the recipe was right and the execution drifted.

The air's calm band

The vapor pressure deficit earns its place beside light and heat for a subtler reason: it governs whether the water and the dissolved nutrients we metered actually move through the plant. VPD is the gap between how much water vapor the air *could* hold at its temperature and how much it actually holds — the dryness felt as a pull on the leaf. It is computed from the saturation vapor pressure, which rises steeply and non-linearly with temperature; the Tetens approximation gives it in kilopascals as

```
SVP(T) = 0.6108 × exp( 17.27 × T ÷ (T + 237.3) ), VPD = SVP(T) × (1 − RH ÷ 100).
```

At 25 °C, SVP = 0.6108 × exp(1.646) = 3.17 kPa. Pair that with 70% relative humidity and VPD = 3.17 × 0.30 = **0.95 kPa**, squarely inside the calm band of roughly 0.8 to 1.2 kPa where the plant neither gasps nor drowns. Push the deficit too high and the plant shuts its stomata against the thirst, throttling the gas exchange that feeds photosynthesis — the light budget goes unspent. Let it fall too low and transpiration stalls, and with it the upward flow that carries nutrients from the root. There is only a narrow band where uptake runs steady, and holding the room inside it, hour by hour, is how the metered feed actually arrives where it was meant to go.

The non-linearity is exactly why VPD must be governed as temperature *and* humidity together, never one alone. Suppose we hold 68.4% RH to sit VPD at a clean 1.00 kPa at 25 °C (since 1 − 1.00 ÷ 3.17 = 0.316). Let the air cool just two degrees to 23 °C with the humidifier untouched: SVP falls to 2.81 kPa, and VPD slides to 2.81 ×

0.316 = **0.89 kPa** — out toward the drowsy low edge, from a temperature drift alone, with the humidity setpoint still reading perfect. A room that controls only what its humidity sensor sees is not controlling VPD. The calm band is a two-handed hold.

The feed

The nutrient solution is the smallest lever to describe and the easiest to take for granted. Metered as electrical conductivity — EC, a fast proxy for the total molarity of dissolved ions — it sets how much machinery the plant builds in the first place: nitrogen into the enzymes and the chlorophyll, phosphorus into the energy currency, potassium into the osmotic work that opens a stomate. Run the feed rich and the plant over-invests in vegetative hardware at the expense of the fruit; run it lean and the enzymes that should pour out flavor are simply never made in quantity. EC is pinned to a target molarity for the same reason as everything else: a flavor is only as reproducible as the least-controlled thing it depends on, and an unwatched feed is a slow drift hiding inside an otherwise perfect room.

Collapsing the variance

Pin them all, each to a tolerance of its own, cycle after cycle, and σ^2(environment) collapses toward the floor set by the plant's own biology. This is the whole table the room is built to hold:

VARIABLE	WHAT IT GOVERNS	REPRESENTATIVE PINNED SETPOINT	HELD TO	WHY DRIFT HERE MOVES THE SCORE
PPFD	the instantaneous carbon supply	~400 $\mu mol \cdot m^{-2} \cdot s^{-1}$	a few percent	sets the photons available this second to build every molecule
Photoperiod → DLI	the day's total flavor budget	~18 h → ~26 $mol \cdot m^{-2} \cdot d^{-1}$	±~1 $mol \cdot m^{-2} \cdot d^{-1}$	the master lever on total compound laid down
Root-zone temperature	enzyme rates ($Q_{10} \approx 2$)	held within ~1 °C	±1 °C	a single degree moves a rate ~7%; drifts the burn's shape
Nutrient EC	the machinery the plant builds	metered to target molarity	tight	too rich or too lean and the flavor enzymes are mis-provisioned
VPD	transpiration and uptake	calm band ~0.8–1.2 kPa	within band	outside it the metered feed never lands where intended

Now put a number on what that collapse buys. Independent variances add, so the coefficients of variation combine in quadrature:

> $CV(total) \approx \sqrt{[\ CV(genetic)^2 + CV(environment)^2 + CV(measurement)^2\ ]}$.

The reference values anchor the analysis — **42%** in wild type (from field observations) and **≤7%** as the design target for our own line — and the decomposition below is illustrative, meant only to show where each term comes from. In the field, nothing is held: the plants are a seed population (a large genetic term), the sky is a lottery (a large environmental term), and only the assay itself is tight. Those sources stack to the number that ruins formulation:

SETTING	GENETIC CV	ENVIRONMENTAL CV	BIOLOGICAL / MEASUREMENT CV	TOTAL CV = √(Σ OF SQUARES)
Field — wild type	~30%	~29%	~5%	**≈ 42%**
The room — a fixed, designed line	~3%	~4%	~5%	**≈ 7%**

$\sqrt{(30^2 + 29^2 + 5^2)} = \sqrt{1{,}766} \approx 42.0$, and $\sqrt{(3^2 + 4^2 + 5^2)} = \sqrt{50} \approx 7.1$. The first door — design — drives the genetic term from thirty toward three by making the plants one genotype. The second door — this room — drives the environmental term from twenty-nine toward four by pinning the table above. What remains is the biological floor, the σ^2(noise) no grower can delete, and the whole art is to push the controllable terms beneath it so the floor is all that is left. Seven percent is not a boast; it is $\sqrt{50}$, the arithmetic of two doors shut hard against a third that will not close.

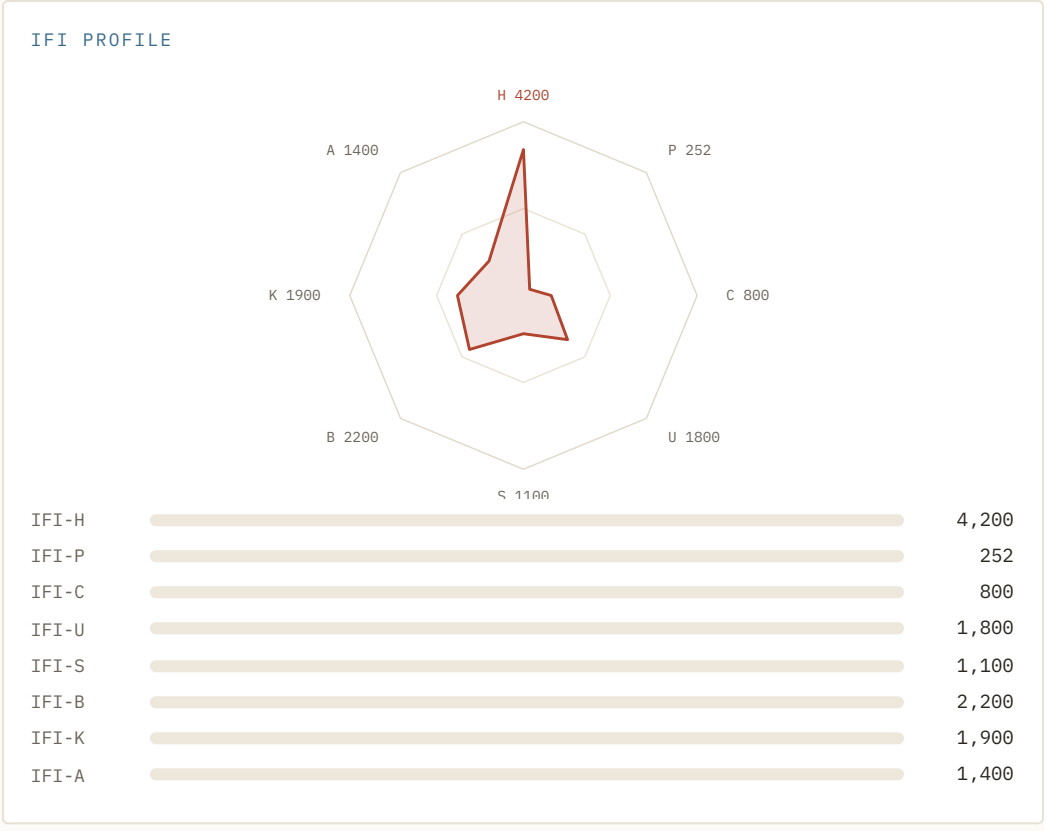


Phoenix (SS-0096) in the controlled chamber — a cultivar written to register on all eight axes at once, grown where every morning is the same morning so that all eight arrive at spec, lot after lot.

The room is what lets a ladder of designed heat behave like a set of specifications rather than a set of hopes. Each rung below is a distinct cultivar at a distinct predicted Scoville band, and the room's single job is to make each rung mean the same thing every lot — so that Blaze meets its design target of approximately 1.72 million SHU this winter and next, the same rung it was designed to occupy, every cycle:

CULTIVAR	SKU	SHU (PREDICTED)	IFI-H	AXES
Ember	SS-0087	866K	1,570	2/8
Cinder	SS-0088	1.05M	2,290	3/8
Kindle	SS-0089	1.25M	2,000	3/8
Flare	SS-0090	1.48M	—	4/8
Blaze	SS-0091	1.72M	2,630	4/8
Wildfire	SS-0092	1.96M	2,930	5/8
Scorch	SS-0093	2.21M	3,220	5/8
Firestorm	SS-0094	2.47M	3,490	6/8
Solara	SS-0095	2.72M	3,980	6/8
Phoenix	**SS-0096**	**2.95M**	**4,200**	**8/8**

At the top of that ladder sits the hardest thing the room has ever been asked to hold. Phoenix does not register on one axis but on all eight at once, and consistency on eight axes is not eight times harder than on one — it is harder than that, because the carbon budget, the enzyme temperatures, and the uptake band must land every axis simultaneously, in the same fruit, at the same dawn:



That the whole of that profile arrives at a batch CV of ≤7%, lot after lot, is the room's argument made in a single fruit. Eight numbers, each a receptor half-answered at a designed concentration, each delivered — not hoped for — because the morning that grew them was the same morning that grew the last.

What the room cannot do

Honesty is part of precision, so it is worth naming what a pinned room does not buy. It does not delete the $\sigma^2(G\times E)$ term — the interaction where a genotype expresses differently under a given environment; it fixes the environment to a single chosen point, which makes the interaction *constant* rather than *absent*, and a constant is only as good as the setpoint you chose. It does not survive sensor drift: a PPFD meter that has quietly lost its calibration, or a humidity probe reading three points high, hands the controller a lie and the controller holds the lie perfectly, so the room's discipline is only as sound as its metrology, and sensors are recalibrated on a schedule precisely because a confidently wrong instrument is worse than none. It does not control what it cannot measure — CO_2 enrichment, air movement across the canopy, and the position effect by which a plant at the edge of a bench lives a different light and VPD than one at the center are all real environmental terms, and a room that pins five variables while three breathe freely has only moved its variance, not removed it.

There is a floor it cannot cross. $\sigma^2(\text{noise})$ — the stochastic flutter of transcription, the developmental jitter from fruit to fruit on the same plant, the simple fact that two cells given identical instructions do not behave identically — is irreducible, and it is why the target is $\leq 7\%$ and not zero. The honest claim is not that the room makes a flavor perfectly repeatable; it is that the room drives every *nameable* source of drift beneath the unnameable one, so that what remains is biology's own small hum and nothing of the grower's making. And there is a limit of kind, not degree: the room holds a setpoint, but the plant feels an experience, and the two are not the same thing. A canopy that has grown taller since the sensors were placed lives in a different light than the one the controller thinks it is lighting. Precision is not a state you reach and leave; it is a thing you keep, by measuring the plant's world as the plant actually meets it, again and again.

The kept promise

And here the quiet through-line of this work surfaces, because it turns out that consistency and nourishment are the same act seen from two angles. A plant grown the same way, to the same fullness, every cycle, carries not only the same flavor every cycle but the same of everything else it was grown to hold — the same density of what makes a food worth eating. When the number is a standing property rather than a lucky harvest, the body gets what the number promised, lot after lot, with no good years and no lean ones to even out. We do not pin the room out of a love of control, and the reverence in that still green hour before the lights is real. We pin it because the deepest kindness a grower can offer is a fruit that is fully itself, reliably — so that the care put in at the root arrives undiminished at the table, every single time. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory. The room is not a cage. It is a kept promise: the same good morning offered to every plant, so that every plant can become, fully and without a wasted season, exactly what it was written to be.

Chapter 7 · The process — *the compounds that survive the journey.*

A flavor is not finished when the fruit is picked. Between the vine and the formulator's tank lies a gauntlet — extraction, drying, milling, standardization, blending — and every step of it is a chance for the number to move. Variance reaches a flavor through three doors: the gene, which decides what the plant can make; the field, which decides how evenly it makes it; and the process, which decides how much of that survives the trip to the tongue. This chapter is about the third door, and it is the one that belongs most fully to the maker rather than the grower. We can hand you an input held tight at harvest — a capsaicinoid load pinned to a batch-to-batch coefficient of variation of **≤7%**, against the **42%** a wild-type pepper scatters across a season — but what happens to it inside your own process is variance you author yourself. So this is the chapter where we stop talking about our work and start talking about yours, and tell you exactly how robust the thing is that you are about to put through it.

There is a reverence in that handoff that we do not want to lose under the arithmetic. A lot of fruit is the end of a long patience — a genome we read residue by residue, a season we held still on purpose — and when it leaves our hands as a standardized fraction it is still alive with everything the plant built into it. To carry a flavor through a process without spilling it is not a chore. It is the last act of care in a chain of care, and it rewards the maker who treats it as one.

The arithmetic of recovery

Begin with the plain physics, because the first place a beautifully grown lot quietly loses its number is here, in the recovery. Any extraction captures some fraction of the compound the fruit holds and leaves the rest behind. Call that fraction the yield, *Y*. If the fruit carried a mass *m* of capsaicinoids and your step recovers *Y* of it into a final volume *V*, the concentration leaving the extractor is

```
C(out) = m × Y ÷ V
```

and nothing in that sentence is fancier than mass balance. The trouble is that *Y* is not a constant. A little hotter solvent one day, a little longer contact the next, a filter that loads differently as it fouls — and the yield wanders. Because *C(out)* is linear in *Y*, a yield that moves by five percent hands you a concentration that moves by five percent, no matter how still the plant was held. An extraction running at **90%** recovery one morning and **85%** the next afternoon delivers two lots that differ by **(90 − 85) ÷ 90 = 5.6%** before the plant has had any say in the matter. The process has written its own scatter directly on top of ours.

The remedy is the discipline of this entire book, applied one more time: *pin the step, measure the recovery against a spiked control, and standardize the output.* A spiked control — a known mass of pure capsaicinoid carried through the same step — tells you *Y* for that run, so you are never guessing at it. And standardization means

that after the measurement you dilute or concentrate every batch to a declared target, so that what leaves the process is defined by its number rather than by the mood of the solvent that day. A standardized oleoresin is not a *purer* thing than the fruit; it is a *more defined* thing, and to a formulator a defined thing is worth more than a pure one.



A lot drawn down into standardized fractions — recovery held flat, measured against a spiked control, and cut to a declared target. The process door is the one the maker holds the key to, and holding this number is most of holding the number.

Now watch how the scatter from each step adds up, because the way it combines is the single most useful piece of mathematics a formulator can carry. The errors from independent steps do not stack arm in arm; they add in quadrature. If your material passes through the plant and then a chain of process steps, each contributing its own coefficient of variation, the output CV is the square root of the sum of the squares:

```
CV(out) = √( CV(plant)² + CV(extraction)² + CV(drying)² + CV(milling)² + CV(blending)² )
```

This is not an accounting convenience; it is what independent relative errors genuinely do. And it carries a lesson that reshapes where you spend your effort. Take our input at its specified floor, **CV(plant) = 7%**, and fold in an illustrative process: extraction at 5%, drying at 2%, milling at 3%, blending at 4%. The combined process term alone is $\sqrt{(25 + 4 + 9 + 16)} = \sqrt{54} =$ **7.35%**, and the whole chain lands at $\sqrt{(49 + 54)} = \sqrt{103} =$ **10.15%**. A seven-percent input left as a ten-percent output — the process nearly matched the plant's own scatter and never touched a gene.

Because the squares dominate, the largest term rules the result, and small terms barely register beside a big one. Collapse that same process to a single combined figure, *CV(proc)*, and read the output straight off:

COMBINED PROCESS CV	OUTPUT CV WITH A 7% INPUT	WHAT IT MEANS
0% (a perfect process)	7.00%	the plant's floor, inherited intact
3%	√(49+9) = **7.62%**	a tight process barely moves the number
5%	√(49+25) = **8.60%**	still well inside a formulator's tolerance
7%	√(49+49) = **9.90%**	the process now rivals the plant
10%	√(49+100) = **12.21%**	one sloppy chain swamps everything upstream

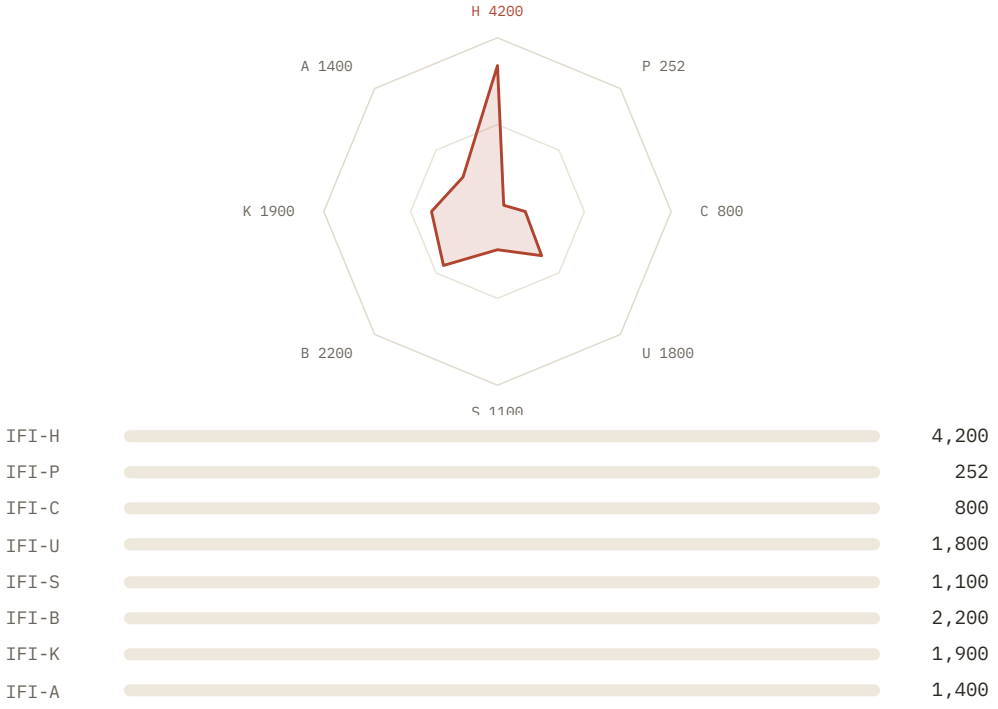
The table says three things at once. First, a process held under 3% is nearly free — it costs the final number only six-tenths of a point. Second, effort belongs on the *biggest* source: one careless step at 10% while the rest sit at 2% gives a combined process term of √(100+4+4+4) = √112 = 10.58%, and halving that one step helps more than perfecting all the others. Third — and this is the quiet payoff of everything we do at the gene — the plant's **7%** is a *floor the whole chain inherits*. A wild-type input at 42% stays at 42% even through a flawless process, because √(42² + 0²) is still 42; our 7% input through a genuinely sloppy 10% process still lands at 12.2%, better than wild-type's best possible case by more than three to one. You cannot process your way under the floor your input sets. We spent the plant so that your floor would be low.

What the index reads, and why a tenth of a percent moves it

A score on the Iyengar Flavor Index is not a measure of how much of a compound a food holds. It is a measure of how *loudly* that compound speaks to its receptor. Each axis is a stated convention: the reference compound defines IFI = 1,000 on that axis, in the same spirit that the Scoville scale is defined by dilution-to-threshold. Each axis has its own defined reference point, so scores across different axes are not directly comparable on a single common scale. Within any one axis, the score is read against that axis's fixed reference:

> $\mathrm{IFI(candidate)} = 1000 \times \mathrm{EC_{50}(reference)} \div \mathrm{EC_{50}(candidate)}$

Invert the candidate's EC_{50} against the reference's, scale by a thousand, and loudness becomes a number. A smaller EC_{50} means a more potent stimulus — half-activation reached at a lower dose — so a louder flavor. Work it on our full-spectrum reference cultivar, **Phoenix (SS-0096)**, whose heat axis reads **IFI-H 4,200**: that score means 4,200 = 1000 × EC_{50}(ref) ÷ EC_{50}(Phoenix), so EC_{50}(ref) ÷ EC_{50}(Phoenix) = **4.2**. Phoenix reaches half-activation of TRPV1 at **1 ÷ 4.2 = 23.8%** of the reference concentration. At the gentle end of our superhot ladder, **Ember (SS-0087)** reads IFI-H 1,570 — half-activation at 1 ÷ 1.57 = **63.7%** of the reference. Across the ladder the EC_{50} falls from roughly two-thirds of the reference to under a quarter of it; that falling EC_{50} *is* the rising score.



Why does a process engineer need the receptor mathematics? Because it tells you exactly how a wobble in concentration becomes a wobble in flavor — and it is not one-to-one everywhere. A receptor does not respond to a stimulus in a straight line; it responds along a saturating curve. For a channel like TRPV1, the fraction of channels open follows a Hill relation,

$$\theta([A]) = [A]^n \div (EC_{50}{}^n + [A]^n)$$

where n is the cooperativity — near 2 for capsaicin at TRPV1. At the half-activation point the curve is at its steepest in log-concentration, with a local slope of $n \div 4$. With $n \approx 2$ that slope is about 0.5, which means that right at the midpoint a **10%** swing in delivered capsaicin produces only about a **5%** swing in receptor response. The receptor compresses your variance. Climb onto the upper, flattening shoulder of the curve — where our hottest cultivars live — and the compression grows stronger still: the same fractional change in concentration moves the perceived loudness less and less.

You can read that compression straight out of our own canonical numbers, and it is one of our favorite facts in the whole library. Convert each cultivar's predicted Scoville heat level to a standardized capsaicinoid fraction through the pure-capsaicin anchor — **16,000,000 SHU** for the neat compound, so that a fraction f of capsaicinoids reads as $f \times$ 16,000,000 SHU, equivalently **1 ppm ≈ 16 SHU** — and set it beside our predicted IFI-H:

CULTIVAR · SKU	PREDICTED SHU	CAPSAICINOID FRACTION (STANDARDIZED OLEORESIN)	IFI-H	IFI-H PER MILLION SHU
Ember · SS-0087	866,000	5.41%	1,570	1,813
Cinder · SS-0088	1,050,000	6.56%	2,290	2,181
Blaze · SS-0091	1,720,000	10.75%	2,630	1,529
Wildfire · SS-0092	1,960,000	12.25%	2,930	1,495
Firestorm · SS-0094	2,470,000	15.44%	3,490	1,413
Solara · SS-0095	2,720,000	17.00%	3,980	1,463
Phoenix · SS-0096	2,937,948	18.36%	4,200	1,430

All SHU and IFI-H values are model predictions based on designed capsaicinoid profiles.

Read the last column top to bottom. At the gentle end, a million Scoville units buys you roughly **1,800 to 2,180** points of loudness; at the hot end that same million buys only about **1,410 to 1,465**. Loudness per unit of concentration falls by nearly a third as you climb — the unmistakable fingerprint of a saturating receptor. (The scatter is honest and real: Cinder sits high, Kindle's narrower axis set reads lower, and we do not sand those bumps flat.) The practical consequence for your process is a gift: the hotter your input, the *less* a given percentage of concentration scatter shows up as flavor scatter. The receptor forgives you most precisely where you are working hardest. It is the nearest thing to a free margin in the whole discipline, and it exists because of chemistry, not luck.

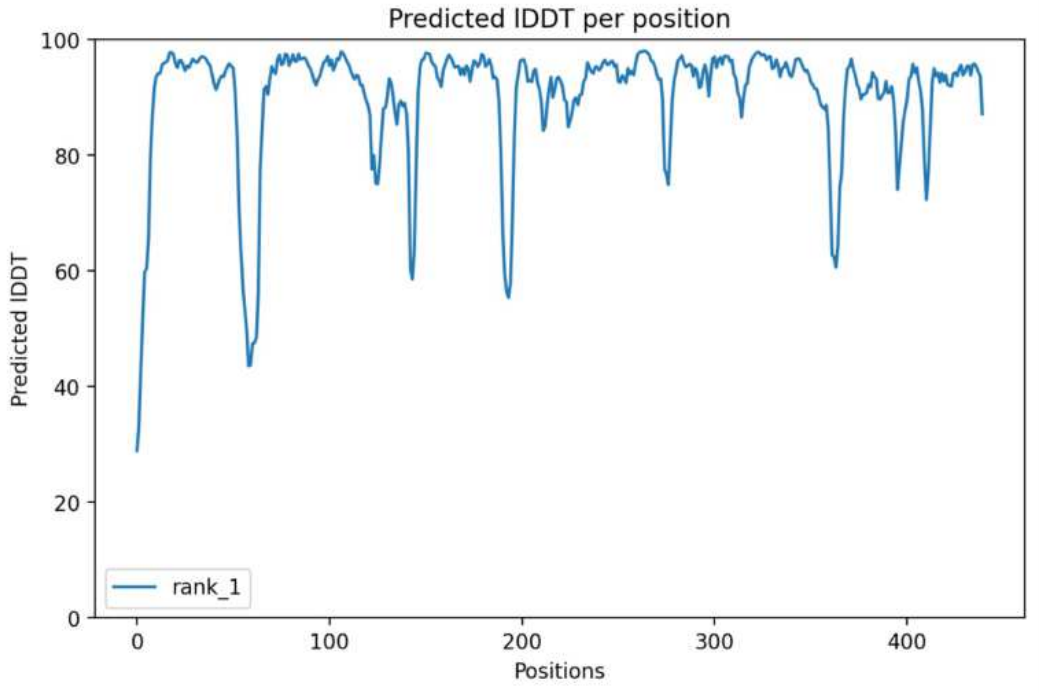
That forgiveness bounds your worst case from above. Because the delivered score can grow no faster than the delivered concentration — and on the saturating shoulder it grows slower — the concentration CV you control through standardization is a *ceiling* on the flavor CV your customer tastes. Hold the concentration to **±7%** and the heat a diner perceives varies by at most seven percent, and usually by less. Control the chemistry you can measure, and the biology quietly works in your favor.

Which molecules survive the journey

The second fact of the process door is that not all molecules travel equally well, and knowing which is which is most of the craft. To know it, you have to know what the molecule *is* — and this is where the chemistry we built pays the maker back.

Capsaicin is a stubborn, hardy molecule, and its ruggedness is written into its structure. It is an amide: a vanillylamine head joined to an 8-methyl branched-chain fatty acid tail by a single, thermally robust amide bond. That bond is formed in the fruit by the terminal acyltransferase **Pun1/CS**, the final capsaicin synthase in the pathway — the enzyme that labels the product. The head group is built along a pathway that runs in every vascular plant — the phenylpropanoid trunk, **PAL → C4H → 4CL → HCT → COMT**, which walks phenylalanine out to ferulate and on toward the vanillyl moiety —

and is handed to the aminotransferase **pAMT** to become vanillylamine. Our flux model places the principal control of capsaicinoid output at the upstream phenylpropanoid steps, primarily **4CL** and **HCT/CCoAOMT**; Pun1/CS operates at the terminal position and labels the product without governing the rate. The branched tail is assembled by a ketoacyl synthase, **KAS**, from a starter unit delivered by a branched-chain aminotransferase, **BCAT**. Only those four terminal enzymes — **pAMT · Pun1/CS · KAS · BCAT** — are specific to pepper heat; everything upstream is the common inheritance of plants. We know this chemistry because we engineered it, and knowing it means knowing exactly how the finished molecule behaves when the heat and the shear and the time of a real process come for it.



The Pun1/AT3 acyltransferase model, coloured by per-residue confidence — the enzyme that forges the amide bond at the heart of capsaicin. The bond it makes is the reason heat is the most forgiving axis to carry through a process.

What the structure predicts, the process confirms. Capsaicin carries no labile glycosidic linkage to hydrolyze, no low-boiling fragment to flash off with the steam; it is lipophilic, non-volatile, and chemically patient. It shrugs off a hard sear, holds its IFI-H through drying and milling almost indifferently, and that is precisely why heat is the most forgiving axis in the entire library to carry through a process. It is also why the two substitutions we wrote into Pun1 — **S39L** at the N-terminus for binding affinity and **L345G** for steric relief in the active-site cleft — matter at the fruit and need no defending downstream. They improve the enzyme's precision at the final synthesis step;

overall pathway output is controlled further upstream in the phenylpropanoid trunk, primarily at 4CL and HCT. The molecule then protects the rest of the journey itself.

The delicate end of the spectrum is another matter, and here honesty is the whole of the service. Volatile aromatics — the small terpenes and short esters that carry a top-note — have low boiling points and leave with the first steam of drying or cooking. Certain glycosides hydrolyze under heat or acid, their sugar cleaved away and their character changed. Peptides that carry richness are cleaved by heat and by stray proteases. The compounds behind the quietest, most beautiful axes are, as a rule, exactly the ones a careless process strips away. We do not hide this. We characterize the process-sensitivity of every compound we ship and tell a formulator plainly which inputs are rugged and which must be treated gently, added late, kept cool.

The eight axes through heat, shear, and time

Gather that characterization into one ledger, axis by axis, because a formulator building across the index needs to know not just *how much* of each to add but *when and how*. The receptors and carrier compounds are fixed by the index; the handling is what the process door asks of you.

AXIS · RECEPTOR	CARRIER COMPOUND(S)	WHAT THE PROCESS THREATENS	HOW IT IS HELD
IFI-H Heat · TRPV1	capsaicin, dihydrocapsaicin	almost nothing — amide-bonded, non-volatile, lipophilic	ship as standardized oleoresin; rugged through sear, dry, mill
IFI-P Pungency · TRPA1	allyl isothiocyanate, allicin, cinnamaldehyde	enzyme-released and volatile — allicin decays, isothiocyanate flashes off with steam	let the native enzymatic step fire late; add cool; keep sealed
IFI-C Cool-Burn · TRPM8	menthol, eucalyptol, camphor	volatile monoterpenoids — they leave with the steam	low-temperature handling; encapsulate; add last
IFI-A Astringency · salivary PRPs	catechins, tannins, proanthocyanidins	oxidize and polymerize with air, heat, and metal ions	inert atmosphere; control pH; minimise time at heat
IFI-U Umami · mGluR1 / TAS1R1·R3	glutamate, 5′-nucleotides	water-soluble and heat-stable — lost to leaching, not to destruction	control the water phase; protect the nucleotide synergy
IFI-S Sweetness · TAS1R2·R3	steviol glycosides, mogrosides, brazzein	glycosides can hydrolyse at low pH and heat; sweet proteins denature	hold pH; add protein sweeteners after the thermal step
IFI-B Zero-Bitter · TAS2R38	the attenuating modifier itself	the effect is a *ratio* to the bitter load; dilution shifts the balance	standardize the modifier-to-substrate ratio, not just the dose

AXIS · RECEPTOR	CARRIER COMPOUND(S)	WHAT THE PROCESS THREATENS	HOW IT IS HELD
IFI-K Kokumi · CaSR	γ-glutamyl peptides (e.g. γ-Glu-Val-Gly)	peptide bonds cleaved by heat and proteases	gentle thermal load; add post-process; watch enzymatic carriers

The ledger repays a careful read. Heat sits at the top because it asks almost nothing of you. Pungency and cool-burn sit near it in the list but at the opposite end of ruggedness — their molecules are volatile or enzyme-dependent, born late in the plant and quick to leave, so they want to be generated or added last, cool, and sealed. Astringency is a chemistry of oxidation: tannins and catechins polymerize in air and at heat, so the astringent axis is held by *excluding* oxygen and time more than by dosing. Umami is the quiet success story — glutamate is water-soluble and heat-stable, so you lose it only by pouring off the water it lives in, never by cooking it away. And two axes hide a subtlety that a naïve dosing strategy misses entirely: zero-bitter is a *ratio*, not an amount — the modifier attenuates bitterness relative to the bitter load present, so if a dilution step changes that ratio your attenuation drifts even when the modifier concentration is dead steady. You standardize the ratio, or you standardize nothing.

The blend is its own variance

There is one more trap, and it waits at the very end, after every compound has survived the heat. When a potent extract is cut into a carrier — a salt, a maltodextrin, an oil — the uniformity of the mix becomes its own variance term, with nothing to do with the plant and everything to do with the blade. A mixer that leaves hot and cool pockets in the carrier delivers a sample CV written entirely by the geometry of the vessel. Worse, dry carriers *segregate*: particles of different size and density drift apart under vibration and flow, so a blend that was uniform leaving the mixer can un-mix itself in the tote on the way to the line.

We hold it the way the pharmacopeias hold blend uniformity, because the pharmacopeias solved this problem first and solved it honestly. Sample the blend at many points across the vessel — not one grab from the top, but a spread drawn from the corners and the core — assay each, and compute the relative standard deviation of the set. Release only when that in-blend CV sits well under your line. The number of sampling points is itself a design choice with teeth: because the uncertainty in an estimated CV shrinks roughly with the square root of the sample count, doubling your confidence costs you four times the samples. Ten well-placed thieves tell you far more than one, and they tell you *before* release rather than after a customer complaint. The arithmetic is the same arithmetic as the recovery step and the quadrature sum: measure the spread, prove it is small, and let no batch leave on faith.



The handoff — a finished, standardized lot passing from our hands to the maker's. Everything before this moment is ours to hold; everything after it is yours, and the ledger tells you exactly how robust the thing in the vial is.

Failure modes, and the honest edge of what we can hold

A chapter on process that only lists victories is not telling the truth, so here are the ways it goes wrong and the places our control genuinely ends.

The first failure mode is thermal, and it cuts both ways. Heat does not only destroy flavor; through Maillard and caramelization it *creates* it, generating new aromatic compounds that were never in the fruit. This is a gift in a roast and a contaminant in a standardized input, because a compound the process invents is a compound the index did not anchor. The discipline is to decide which you want: a defined input carried gently and flavored deliberately downstream, or a transformed one whose new character you then measure and standardize in its own right. What you cannot do is leave it to chance and still claim a number.

The second is oxidation and the slow clock of shelf life. Capsaicin is patient, but the astringent and aromatic compounds beside it are not; air, light, trace metals, and water activity run a reaction you did not schedule. A lot measured at release and a lot measured at month six are not the same lot unless you have controlled the atmosphere, the packaging, and the moisture it sits in. Standardization is a measurement at a moment; stability is the promise that the measurement still holds when the input is finally used, and that promise is kept with inert headspace, cool storage, and a water activity low enough to still the chemistry.

The third is the honest edge itself. We can characterize how a compound behaves through heat, shear, and time, and we can hand you the sensitivity of every molecule we ship. We cannot follow the molecule into your particular kettle, your particular line speed, the particular hard water in your particular plant. The handoff in the figure above is a real boundary. Before it, the variance is ours, and we have spent a genome and a season to make it small. After it, the variance is yours — and the whole purpose of this chapter is to make sure you meet it with the recipe in hand: which input shrugs off the heat, which must be kept cool and added last, which is a ratio and not a dose, and which wants no air at all.

The door the maker holds

Process variation is the one door the grower cannot close for you; it is the maker's to hold. But it is not held alone, and it is not held in the dark. The care we built in at the gene — the steadied enzyme, the pinned recovery, the input designed to achieve **≤7%** where the wild plant scatters at **42%** — is a floor laid under your whole process, and the quadrature sum is the proof that a disciplined process barely lifts off that floor. The receptor mathematics is a second ally, compressing your worst-case concentration scatter into a smaller flavor scatter precisely where your inputs run hottest. And the ledger of sensitivities is the map, so that nothing fragile is lost to a step that could have been run cool.

We think about this the way we think about everything we grow: that a flavor is a kind of nourishment, something you give a body with intention, and that spilling it carelessly between the tank and the table is a small failure of care we would rather help you avoid. Hold the recovery flat. Standardize to the number, not the mood of the day. Add the delicate things late and cool. Prove the blend before you release it. Do those four things, and the long patience that built the plant arrives intact on the plate — not as our achievement handed down, but as a shared one, authored at the gene and finished, exactly, by your own hands.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

Chapter 8 · The assay, and the ruler's own wobble — *gauge R&R.*

You cannot hold a number you cannot measure twice. Everything in this book rests on a single quiet promise — that when we say a lot of fruit scores **IFI-H 4,200**, the fruit truly scores 4,200, and would score it again tomorrow, in another analyst's hands, on another plate. That promise is only as good as the instrument that keeps it. So before we trust any measurement of the plant, we turn the instrument on itself and measure the measuring.

There are four doors through which variation enters a number. The first is the plant's own genetics — the sequence we wrote and the sequence that grew. The second is the process — the light, the water, the temperature of a controlled chamber, the hour of harvest. The third is sampling — which fruit, which pericarp, which gram of a milligram-scale compound. The fourth door is the one almost everyone leaves unguarded: analytical variation, the scatter in the act of measuring itself. It is the most dangerous of the four precisely because it hides. A drifting assay reports chaos in a plant that was perfectly calm; a wobbling ruler makes every object look the wrong length, and a lab that does not watch for it will spend a fortune and a year re-engineering a cultivar to fix a fault that lives in its own pipette.



The assay stage — before we trust what it says about the fruit, we make it tell us the truth about itself.

The discipline that watches the fourth door has a name borrowed from manufacturing — measurement-system analysis — and its central instrument is the **gauge R&R** study, which splits the ruler's own error into two honest halves and then holds each one small. This chapter is the whole of that work: what the assay physically does to a molecule, how a dose becomes an EC_{50} and an EC_{50} becomes an index score, where the scatter in that chain comes from, and how we subtract the ruler's tremble back out so that the number we report is the plant and nothing but the plant.

Repeatability and reproducibility — the ruler's two tremors

The gauge R&R study answers a question with two parts, and the names of the two parts are the whole of its vocabulary.

Repeatability is the scatter you get when the *same* analyst measures the *same* sample on the *same* instrument, back to back, changing nothing — the irreducible tremble of the method itself. In the manufacturing literature it is called equipment variation, and we will write its standard deviation σ(EV). It is the hand that cannot hold perfectly still even when nothing else moves.

Reproducibility is the extra scatter that appears when you change something that *should not matter*: a different analyst, a different day, a different plate, a fresh aliquot of reagent from the same qualified lot. None of those things is the fruit; all of them move the number a little. That is appraiser variation, σ(AV), and in a well-run lab it should be the smaller of the two, because the parts that differ between analysts and days are exactly the parts we work hardest to standardize.

Together the two make the gauge — the ruler's own contribution to everything we see:

> **σ(gauge) = √(σ²(EV) + σ²(AV))** — repeatability and reproducibility, added in their squares, because independent scatters combine as the root of the sum of their variances, never as a plain sum.

That quadrature rule is not a convenience; it is the geometry of independent error, and it governs the entire chapter. Variances add; standard deviations do not. Two tremors of 1.6 and 1.0 do not make a tremor of 2.6 — they make $\sqrt{(1.6^2 + 1.0^2)} \approx 1.9$, noticeably less, because they are as likely to cancel as to compound. Hold that fact and the rest of the mathematics follows from it.

What the assay actually does to a molecule

Before we can measure the measuring, we have to be exact about what the measuring *is* — because the shape of the error depends entirely on the shape of the method. Every axis of the Index is anchored to a receptor, and the assay for each axis is, at bottom, a way of asking a living receptor how loudly a molecule speaks to it.

Five of the eight axes read ionotropic and metabotropic receptors through the same reporter: a rise in intracellular calcium. For heat, capsaicin binds the vanilloid pocket of **TRPV1**, a non-selective cation channel on the trigeminal endings of the tongue. The molecule does not warm the tissue; it lowers the channel's thermal threshold so the body reports a fire that is not there, and the channel opens and lets calcium pour in. TRPV1 is a homotetramer with four vanilloid sites, and the four fill cooperatively — which is why its dose-response curve climbs with a Hill coefficient near 2 rather than 1, a steeper, more switch-like rise than a single-site receptor would give. Cool-burn reads **TRPM8** and pungency reads **TRPA1** by the same direct-calcium route; TRPA1 is the strange one, gated not by a snug fit but by covalent chemistry — electrophilic agonists such as allyl

isothiocyanate and allicin bond to specific cysteine residues on the channel's inner face, so its "binding" is a chemical reaction and its kinetics are, honestly, harder to fit cleanly.

The savory and sweet axes read G-protein-coupled receptors, which do not pass calcium themselves. **TAS2R38** (bitterness), **mGluR1** with the **TAS1R1·TAS1R3** heterodimer (umami), **TAS1R2·TAS1R3** (sweetness), and **CaSR** (kokumi) all couple through a G protein to phospholipase C, to IP_3, to a release of calcium from internal stores — and we watch that release with the same fluorescent calcium reporter, routed through a promiscuous G-protein chimera so a taste receptor's signal lands in a channel we can read. Kokumi is the subtle case: its agonists, glutathione and the tripeptide γ-Glu-Val-Gly, are not even true agonists but positive allosteric modulators, compounds that do not shout at CaSR so much as lean on it and make every other voice louder.

The eighth axis is the one that refuses the pattern, and it matters that we say so. Astringency is read not through a receptor channel at all but through the **salivary proline-rich proteins**, and the "signal" is a precipitation — tannins cross-link the proteins and the solution clouds. We measure it by turbidity, by light scattered off aggregates, which means it has its own units, its own standards, and its own, separately-run measurement-system analysis. A calcium assay's gauge study tells you nothing about a turbidity assay's wobble. Honest measurement science does not pretend one ruler grades every axis.

AXIS	RECEPTOR	CLASS	ASSAY READOUT	CHARACTERISTIC AGONIST
IFI-H · Heat	TRPV1	ionotropic cation channel	direct Ca^{2+} influx	capsaicin
IFI-C · Cool-Burn	TRPM8	ionotropic cation channel	direct Ca^{2+} influx	menthol
IFI-P · Pungency	TRPA1	electrophile-gated channel	direct Ca^{2+} influx	allyl isothiocyanate · allicin
IFI-S · Sweetness	TAS1R2 · TAS1R3	GPCR heterodimer	Ca^{2+} via Gα chimera	sucrose · brazzein
IFI-U · Umami	mGluR1 · TAS1R1·TAS1R3	GPCR	Ca^{2+} via Gα chimera	glutamate
IFI-B · Zero-Bitter	TAS2R38	GPCR (gustducin)	Ca^{2+} via Gα chimera	PROP · PTC
IFI-K · Kokumi	CaSR	GPCR	Ca^{2+} (allosteric)	glutathione · γ-Glu-Val-Gly
IFI-A · Astringency	salivary PRPs	protein precipitation	turbidity / nephelometry	tannic acid

From dose to EC_{50} — reading the curve

Whatever the receptor, the measurement is the same gesture: present the molecule at a ladder of concentrations, usually eight to twelve steps a half-log apart, and watch the response climb. The curve that results is the sigmoid every pharmacologist knows, the Hill equation:

```
R(c) = R_max · cⁿ / ( EC₅₀ⁿ + cⁿ )
```

where *c* is concentration, *n* is the Hill coefficient that sets the curve's steepness, and **EC_{50} is the concentration at which the response reaches half of its maximum** — the single number that says how loudly the compound speaks. Set $c = EC_{50}$ in that equation and the two c^n terms make the fraction exactly one-half, R_max/2, by construction. We fit all four parameters — bottom, top, EC_{50}, and *n* — to the measured points by nonlinear least squares, and the fit returns not just the EC_{50} but a standard error on it, the first and most important honest number in the whole chain.



Assay plates on the bench — each well a concentration, each column a dose-response, the whole tray one reading of one lot.

A low EC_{50} means a loud molecule: it reaches half-activation at a low concentration, so a little goes far. That is the quantity the Index inverts into a score.

The index is a ratio — and the ratio is why the ruler is quiet

The Iyengar Flavor Index does not report an EC_{50} directly. It reports a loudness relative to a fixed reference:

```
IFI(candidate) = 1,000 × EC₅₀(reference) ÷ EC₅₀(candidate)
```

The reference — native wild-type — is anchored at 1,000 by convention, so a score of 4,200 means the candidate reaches half-activation at EC_{50}(candidate) = EC_{50}(reference) × 1,000 / 4,200 = EC_{50}(reference) / 4.2. The fruit is 4.2 times more potent than the reference; it speaks four times louder from a quarter of the dose. Our predicted IFI-H design scores across the flagship heat ladder invert to EC_{50} fractions by the same arithmetic, every value a clean function of a canonical score and nothing invented:

CULTIVAR	SKU	PREDICTED IFI-H	RELATIVE POTENCY	EC_{50} (FRACTION OF REFERENCE)
Ember	SS-0087	1,570	1.57×	0.637
Blaze	SS-0091	2,630	2.63×	0.380
Firestorm	SS-0094	3,490	3.49×	0.287
Solara	SS-0095	3,980	3.98×	0.251
Phoenix	SS-0096	4,200	4.20×	0.238

Every EC_{50} fraction in that last column is simply 1,000 ÷ IFI-H: a candidate that scores 4,200 reaches half-activation at 1,000/4,200 = 0.238 of the reference concentration. The arithmetic is deliberately that plain, because a score that cannot be inverted back to the physical quantity it came from is not a measurement, it is a decoration.

And here is the structural gift the ratio hands us, the reason the gauge in this chapter is so small. Because the index is a *ratio* of two EC_{50} values, and because we measure the reference on every single plate alongside the candidate, any error that multiplies the whole plate — a reporter dye a touch hot, a cell passage a day old, a plate reader running a hair warm — multiplies the numerator and the denominator alike and *cancels*. What survives is only the independent uncertainty in the two curve fits themselves. Propagate those through the ratio the way variances always propagate through a quotient:

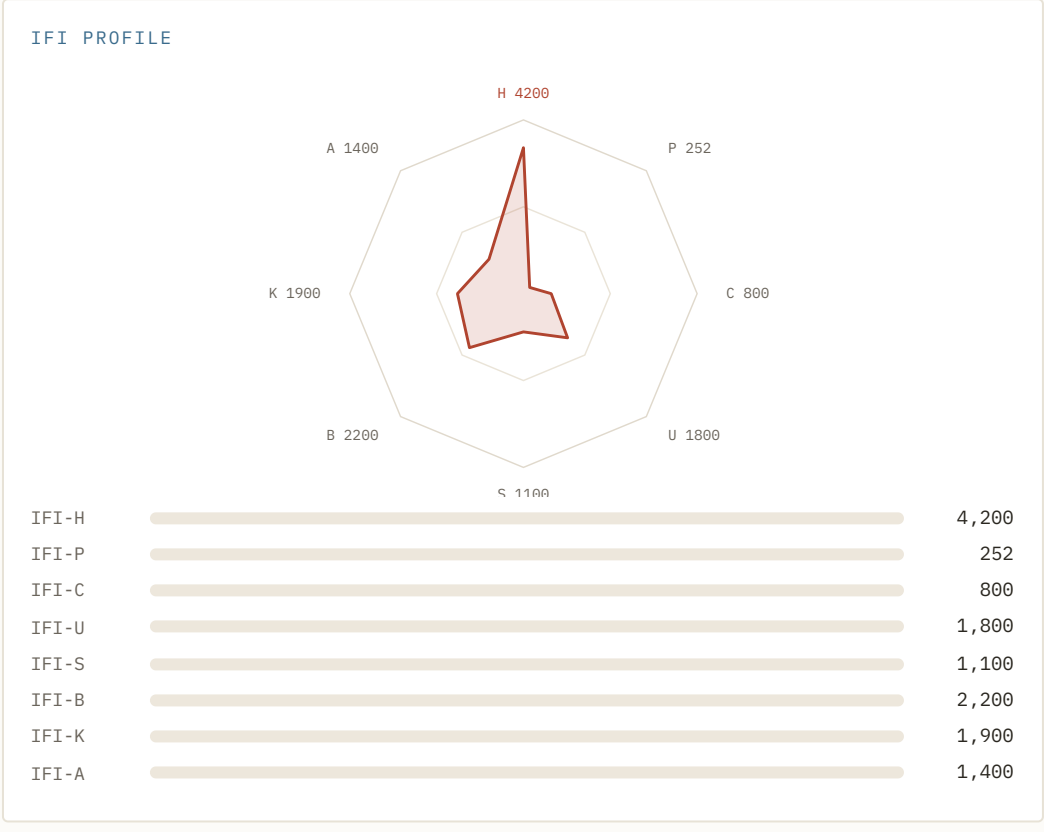
> **$CV(IFI) = \sqrt{\,CV^2(EC_{50}\ \text{reference}) + CV^2(EC_{50}\ \text{candidate})\,}$**

The reference EC_{50} is measured on every plate and averaged across many, so its scatter is beaten down to roughly 0.6 percent. The candidate is a single on-plate fit, carrying about 1.8 percent. Combine them in quadrature and the index score should carry

> $\sqrt{\,0.6^2 + 1.8^2\,} = \sqrt{\,0.36 + 3.24\,} = \sqrt{3.60} \approx$ **1.9 percent**.

Hold that number. It is a *prediction*, built from the bottom up out of the physics of two curve fits, and in a moment the empirical gauge R&R study will arrive at the same 1.9 percent from the top down. When a bottom-up error budget and a top-down variance study agree, it is not a coincidence to wave at — it is proof that no variance source has gone missing, that the model of the measurement is complete.

The cultivar those numbers describe most fully is the one engineered to register on every axis at once:



Eight receptors, eight independent EC_{50} fits, eight ratios — and eight separate gauge studies standing behind them, because the ruler for heat is not the ruler for kokumi and a single blanket claim about precision would be a lie of convenience.

The study that measures the measurement

The number that comes off the bottom-up propagation has to be confirmed by a real experiment, and the experiment has a deliberate shape. We take a set of known samples that span the range we intend to grade, and we have several analysts each measure each sample several times, across several days and plates — a crossed, nested layout, every analyst crossed with every sample, every measurement nested inside its day and plate. Then the analysis of variance partitions the total scatter into the sources that made it:

$$\sigma^2(\text{total}) = \sigma^2(\text{part}) + \sigma^2(\text{EV}) + \sigma^2(\text{AV})$$

the variance *between the samples* (the signal we actually want — real differences between real fruit), the variance of one analyst measuring one sample twice (repeatability), and the variance that enters across analysts and days (reproducibility). The gauge is the root of the sum of the two error terms, exactly as the receptor-level propagation said it should be: $\sigma(\text{gauge}) = \sqrt{\sigma^2(\text{EV}) + \sigma^2(\text{AV})}$.

Run that partition on a tightly-controlled lot and the budget resolves into four honest rows. We express every term as a coefficient of variation — a standard deviation as a percentage of the mean, CV = σ/μ — because CV is the scale-free language in which a heat score and a kokumi score can be compared on one page.

SOURCE OF VARIATION	SYMBOL	SD (AS CV)	VARIANCE (CV²)	SHARE OF TOTAL VARIANCE
Part — biology + process	σ(part)	4.30 %	18.49	83.7 %
Repeatability — equipment	σ(EV)	1.60 %	2.56	11.6 %
Reproducibility — analyst / day	σ(AV)	1.02 %	1.04	4.7 %
Gauge — EV ⊕ AV	σ(gauge)	**1.90 %**	3.60	16.3 %
Total observed	σ(total)	**4.70 %**	22.09	100 %

Read the rows as arithmetic, because they are. The three real sources add in their squares: 18.49 + 2.56 + 1.04 = 22.09, whose root is $\sqrt{22.09} = 4.70$ percent, the total scatter an unexamined lab would simply report and stop. The gauge is not a fourth addend but the quadrature of the two error terms, $\sqrt{(2.56 + 1.04)} = \sqrt{3.60} = 1.90$ percent — the very 1.9 percent the receptor-level propagation predicted. The two roads meet. And reproducibility, at 1.02 percent, is comfortably the smallest term in the table, which is the signature of a method whose handoffs between people and days have been engineered down to almost nothing. Drift between analysts is the error that most often masquerades as a bad harvest; here it is the quietest voice in the room.

Reading the ruler — %GR&R, distinct categories, and an honest subtraction

Two summary readouts come off that partition, and the discipline holds the ruler to both before a single lot is graded against it.

The first is **%GR&R** — the fraction of the scatter that is the ruler rather than the thing:

> **%GR&R (standard-deviation basis) = σ(gauge) / σ(total) = 1.90 / 4.70 = 40.4 % %GR&R (variance basis) = σ²(gauge) / σ²(total) = 3.60 / 22.09 = 16.3 %**

Those two percentages describe one situation and must be read together, because their difference is the subtlest and most honest point in the chapter. On the variance basis — the one that answers *how much of the real variation is the gauge* — the gauge takes 16.3 percent, roughly a sixth, and the rest is genuine fruit. That is well inside the line a careful lab draws. On the standard-deviation basis, the same gauge reads a startling 40.4 percent, a figure the old manufacturing convention would flag as unfit.

Both are true, and the reason they diverge is worth more than either number alone. The standard-deviation %GR&R is a ratio against the *total* scatter, and our total scatter is tiny — 4.7 percent — because our process is so tightly controlled. A gauge looks worse the better your process becomes, because there is less and less part variation to compare it against; the denominator shrinks and the fraction swells. A lab reading 40 percent here has not built a bad ruler. It has built a superb process and is measuring its ruler against almost nothing. The fix is to judge the gauge where judgment belongs — against the thing it must resolve.

That is what the second readout does. The **number of distinct categories** asks how many genuinely different levels the ruler can tell apart across a given spread:

> **ndc = 1.41 × σ(part) / σ(gauge)** — where 1.41 ≈ √2.

Evaluate it against the range we actually grade, and the ruler is sharp. The flagship heat ladder runs from Ember at IFI-H 1,570 to Phoenix at 4,200; the part-to-part standard deviation across that ladder is about 890 index units on a mean near 2,920 — a part CV of roughly 30.4 percent. Then

> ndc = 1.41 × 30.4 / 1.9 ≈ **22.6** — call it twenty-three distinct rungs the ruler resolves across the ladder,

far above the floor of five below which an instrument is too blunt to support a specification at all. But evaluate the *same ruler* against a single tightly-held lot, where the only part variation is the 4.3 percent of biology and process inside one cultivar, and the answer is honest and humbling:

> ndc = 1.41 × 4.3 / 1.9 ≈ **3.2** — barely three categories.

That is not a defect. It is the plain consequence of how little two lots of the same engineered cultivar differ from each other. The ruler can place any fruit precisely on the long ladder of cultivars; it cannot finely rank two nearly-identical lots of Phoenix against each other,

because there is almost nothing there to rank. We state that limit out loud rather than let a flattering number imply a precision we do not have.

And now the payoff the whole study exists to deliver. Knowing the gauge on its own, we can clear it out of the observed scatter — not by subtracting the percentages, which would be wrong, but by subtracting the *variances* and taking the root, the same quadrature run in reverse:

> **σ(biology + process) = √(σ²(total) − σ²(gauge)) = √(4.70² − 1.90²) = √(22.09 − 3.61) = √18.48 ≈ 4.30 percent.**

The true scatter of the fruit — the plant and the process, cleaned of the instrument — is 4.3 percent, not the 4.7 the raw data showed. The 0.4-point difference is small, and that is exactly the trap: a lab that skipped this step would report 4.7, go hunting for half a percent of improvement, and spend a season chasing a fault that was never in the greenhouse. It was in the pipette, and it was always 1.9 percent, and it was knowable from the first day anyone bothered to measure the measuring.

GAUGE R&R READOUT	FORMULA	OUR RESULT	LINE IT ANSWERS TO
%GR&R — variance basis	σ²(gauge) / σ²(total)	16.3 %	below ~20 % — a CV that is mostly the plant, not the lab
%GR&R — SD basis	σ(gauge) / σ(total)	40.4 %	high only because the process is so tight; read against tolerance, not total
ndc — across the grading ladder	1.41 × σ(part) / σ(gauge)	≈ 23	≥ 5 — a ruler sharp enough to grade the range
ndc — within a single lot	1.41 × σ(part) / σ(gauge)	≈ 3	the honest limit on ranking near-identical lots
Recovered biology + process CV	√(σ²total − σ²gauge)	4.30 %	the true scatter of the fruit, instrument removed

The quiet discipline that keeps the ruler honest

None of these numbers stay true on their own. A gauge held to 1.9 percent in a qualification study and then left alone will drift, and the discipline that keeps it where it was found is deliberately dull — which is the point. We hold the analytical term to the smallest of the four doors by the plain means that work, run every day, for every axis, without exception and without drama.



The controlled chamber — fixed light, fixed temperature, a reference measured on every plate; the conditions under which a number means the same thing twice.

A **positive and a negative control sit on every plate** — a known agonist at a known concentration and a vehicle blank — so that a plate whose controls have moved is caught before its data is ever believed, not a month later in a confused meeting. **Every sample is read in replicate**, three wells at minimum per concentration, because a single well is an anecdote and a dose-response curve is only as trustworthy as the scatter around each of its points. The **reference is carried on every plate**, which is what makes the index a ratio and lets the plate-wide errors cancel. And the **calibration is re-qualified on a fixed schedule** — a standing calendar, not a nervous impulse — because requalifying only when someone is worried biases you toward confirming the worry, while requalifying on a clock catches the slow drifts that no one was worried about yet.

The honest limits are stated plainly, because a measurement system you oversell is worse than one you distrust. The ruler resolves the long ladder of cultivars beautifully and ranks two lots of one cultivar only coarsely — and we design our specifications around what it can do, grading against the ≤ 7 percent batch CV we promise rather than pretending to split differences finer than three categories. The astringency axis rides an entirely separate turbidity method with its own gauge study, and we never let a calcium assay's clean 1.9 percent borrow credibility for a precipitation readout that must earn its own. TRPA1's covalent, reaction-gated kinetics fit less cleanly than TRPV1's cooperative binding, so its curves carry a wider fit error that we report rather than smooth over. And the whole subtraction that recovers the 4.3 percent depends on the gauge and the fruit being statistically independent — the moment an instrument's error begins to depend on the sample it is reading, the quadrature stops being clean and the study has to be rebuilt. We watch for that coupling as closely as we watch the drift.

For context on how much this matters: a wild-type pepper, grown and measured without any of this, scatters its heat at a batch CV near 42 percent — nearly half its own signal is noise, and much of that noise, uninvestigated, is the measurement. The engineered lots we grade hold to ≤ 7 percent, and the gauge inside that seven is 1.9 — less than a third of the whole promise, a ruler small enough that it cannot by itself push an in-spec lot out of spec, and still we subtract it in quadrature so the number we publish is the fruit alone. The

difference between 42 percent and 7 is the plant we rewrote. The difference between 4.7 and 4.3 is the ruler we refused to confuse with it.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

That is the whole of the fourth door. Measured on its own, held low on a fixed schedule, and subtracted back out in the only arithmetic that honors independent error, the ruler becomes the one thing in the building we are allowed to trust completely — which is exactly what a ruler is for.

Chapter 9 · The reference standard — *the fixed point the whole scale hangs on.*

Look one more time at the equation every book in this library is built on, and watch how much of its weight comes to rest on a single word.

> IFI(candidate) = 1,000 × EC_{50}(reference) ÷ EC_{50}(candidate)

The score of any candidate — any pepper, any tomato, any cup of tea, on any of the eight axes — is the half-maximal activation of a *fixed reference*, divided by the candidate's own, scaled by a thousand. Everything downstream is a ratio, and a ratio needs a denominator that never moves. The reference is that denominator. It is the one quantity in the whole system that is not measured *against* anything; everything else is measured against *it*. Which is the quiet terror of the thing: should the reference ever drift — degrade a few percent in a freezer over a winter, be re-made a shade off, be handed to a new lot nobody re-qualified against the old — then every number in every book drifts with it, silently, all at once, in exactly the same direction, and no tongue catches the error until a product fails a year downstream and the cause turns out to be a decade upstream.

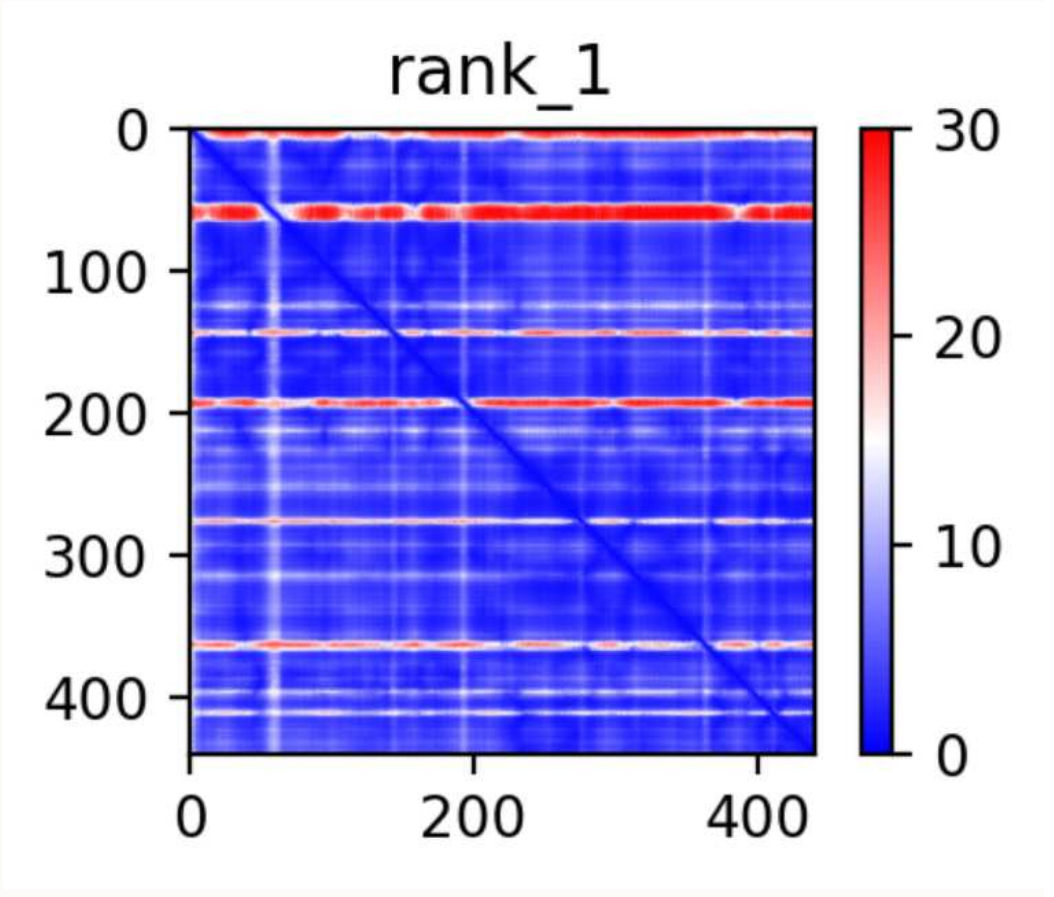


FIG. 9 — the predicted aligned error map of the Pun1/AT3 model; a picture, really, of knowing exactly how much you don't know, which is the first thing a reference standard is for.

So the reference is the most carefully kept thing we grow, and it is kept by the oldest discipline in all of measurement. Let us set the arithmetic in the open and show exactly what it anchors, exactly how it is held still, and exactly how far its honesty can reach — because a standard is worth only as much as the doubt it can state about itself.

Why we anchor at the half-maximal point

Capsaicin is the reference ligand of the heat axis, and TRPV1 is the lock it turns: a non-selective cation channel on the trigeminal endings of the tongue, the same channel gated open by temperatures above 43 °C and by a fall in pH. The molecule does not warm the tissue. It binds in the vanilloid pocket and lowers the channel's thermal threshold, so the nerve reports a fire that is not there and the body believes it. The number we anchor to is not *how much* capsaicin a fruit holds but *how loudly* a given amount of it speaks to that channel — its EC_{50}, the concentration that drives the receptor to half of its full opening.

Why half, and not the first flicker of a response, or the full roar of saturation? The answer is error, and it is worth deriving, because it is the reason the whole scale stands where it stands. A receptor does not answer a stimulus in a straight line; it answers in a sigmoid. Write the fractional activation θ against ligand concentration [L]:

$$\theta([L]) = [L]^n \div (EC_{50}{}^n + [L]^n)$$

Here n is the Hill coefficient — the cooperativity of gating, near 2 for the way capsaicin opens TRPV1, which makes the curve steeper than simple one-to-one binding would. At $[L] = EC_{50}$ the two terms in the denominator are equal and θ = ½: the channel is exactly half-open, by construction. The sensitivity of the readout — how much the response moves for a fractional change in concentration — is the slope of that curve in log-concentration:

$$S = d\theta \div d(\ln[L]) = n \cdot \theta \cdot (1 - \theta)$$

and θ(1 − θ) is largest precisely where θ = ½, at $[L] = EC_{50}$, with a peak value of $n/4$. Everywhere else the curve is flatter. Near the threshold the receptor is nearly silent; near saturation it is nearly full; and in both places a large change in concentration produces only a whisper of change in response, so reading a concentration *back* from a measured response is worst-conditioned exactly where the curve lies flat and best-conditioned at the half-max. Anchoring the entire scale at EC_{50} is therefore the single choice that lets the least measurement noise through. The biology and the metrology turn out to be one sentence: the steepest point of the receptor's answer is also the most honest place to stand, and for a cooperative channel with n near 2 the steepness — and so the honesty — is sharper still. Put a number on it. Read a response up at 95% of maximum, where θ = 0.95, and the sensitivity has fallen to $n \cdot 0.95 \cdot 0.05 \approx 0.048n$ — barely a fifth of its half-max value of $n/4 = 0.25n$ — so identical response noise now maps to roughly five times the concentration error. The flat shoulders of the sigmoid are not merely less

convenient than its middle; they are quantitatively blind, and a scale anchored out on them would inherit that blindness in every reading it ever produced.

A reference enters every score as a multiplier

Now make the terror precise, because precision is the only thing that disarms it. Hold a candidate fixed and let the reference's measured EC_{50} move by a small fraction δ — call the true value R and the drifted value R(1 + δ). The reported score becomes

> $IFI'(c) = 1{,}000 \times R(1 + \delta) \div EC_{50}(c) = (1 + \delta) \times IFI(c).$

Every score on that axis, multiplied by the same (1 + δ). This is the mathematical shape of the danger: the reference is not one error among many, averaging out in a crowd — it is a *common-mode* factor that rides every number at once. A reference that has quietly lost 3% of its potency, so that it now takes 3% more of it to reach half-activation and its EC_{50} reads δ = 0.03 high, lifts every heat score by 3% in the same stroke. A flagship's true IFI-H of 4,200 reads 4,326; a line candidate sitting exactly on a 2,630 spec reads 2,709; a gentle 1,570 reads 1,617. Hold them up against one another and nothing looks wrong at all — the rank order is untouched, every ratio *between* candidates is preserved, every internal sanity check passes — because the error is identical in all of them. It is invisible to every comparison the laboratory can make with itself, and fatal to every absolute promise the laboratory makes to the world: the formulator a continent away, working to a locked target, is now a silent 3% off on every incoming lot. An error that hides from all your mirrors and shows only to your customer is the exact failure a reference standard exists to forbid, and the only defense against a common-mode error is a reference kept so still that δ never leaves the noise. The sign runs both ways and the arithmetic is blind to how hot the fruit is: a reference re-made a touch too concentrated reads low and *deflates* every score by the same factor instead — a gentle 1,570 and a ferocious 4,200 shaved by the identical fraction. The error cannot see the candidate at all; it lives wholly in the denominator, which is the one place no candidate, and no internal comparison, can ever catch it.

Reading the ladder in units of the reference

Because the score is a ratio, it can be read backwards, and reading it backwards is how the abstraction becomes a bench quantity. Rearrange:

> $EC_{50}(candidate) = 1{,}000 \times EC_{50}(reference) \div IFI(candidate) = (1{,}000 \div IFI) \times R.$

A candidate's own half-activation concentration is pinned to the reference's as a pure multiple — no absolute micromolar figure required, which is the whole convenience of a reference-anchored scale. A score of 1,000 *is* the reference, EC_{50} = R by definition. A score of 4,200 means the candidate reaches half-activation at 1,000 ÷ 4,200 = 0.238 of the reference's concentration: it is 4.20× louder,

and that 4.20 is simply the score divided by a thousand. The heat ladder of the flagship superhots, read in these units, lays the whole relationship bare:

CULTIVAR · SKU	PREDICTED SHU	IFI-H	$EC_{50}(C) \div R$	PUNGENCY VS. PURE CAPSAICIN
Ember · SS-0087	866,000	1,570	0.637	5.4%
Blaze · SS-0091	1,720,000	2,630	0.380	10.8%
Firestorm · SS-0094	2,470,000	3,490	0.287	15.4%
Solara · SS-0095	2,720,000	3,980	0.251	17.0%
Phoenix · SS-0096	2,937,948	4,200	0.238	18.4%

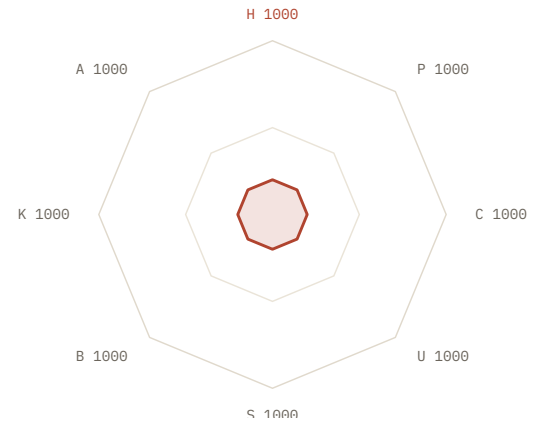
(SHU values are computational predictions from our design model, not certified measurements.)

The two rightmost columns tell the same heat in two older dialects. $EC_{50}(c) \div R$ is loudness measured in units of the reference — the fraction of the anchor's concentration this fruit needs to shout equally loud. The Scoville figure is its concentration cousin, not how loudly the molecule speaks but how much of it the fruit carries, and it too is meaningful only against a fixed point. Normalized so that pure capsaicin reads 16,000,000 SHU, the conversion is one part per million of capsaicinoids ≈ 16 SHU, and our model predicts 2,937,948 SHU for Phoenix, which works out to 2,937,948 ÷ 16,000,000 = 18.4% of the pungency of pure capsaicin — the rightmost column, each rung a step up a single, honest ruler.

It is worth remembering that Scoville's own scale spent its first century anchored to nothing steadier than a panel of human tongues diluting a sample until the burn vanished — a reference that aged, tired, acclimatized and disagreed, drifting with every taster and every long afternoon. High-performance liquid chromatography finally gave it a chemical anchor: quantify the capsaicinoids, convert, and the dilution panel's wandering fixed point was replaced by a molecule that does not get tired. The IFI gives *loudness* the same gift one layer deeper, at the receptor itself. The lesson simply repeats at every scale of the problem: a measurement is only ever as steady as the thing it is measured against.

On its own eight axes, by construction, the reference set reads a flat thousand straight across —

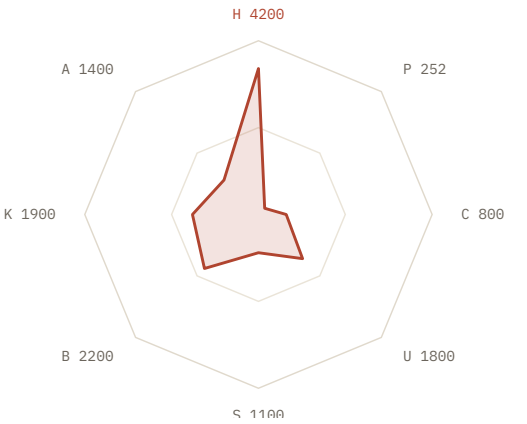
IFI PROFILE



IFI-H	1,000
IFI-P	1,000
IFI-C	1,000
IFI-U	1,000
IFI-S	1,000
IFI-B	1,000
IFI-K	1,000
IFI-A	1,000

— not because any one anchor is bland, but because each axis is measured against its own reference, so each reference scores unity, times a thousand, against itself. This is the fixed point drawn as a profile: eight flat thousands, the still center the whole library turns on. Against it stands the hardest possible audit of whether every anchor is holding — Phoenix, the one cultivar designed to register on all eight axes at once:

IFI PROFILE



IFI-H	4,200
IFI-P	252
IFI-C	800
IFI-U	1,800
IFI-S	1,100
IFI-B	2,200
IFI-K	1,900
IFI-A	1,400

Eight axes, eight independent references, eight denominators that must each stay still for this single profile to mean the same thing next year as it does today. A full-spectrum fruit is where reference discipline is tested most, because it is only ever as honest as the shakiest of its eight anchors.

The traceability chain

Keeping a fixed point still is not a feeling; it is a hierarchy, and every number in the library can be walked back up it to one defined origin.

TIER	WHAT IT IS	HOW IT IS HELD
Primary standard	the master characterization of each axis's reference	defined once, exhaustively, against first principles
Working standards	the aliquots that actually touch the daily assays	drawn from the primary; each re-qualified against it before use
Daily controls	the check samples on every plate	read every run; a drift here halts the line

The quiet heart of the chain is the handoff between working lots, and it is governed by statistics, not by nerve. A new working lot is never simply swapped in. It is run *alongside* the outgoing lot across a block of shared assays — an overlap — and released to anchor measurements only once the two agree inside the budget. The size of that block is set by how small a disagreement we insist on being able to see. A single assay carries a coefficient of variation we hold near 2%; the uncertainty in the *mean* of n replicates shrinks as

$$SE = CV \div \sqrt{n},$$

so to resolve a bias between lots as small as the budget itself — to catch a 2% step before it can hide — we want the standard error of the comparison down near a quarter of that, about 0.5%. Then

$$n \geq (CV \div SE)^2 = (2 \div 0.5)^2 = 16,$$

sixteen paired reads per lot as a floor, and in practice we run more. The release rule is deliberately an equivalence test, not a test of difference: the new lot passes only when the confidence interval on (new − old) lies *entirely inside* ±budget. The burden is on proving sameness, never on merely failing to prove a difference — a distinction that looks pedantic and is in fact the whole game, because a small overlap can always fail to find a difference that a larger one would have caught. Done this way, the scale is handed from one lot to the next the way a relay baton passes without the race pausing. Skip the overlap once and the whole library can lurch by the gap between two lots, with no single measurement ever looking wrong.

The cadence between re-qualifications is set the same honest way — by the stability study, not by habit. If the storage allocation in the budget is 1.0% and the standard's own stability study bounds its loss at, say, about 0.1% of potency per month, then the outer limit is simply $1.0\% \div 0.1\%\cdot\text{month}^{-1} = 10$ months: the point at which accumulated drift would just reach its allocation. The working lot is re-qualified comfortably inside that window, on the calendar, before the drift arrives rather than after it has already quietly spent the budget. A reference is retired on a schedule chosen so that even its worst permissible wander still fits under the doubt it was promised to keep.



The controlled chamber where a working standard lives out its stability study — cold, logged, and re-read against its own predecessor before a single aliquot is ever allowed to anchor a number.

The uncertainty budget

A standard is only as trustworthy as the doubt it can state about itself, so we keep a budget for every reference and carry the total openly rather than pretend the anchor is perfect. Each step down the chain adds a little uncertainty, and because these contributions are independent they do not simply add — they add in quadrature, the square root of the sum of squares, which is how independent errors combine:

> $\sigma_{total} = \sqrt{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \ldots)}$.

Worked out for the four real contributors, with representative component figures to show the arithmetic, the reference's budget comes to a little over 2%:

CONTRIBUTOR TO THE REFERENCE'S UNCERTAINTY	HELD BY	CV
Purity & characterization of the primary	exhaustive assay against first principles, documented	1.0%
Aliquoting & dilution to working standards	gravimetric dispensing; verified volumes	1.2%
Storage stability between re-qualifications	cold chain; stability study on the standard itself	1.0%
Re-qualification against the predecessor	overlap testing before any new lot anchors a number	0.8%
Combined, in quadrature	**$\sqrt{(1.0^2 + 1.2^2 + 1.0^2 + 0.8^2)}$**	**≈ 2.0%**

That 2% is not a number we keep to ourselves. It is the floor beneath every score on the axis, and the reason is one more ratio of errors. A score is R ÷ C, and the relative uncertainty of a ratio of independent quantities is itself a quadrature sum:

> $CV(IFI)^2 = CV(R)^2 + CV(C)^2$.

The reference's CV sits inside every score, permanently, under that square root. The best a candidate can ever be reported — even a flawless one, with its own scatter CV(C) driven to zero — is CV(R) itself. Hold the reference at 2% and nothing on that axis can be

known tighter than 2%, ever. That is precisely what a floor means, and it is why the reference is held an order of magnitude more carefully than anything it measures.

REFERENCE CV	CANDIDATE CV	REPORTED SCORE CV = √(CV_REF² + CV_CAND²)
2%	0% (ideal)	2.00%
2%	3%	3.61%
2%	7% (batch spec)	7.28%
1%	7%	7.07%
2%	42% (wild-type)	42.05%

The table reads as relief more than constraint. A candidate held at our ≤7% batch specification, measured against a 2% reference, is reported at $\sqrt{(2^2 + 7^2)} = 7.28\%$ — the anchor adds barely a quarter-point over the candidate's own scatter. Tighten the reference to 1% and the same candidate reports at 7.07%, almost the candidate alone. What the reference cannot do is rescue a sloppy candidate: a wild-type line swinging at 42% reads $\sqrt{(2^2 + 42^2)} = 42.05\%$ no matter how perfect the anchor beneath it. The reference only ever guarantees that it is not *adding* confusion of its own; it can never subtract someone else's.

The daily controls turn that budget into a tripwire. A check sample reads on every plate, on every run, and its action limits sit at ±3σ — at a 2% assay CV, a band of ±6% around target. Anything landing outside it halts the line before a single candidate number is trusted. The 6% is not an accident of the arithmetic: it is the same order as the ≤7% we promise the world, so the control physically cannot pass a plate that would quietly spend the whole batch budget on measurement drift alone. A drift there stops everything, which is the entire point of reading it every single time rather than now and then.



Daily control plates on the assay tray — the check samples that are read on every run, so a shift in the fixed point is caught on the day it happens, not the year a product fails.

Why engineered output can even be held this tight

There are two different consistencies at work here, held by two different disciplines, and it clarifies everything to see that they are separate. The reference's steadiness is held by metrology — the cold chain, the overlaps, the budget. The *candidate's* steadiness, the ≤7% batch CV itself, is held a layer deeper, in the protein, and it is the reason an engineered fruit can be promised at all.

Capsaicin is assembled at the Pun1/AT3 acyltransferase, which condenses vanillylamine with an 8-methyl branched-chain fatty acid — the terminal, pepper-specific step at the far end of a long phenylpropanoid road that every vascular plant already walks. The reference for the heat axis is deliberately the *native* enzyme making *native* capsaicin: the un-engineered genotype, chosen as the anchor precisely because it is the universal, unmoving baseline, the thing no edit has touched. The shipped construct is a double variant, S39L+L345G — two substitutions a protein language model proposed by scoring every possible single-residue change, and that a stability-and-efficiency ranking selected: two edits a full century of breeding-by-phenotype never found:

EDIT	WHAT IT DOES	ΔΔG (KCAL/MOL)	EQUILIBRIUM SHIFT, EXP(−ΔΔG ÷ RT)
S39L	binding affinity ↑, N-terminus	−2.1	33.9×
L345G	steric relief, active-site cleft	−1.8	20.5×
C175S *(structural annotation — not in shipped construct)*	disulfide-misfold prevention	−0.9	4.5×
S39L+L345G shipped double variant (if additive)	**designed fold favored**	**−3.9**	**≈ 697×**

The right-hand column is where consistency is actually purchased. A folding free-energy change ΔΔG — negative for a stabilizing edit — shifts the balance between the correctly folded, catalytically active enzyme and its misfolded dead ends by a Boltzmann factor,

```
K_mut ÷ K_wt = exp(−ΔΔG ÷ RT),
```

with RT ≈ 0.596 kcal/mol at the plant's ~300 K growing temperature. Our structural model also identified C175S as a potential disulfide-misfold prevention site, annotated in the table; it was not carried into the shipped construct after the final design review. The two shipped edits, S39L and L345G, together move the equilibrium toward the designed fold by exp(3.9 ÷ 0.596) ≈ 697× — approximately, because free energies add cleanly only when the residues do not talk to one another, and we verify that empirically rather than trust it to the sum. A misfold branch suppressed by that factor is exactly the kind of stochastic loss that wanders with every degree of temperature and every small shift in cellular redox — the kind of thing that makes a wild-type plant's capsaicin output swing by 42% from one batch to the next. Suppress it, and the swing collapses. A ≤7% batch CV, in the end, is not willed into being. It is the macroscopic shadow of an enzyme that has been told, in the language of kilocalories, to fold the same way every time — which is what lets the metrology hold it to a number at all.

It is worth being exact about which ruler is which. The Scoville column is anchored to *mass* — grams of capsaicinoid, however loud or quiet each one happens to be — while the IFI-H column is anchored to *loudness*, the half-activation of TRPV1 itself. Up this particular ladder the two track each other closely, because capsaicin is a single, well-characterized molecule that is both abundant and loud, so more of it is simply more heat. The receptor anchor earns its

keep on the messier axes, where a food carries several compounds of very different loudness and a mass assay would miscount them all as equal — there the fixed point is not a convenience but the only honest way to add the voices up. Mass and loudness each need their own unmoving reference, and on every axis we keep both.

What the fixed point does not promise

A reference this carefully kept can make two promises and must honestly refuse a third. It promises *precision* — read the same sample twice and get the same number. It promises *consistency* — read it next year, on another continent, from the same revision of the same standard set, and get that number still. What it cannot promise is *accuracy to the whole of your experience*. EC_{50} is a receptor constant; it fixes how loudly a molecule speaks to a channel, not how the mind downstream will hear it. Context, memory, adaptation, the warmth of the room and the people at the table all ride on top of the signal, and the index measures the sound entering the ear, not the music the listener finally makes of it. We anchor the loudness so that the poetry has something true to stand on. We do not pretend the loudness is the poem.

Three honest edges remain, and naming them is part of the discipline:

- **The anchor is a material *and* a method.** EC_{50} depends on the assay it is read in — temperature, the membrane lipid, the expression system — so the standard is really the reference compound plus the procedure that measures it. Change the procedure and the fixed point moves, which is why the method is locked as tightly as the material, and re-qualified with it.
- **There are eight references, not one.** Each axis carries its own anchor and its own budget, and a drift confined to one of them — the heat reference alone, say — tilts only that column while the other seven stay true. Eight independent fixed points are eight separate things to keep still, and any composite built across them is only as trustworthy as its weakest anchor.
- **The free energies are only approximately additive.** Stacking stability edits multiplies their Boltzmann factors cleanly only when the residues do not interact; epistasis is real, so every stacked construct is re-qualified on the bench rather than trusted to the arithmetic that first proposed it.

The scorecard, finally, rails each axis at a normalization maximum of 5,000, and the reference must stay linear and re-qualifiable the whole way up to that rail — a fixed point is no use if the ruler itself bends before the top of the scale. A heat of 4,200 sits well inside the range the reference is characterized to hold straight, with headroom deliberately kept above the hottest thing we have yet grown.

Across more than sixty books and the tens of thousands of indexed pages they generate, the one failure that would read as carelessness is drift — a figure that says one thing here and something else three pages on. We lock the *written* numbers to a single source so no reader ever catches the library disagreeing with itself; the reference standard does the identical work for the *tasted* ones. It is why an

umami score of 2,600 is the same round, brothy, cared-for depth on a Tuesday, in a decade, in a kitchen we will never visit, read from the same revision of the same anchor. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory. Lose the fixed point and the whole sky wheels at once, invisibly — every star still in its familiar place relative to the others, and the entire map quietly wrong. Keep it still, re-qualified by the calendar and not by nerve, its doubt counted and carried in the open, and everything else — every plant we grow toward the fullest flavor, and the fullest nourishment, it can be made to hold — can at last be measured.



PART III OF V

The Mathematics of a Spec

the variance budget · stacking and the 1/√n reward · process capability · the control chart · lot release

Chapter 10 · The variance budget — *spending your tolerance on purpose.*

A finished lot of fruit comes off the line the way a choir arrives at a held note: hundreds of separate voices, each grown in its own corner of the chamber, each ripened on its own afternoon, and the thing that makes it beautiful is that they land together. In our family we have stood over a full tray of sibling fruit and felt the particular quiet of it — not the quiet of nothing happening, but the quiet of everything agreeing. That agreement has a name in the lab. We call it tolerance, and it is the whole subject of this chapter: how much a product is allowed to wander at the mouth before the person eating it feels the difference, and how a formulator gets to decide, on paper and in advance, exactly where that wander is permitted to live.

The single most freeing idea in all of quality is that tolerance is not a wall you press up against and hope to survive. It is a *budget* — a fixed sum of permitted variation, to be spent deliberately across the whole chain from seed to spoon. So much of it goes to the biological scatter of the input we ship you. So much to your own process: the weighing, the blending, the heat of the pan. So much to the assay you run to clear a lot for release. Every one of those is a line item in an account, and the art of the thing is deciding, before you have weighed a single gram, where each percent of the budget is allowed to go. The sum does not behave the way intuition insists it should, and that is precisely why the account must be written down.

The arithmetic of sameness — why percents refuse to add

Begin with the rule, because everything else is a consequence of it. The sources of variation along the chain are independent — the plant's scatter does not know what your scale is doing, and your scale does not know how your assay was calibrated — and independent variations combine not by adding but by adding in *quadrature*: the square of the total equals the sum of the squares of the parts.

> **σ²(product) = σ²(input) + σ²(process) + σ²(release)** — and the lot is in spec when σ(product) ≤ the perceptual tolerance.

This is not a convention we chose for tidiness; it falls out of the definition of variance. For any two random quantities, Var(X + Y) = Var(X) + Var(Y) + 2·Cov(X, Y), and when X and Y are independent their covariance is zero, so the cross term vanishes and the variances simply sum. Because the chain from fruit to spoon is a product of small, independent relative wobbles — the attribute you taste is the input level multiplied by whatever your process does to it multiplied by whatever your assay fails to catch — the natural currency is not the absolute standard deviation but the *relative* one, the coefficient of variation.

> **CV = σ / μ**, usually written as a percent — the scatter measured as a fraction of the mean.

Work in CV and the quadrature rule reads: CV^2(product) ≈ CV^2(input) + CV^2(process) + CV^2(release). We prefer this form for two reasons, one mathematical and one carved into the body. The mathematical one: when an attribute is a chain of multiplied factors, $A = A_0 \cdot (1 + \varepsilon_i)(1 + \varepsilon_p)(1 + \varepsilon_r)$, the small relative deviations ε add to first order, and their variances — the CVs squared — add exactly as the rule says. The bodily one is older than any of our instruments. Perception is relative. Weber's law holds that the just-noticeable difference in a stimulus is a roughly constant *fraction* of the stimulus, not a fixed absolute step: ΔI / I ≈ k. The tongue does not count molecules; it compares ratios. A budget denominated in percents is a budget written in the only units the mouth actually reads.

The consequence is the one that catches every formulator the first time. Hand someone an input held at five percent against a product that tolerates six, and it is irresistible to think a whole percent of headroom remains. It does not. Percents do not subtract; variances do. The input already spends $\sqrt{(5^2)}$ of a budget whose ceiling is six, and the room left for *everything else* — your process and your release assay together — is only $\sqrt{(6^2 - 5^2)} = \sqrt{11} \approx 3.3$ percent. Add one careless blending step at four and the total is not nine, and not six, but $\sqrt{(5^2 + 4^2)} = \sqrt{41} \approx 6.4$ percent — just over the edge, and now the lot fails for a reason that was visible weeks earlier on a single sheet of paper, had anyone thought to spend the budget before spending the fruit.

Where the wander begins — the enzyme that decides how steady a pepper can be

A budget is only as honest as its first line, and the first line is the biology of the input. To know why we designed a pepper to hold a coefficient of variation at or under seven percent — against forty-two percent in the wild-type plant we started from — you have to follow the molecule that makes the heat and ask where its scatter comes from.

Capsaicin is assembled at a single condensation. Vanillylamine, supplied when the aminotransferase pAMT works on vanillin off the phenylpropanoid branch (PAL → C4H → 4CL → HCT → COMT, the backbone every vascular plant already runs), is joined to an eight-methyl branched-chain fatty acyl group drawn from the BCAT and KAS machinery. The two are condensed by Pun1, the acyltransferase sometimes written AT3, and the length and saturation of that acyl chain set the balance between capsaicin and dihydrocapsaicin — and with it the very shape of the burn, its speed of arrival and its tail. Capsaicin then binds TRPV1, the non-selective cation channel on the trigeminal endings of the tongue, the same channel opened by temperatures above 43 °C and by acid. The molecule warms nothing; it lowers the channel's thermal threshold so the body reports a fire that is not there.

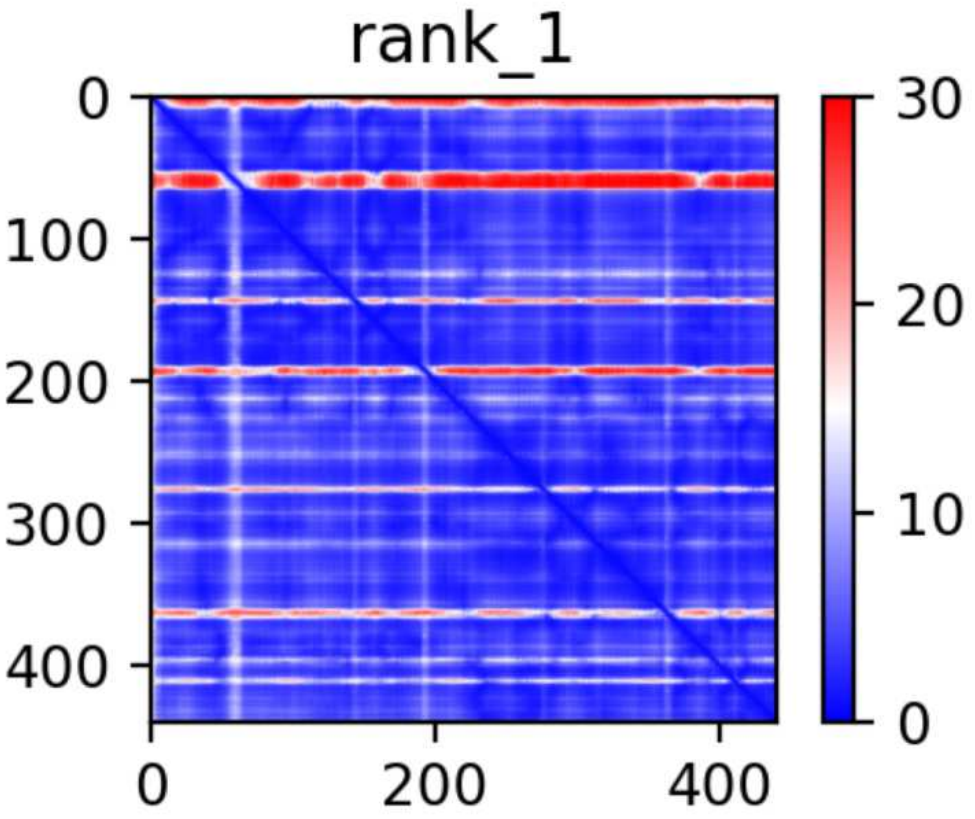


FIG. — Predicted aligned error across the Pun1/AT3 model; the dark diagonal block is the active-site cleft, resolved to near-certainty — the region we had to hold still for the enzyme to hold its output.

Here is the source of the forty-two percent. A biosynthetic pathway is a cascade, and noise in a cascade compounds: if each step's expression fluctuates independently, those fluctuations propagate forward and accumulate. Our flux model predicts that metabolic control in this pathway sits upstream — at 4CL and HCT/CCoAOMT in the phenylpropanoid backbone — not at the terminal Pun1/AT3 step. Pun1 is the final capsaicin synthase; it labels the product, it does not govern the throughput. Yet Pun1's expression is environmentally soft, and because it sits at the end of the chain, every wobble in its activity lands without further dilution in the final capsaicinoid load. It rises and falls with light, with water, with the temperature of a single afternoon, with where on the plant a fruit happened to set and how far into ripening it was picked. Forty-two percent is not carelessness in the field. It is a pathway whose most labile terminal step is also the one least buffered against the weather.

The fix is not to grow the plant more carefully. It is to make the terminal step stop contributing noise. An ESM2 protein-language model scored every single-residue substitution in Pun1 for predicted effect on catalytic efficiency times fold stability; two changes never before combined came through, and CRISPR wrote them in.

EDIT	WHAT IT DOES	ΔΔG (KCAL/MOL)
S39L	lifts substrate-binding affinity at the N-terminus	−2.1
L345G	relieves steric strain in the active-site cleft	−1.8

Read what those two do together and the variance collapse explains itself. A more efficient, more stable enzyme can be driven to operate at substrate saturation — and a saturated step is a buffered step. When an enzyme is working flat-out, the sensitivity of its flux to its own abundance goes slack: the derivative d(flux)/d(enzyme) flattens toward zero, so the day-to-day tremor in expression that used to pass straight through now barely registers. The terminal step stops

being the plant's loudest nerve and becomes its calmest muscle. Our model predicts approximately 2.95 million SHU for the peak design, yes — but the quieter, deeper result is that the batch-to-batch coefficient of variation is designed to fall from forty-two percent to at or under seven. In the language of the budget that is not a small improvement. Forty-two squared is 1,764 units of variance; seven squared is 49. The two edits move the input from 1,764 to 49 — a thirty-six-fold collapse — which is exactly the difference between a fruit that cannot be formulated to any precision at all and one that claims only a modest corner of a well-kept account.



Backlit controlled-growth chamber — identical light, water, and temperature held batch after batch, so the plant never has to argue with its weather.

What the tongue can tell apart — reading the ceiling off a receptor

The top of the budget — the perceptual tolerance, the σ a product is allowed to carry at the mouth — is not a number we get to pick. The receptor sets it. A just-noticeable difference, the smallest change a person can reliably feel, sits for most chemosensory intensities at a few-to-several percent of the baseline, and for heat we run the account against a working tolerance of about six percent: one JND, near enough. Hold a product's total variation under that, and no diner can tell this week's bottle from last week's. This is why we design the input's own CV to land at or under seven percent in the first place — it is built to sit at the edge of a single JND, so that even before your process adds anything, the fruit alone is already quieter than the mouth can resolve.

Measuring where that edge falls means measuring loudness, not content, and loudness is what the Iyengar Flavor Index exists to name. Each axis is anchored to a stated reference constant — set by convention, the way the Scoville scale is defined by dilution-to-threshold and not derived from any particular published measurement — that places IFI = 1,000 on that axis by definition. The index then scores any candidate by comparing its receptor loudness to that defined reference point.

> **IFI(candidate) = 1,000 × EC₅₀(reference) ÷ EC₅₀(candidate)**

EC_{50} is the concentration at which a stimulus drives its receptor to half of full response, and it is read off a dose–response curve. Across a sweep of concentrations the response follows a Hill form, $R(C) = R_{max} \cdot C^n / (EC_{50}^n + C^n)$, where at $C = EC_{50}$ the response is exactly half of R_{max} and n is the Hill slope — the cooperativity of the channel. Capsaicin opens TRPV1 with a slope close to two, a cooperative, switch-like turn-on, and that steepness is a gift to the person measuring it: the curve carries the most information about EC_{50} right where it is steepest, around the midpoint, so doses placed to bracket EC_{50} pin it far more tightly than doses scattered at the flat extremes. The precision of a release assay is, in the end, the precision of this fit — set by how many concentrations bracket the midpoint, how many replicates sit at each, and how clean the plateau is.



A tray of sibling fruit from a single engineered lot on the microscope stage — the release assay reads each one, and they agree.

On the heat axis (H) and pungency axis (P), where IFI-data exists from published EC_{50} measurements, the index is a constant times a ratio of two independently determined values, and the relative errors add in quadrature while the constant drops out:

> $\mathbf{CV^2(IFI) = CV^2(EC_{50}\text{, reference}) + CV^2(EC_{50}\text{, candidate})}$

So if your assay pins each EC_{50} to a coefficient of variation of four percent, the IFI you report carries $\sqrt{(4^2 + 4^2)} = \sqrt{32} \approx 5.7$ percent — already most of a two-percent release line's worth of care, and a reminder that a sloppy reference lot poisons every score read against it. To keep a reported IFI under seven percent you must hold each EC_{50} under about five. The chain of reasoning runs clean from

the channel to the account: the receptor fixes the ceiling, the dose–response fit fixes how tightly you can measure against it, and that measurement precision is the number you write on the release line of the budget.

The budget, written out — the account you keep before you weigh

Make the shape concrete. Take a finished product whose perceptual tolerance is six percent — the ceiling, the total σ it can carry at the mouth. The budget is the square of that, thirty-six "units" of variance, to be parcelled out among the contributors, each claiming the square of its own CV.

LINE ITEM	ALLOTTED CV	VARIANCE CLAIMED (CV²)	SHARE OF THE BUDGET
Input (what we ship you)	3.0 %	9.00	25.0 %
Your process (weigh, blend, heat)	3.5 %	12.25	34.0 %
Your release assay	2.0 %	4.00	11.1 %
Margin held in reserve	3.28 %	10.75	29.9 %
Total	**√36 = 6.0 %**	**36.00**	**100 %**

Read down the variance column and the quadrature logic turns plain. The three working terms — three, three-and-a-half, and two percent — claim only 25.25 of the thirty-six units between them, leaving a genuine reserve of 10.75 units, which is a comfortable 3.28 percent of room still in hand. The reserve is real not because 3 + 3.5 + 2 is a small sum but because their *squares* are, and squares are the currency the budget is denominated in. A line item's cost grows with the square of its CV, which means the cheapest percent to buy back is always the largest one: halving your worst contributor removes four times as much variance as halving your best.

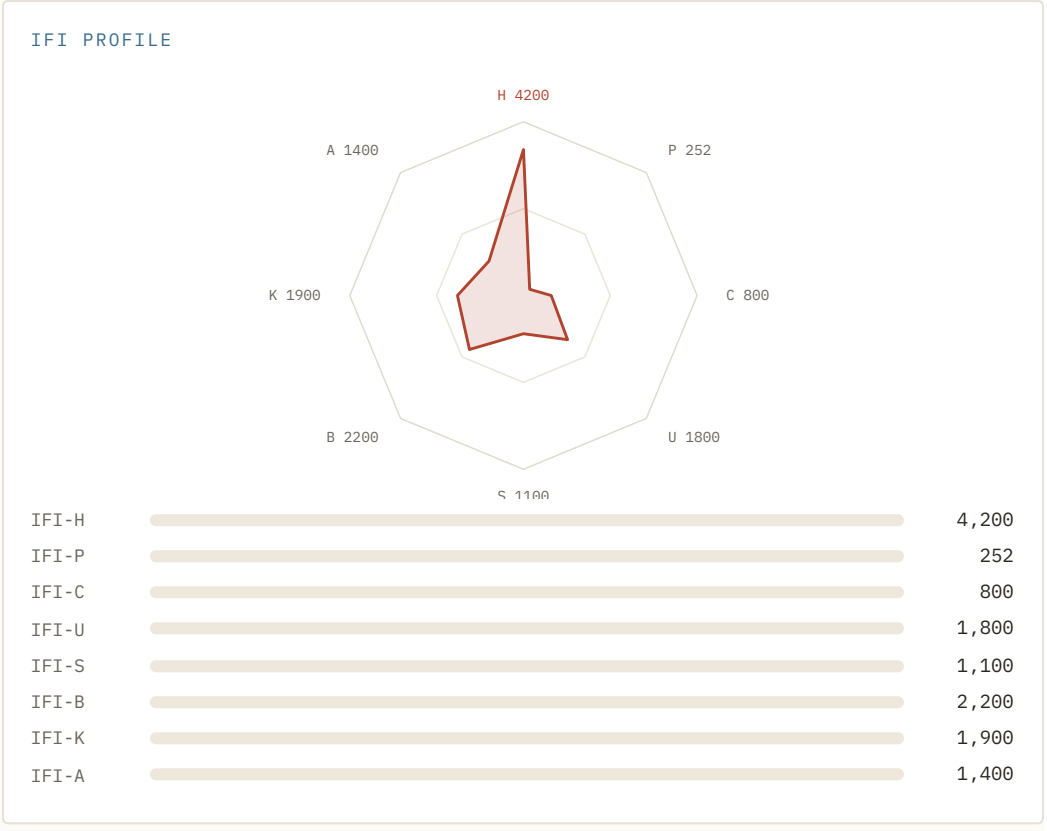
Run the same account backward and it stops being a tally and becomes a design tool — the most useful thing a formulator can do with us. Fix the process and release you can actually hold — say three-and-a-half and two percent — and the question "how tight must the input be?" has a single exact answer: the input may claim no more than $\sqrt{(6^2 - 3.5^2 - 2^2)} = \sqrt{(36 - 12.25 - 4)} = \sqrt{19.75} \approx 4.4$ percent. Any input under that lands the lot in spec; any input looser blows the budget no matter how careful the rest of the line. Generalize it across the tolerances you might be designing to, holding process at 3.5 and release at 2.0, and you can read your maximum affordable input straight off a table.

PRODUCT TOLERANCE (CEILING Σ)	BUDGET (UNITS)	MAX INPUT CV YOU CAN AFFORD
4.03 %	16.25	0.0 % — no input is tight enough
4.5 %	20.25	2.00 %
5.0 %	25.00	2.96 %
6.0 %	36.00	4.44 %
7.0 %	49.00	5.72 %
8.0 %	64.00	6.91 %
10.0 %	100.00	9.15 %

Name your product's tolerance, subtract your own known terms in quadrature, read off the maximum input CV you can afford — then ask us for a lot under it, which, held the way we hold them, we can nearly always supply. The lots that fail in the building almost always failed in the planning, before anyone touched a scale.

Spending it on a real lot — a worked pass, and the limit of inferring loudness from heat

Put the whole account to work on one product. Suppose you are building a finished heat around our most fully specified design tier and you want the mouth to read the same burn, bottle after bottle, to within one JND — a six-percent ceiling, thirty-six units to spend. Our fruit is designed to carry a coefficient of variation of seven percent at the plant; you standardize and blend it down into a base so that the heat-bearing fraction entering your product carries an input CV of 3.0 percent. Your line — weigh, blend, hold the cook temperature — you have characterized at 3.5 percent. Your release assay, an EC_{50} read against a clean reference, sits at 2.0 percent. The account closes at $\sqrt{(3.0^2 + 3.5^2 + 2.0^2)} = \sqrt{25.25} \approx 5.02$ percent, comfortably under the six-percent ceiling with 10.75 units of reserve still in hand. The lot clears not because any one step was heroic but because the squares were small and the plan was made first.



That profile is Phoenix (SS-0096), our flagship full-spectrum cultivar designed to register on all eight axes at once — and it carries a lesson about the difference between what a product is designed to hold and what a mouth hears. Each axis score above is calibrated against that axis's own stated reference constant; the figures across axes are not on a shared common scale and cannot be directly compared to one another. Walk the flagship design tier table and

watch the two scales — predicted SHU, which counts capsaicinoid content, and IFI-H, which counts loudness at TRPV1 — refuse to stay proportional.

CULTIVAR	SKU	PREDICTED SHU	IFI-H	AXES	PREDICTED SHU PER IFI POINT
Ember	SS-0087	866 K	1,570	2/8	552
Cinder	SS-0088	1.05 M	2,290	3/8	459
Kindle	SS-0089	1.25 M	2,000	3/8	625
Flare	SS-0090	1.48 M	—	4/8	–
Blaze	SS-0091	1.72 M	2,630	4/8	654
Wildfire	SS-0092	1.96 M	2,930	5/8	669
Scorch	SS-0093	2.21 M	3,220	5/8	686
Firestorm	SS-0094	2.47 M	3,490	6/8	708
Solara	SS-0095	2.72 M	3,980	6/8	683
Phoenix	SS-0096	2.95 M	4,200	8/8	700

The ratio in the last column is not constant. It climbs from roughly 550 predicted SHU per IFI point near the lower design tiers to about 700 near the upper — which is to say loudness grows *sub-linearly* with content, exactly the signature of a receptor approaching saturation. Capsaicin does not get a linear hearing; as the channel nears R_max, each additional increment of capsaicinoid buys a smaller increment of felt heat. And the column is not perfectly smooth, either — Cinder sits low, Kindle high — because these genotypes differ in the capsaicin-to-dihydrocapsaicin balance that sets the shape of the burn, and two fruits at the same predicted SHU can speak to the nerve with different loudness. That scatter is not noise to apologize for; it is the entire argument for measuring at the receptor. Content is what a pepper is designed to hold. The index is what the body hears, and only the second one is the thing you are actually trying to keep the same.

Full spectrum raises the stakes of that sameness in a way worth naming, because it changes the budget from one account into eight coupled ones. Phoenix speaks on all eight axes — heat at TRPV1, pungency at TRPA1, cool at TRPM8, umami at mGluR1, sweetness at TAS1R2·R3, attenuated bitterness at TAS2R38, kokumi at CaSR, astringency at the salivary proline-rich proteins — and a lot is only "consistent" if every one of those lands within its own JND at once. Were the eight independent, the chance of clearing all of them together would be the per-axis pass rate raised to the eighth power: even at ninety-five percent per axis, $0.95^8 \approx 0.66$, and a third of lots would fail on something. Push each axis well inside its tolerance, to a per-axis pass rate of 0.997, and $0.997^8 \approx 0.98$ — the margin you buy on each axis is what keeps the whole fruit in spec. The eight are not truly independent; they draw on shared phenylpropanoid flux, so they rise and fall together, and that coupling is the quiet reason stabilising the terminal capsaicin-forming step pays dividends across all eight. Hold that final step steady, and you do not buy one axis's consistency — you buy all eight in a single act.

When the budget breaks — the honest edges of the account

A budget is a model, and the discipline of using one is knowing where it stops being true. Four edges matter.

The first is a hard floor, and it is in the table already. Hold process at 3.5 and release at 2.0 and those two alone claim $\sqrt{(3.5^2 + 2.0^2)} = \sqrt{16.25} \approx 4.03$ percent of variance before the fruit contributes anything. If a product demands a tolerance tighter than 4.03 percent, no input on earth can rescue it — even a perfect, zero-variance lot leaves you over budget. The move then is not to source a better pepper; it is to tighten the process or the assay first, because you are trying to pour four units of water into a three-unit glass and the glass is the problem.

The second edge is the assumption the whole rule rests on: independence. Quadrature adds variances cleanly only because the cross term, $2 \cdot \mathrm{Cov}(X, Y)$, is zero — and it is zero only when the sources genuinely do not move together. Weigh your blend and read your release assay on the *same* miscalibrated balance, or score every lot against a reference EC_{50} that is slowly drifting, and the errors become correlated, the covariance turns positive, and the true variance is larger than the sum of squares the budget counts. The account then passes a lot it should have failed, silently. The guard against it is structural: independent references, cross-calibration between instruments, a standard measured fresh rather than assumed.

The third edge is the deepest, and it is the difference between scatter and drift. A variance budget governs how a lot wobbles *around its mean*. It says nothing about the mean itself. If your reference lot ages and its EC_{50} shifts, every score slides in the same direction at once — that is bias, not variance, and no amount of reserve protects against it, because the budget was never watching the center. A process can be beautifully precise and steadily wrong. Keeping the mean honest is a separate practice from keeping the scatter small, and a product needs both: a tight CV around a true target, re-anchored to a fresh reference often enough that the target cannot creep.

The fourth is a limit on the budget's own numbers. Every CV in the account is itself an estimate from a finite run of batches, and a CV estimated from too few is barely a number at all. The relative uncertainty in an estimated CV runs about $1/\sqrt{(2(n - 1))}$: from eight batches you know your own CV only to within roughly twenty-seven percent of itself, which means a "3.5 percent" process line might truly be anywhere from about two-and-a-half to four-and-a-half. To pin a CV to within ten percent of its true value takes $n \approx 51$ — about fifty lots, not five. This is why we treat "at or under seven percent" as a design specification earned across tens of batches rather than asserted from a handful, and why the margin held in reserve is not timidity but arithmetic honesty: it is the room the account keeps for the fact that it does not know its own inputs perfectly.

Spend the budget on paper first, in quadrature, with the receptor setting the ceiling and the enzyme setting the floor — and the physical lot has nowhere left to surprise you. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory. That is the whole of what we mean by consistency: not that nothing varies, which no living thing allows, but that every percent of the variation

was accounted for, assigned, and kept within the hair's breadth the mouth can feel. A tray of fruit that all agree is not an accident. It is a budget, kept.

Chapter 11 · Stacking — *how variation shrinks when you combine.*

Here is the happiest piece of arithmetic in the book, and the one most formulators have exactly backwards. When you combine several of our inputs into one product, their coefficients of variation do not add, and they do not even hold steady. For independent inputs, they *shrink.* The whole becomes steadier than any of its parts — the opposite of what intuition, braced for errors to pile up, expects. A blend of things each wandering a little lands truer than any one of them alone, and it lands truer the more of them you trust. We have watched this turn a nervous formulation — a sauce that read a shade different every week — into something a line could set its clock by, and the whole of the reason is a square root. It is the kind of result our family loves most: a gift the mathematics hands you for free, if you have the discipline to earn it with inputs worth combining.

We should say plainly what a *scatter* is, because everything here rests on one number. The coefficient of variation, written CV, is the standard deviation of a measurement divided by its mean — the width of the wobble expressed as a fraction of the level itself. A capsaicinoid load that averages some figure and swings by seven percent of it from batch to batch has a CV of seven percent. Our engineered lines hold **CV ≤ 7%**; the wild plant they descend from wanders at **42%**. That sixfold gap is the quiet subject of this whole book, and stacking is the multiplier that makes a tight input tighter still.

The rule beneath the gift — variance in quadrature

Start from what a blend actually is. Write the delivered property of a product — a heat, a sweetness, a brothy depth — as a weighted sum of the inputs that carry it:

> $B = \sum w_i \cdot X_i$, with the weights summing to one, $\sum w_i = 1$.

Each X_i is one input's contribution, a quantity that is not a single fixed value but a small cloud: a mean μ_i it is supposed to hit and a standard deviation σ_i it wobbles by. The weight w_i is that input's share of the finished profile. Two laws of expectation carry us the rest of the way, and they are worth stating because the entire reward descends from them. The first is that means combine straight — averages are linear — so the blend's mean is just the weighted mean of the parts:

> $\mu(B) = \sum w_i \cdot \mu_i$.

The second is the one with the teeth. Scaling a random quantity by a constant scales its *variance* by the constant **squared** — $\mathrm{Var}(a \cdot X) = a^2 \cdot \mathrm{Var}(X)$ — and, for inputs whose wobbles are independent, variances simply add. Put those together and the blend's variance is a sum of squared weights against squared scatters:

> **$\sigma^2(B) = \Sigma\ w_i^2 \cdot \sigma_i^2$** — contributions squared, scatters squared, summed.

This is the quadrature rule, and the squares are the whole story. A scatter does not enter the blend at its face value; it enters as its square, scaled by the square of the weight it is carried at. The full law has one more term we will come back to — a covariance that appears when inputs are *not* independent — but hold that thought: for inputs that genuinely do not share a fate, the cross-terms vanish and the clean sum above is exact. Everything generous about stacking lives in those two squarings.



A research plot of independent lots — many plants, many drying trays, many plates, each wandering on its own.

From scatter to the square root of *n*

Follow the rule through for the cleanest case and something lovely falls out. Take *n* inputs, each held at the same scatter and each contributing an equal share of the whole — identical clouds, identical weights. Equal shares means every weight is $w_i = 1/n$. Equal clouds means every input has the same mean μ and the same CV, call it *c*, so each $\sigma_i = c \cdot \mu$. Drop those into the two laws.

The mean is untroubled: $\mu(B) = \Sigma\ (1/n) \cdot \mu = \mu$. An average of things that each aim at μ still aims at μ. The variance is where the gift appears:

> $$\sigma^2(B) = \Sigma\ (1/n)^2 \cdot \sigma^2 = n \cdot (1/n^2) \cdot \sigma^2 = \sigma^2/n.$$

The *n* copies each contribute a variance scaled by $1/n^2$, and there are *n* of them, so the pile collapses to σ^2/n. Take the square root for the standard deviation, $\sigma(B) = \sigma/\sqrt{n}$, and divide by the mean to return to a CV:

> **$CV(B) = CV(input) / \sqrt{n}$** — for *n* equal, independent, well-held inputs.

Two inputs, each wandering at six percent, combined in equal measure, do not give a blend at six. They give a blend at $6/\sqrt{2} \approx$ **4.2 percent.** Four such inputs give $6/\sqrt{4} =$ **3.0.** The errors do not march in step; being independent, they partly cancel — a lot that ran a little hot on one input tends to meet a lot that ran a little cool on the other, and the blend lands nearer the mark than either parent could alone. Picture it in the concrete: one drying tray comes off two percent

warm, the next two percent cool, and the mash made from both reads dead level, because the two misses were never agreeing with each other to begin with. Diversify an intensity across several things you trust and the whole settles toward its target more surely the more of them there are.

BUILD	INPUTS *N*	EACH AT CV	BLEND CV
Single input	1	6.0 %	6.00 %
Two, equal weight	2	6.0 %	4.24 %
Three, equal weight	3	6.0 %	3.46 %
Four, equal weight	4	6.0 %	3.00 %
Five, equal weight	5	6.0 %	2.68 %
Six, equal weight	6	6.0 %	2.45 %

It does not even matter whether you think of the blend as an average of intensities or as a running total of mass. Sum *n* equal independent contributions instead of averaging them and the total's mean grows like *n* while its standard deviation grows only like $\sqrt{n}$, so the ratio of the two — the CV — still falls as $c/\sqrt{n}$. The scatter is simply slower to accumulate than the signal. That is the engine under every row of the table: signal adds in proportion to *n*, noise adds only in proportion to its square root, and the square root always loses the race.

The effective number of inputs

Real blends are never four things in equal quarters. They are lopsided — a dominant base, a couple of supporting players, a whisper of accent — and the clean $\sqrt{n}$ reward has to be generalized to earn its keep in a real formulation. Go back to the master rule with equal clouds but unequal weights. Every input still has CV *c*, so $\sigma_i = c\cdot\mu$, but the weights are whatever they are. Then:

> $\sigma^2(B) = \Sigma\, w_i^2\cdot c^2\cdot\mu^2 = c^2\cdot\mu^2\cdot(\Sigma\, w_i^2)$, so $CV(B) = c\cdot\sqrt{(\Sigma\, w_i^2)}$.

The sum of squared weights, $\Sigma\, w_i^2$, is the quantity that decides how much of the gift you collect. Its reciprocal has a name worth knowing — the **effective number of inputs**, $n_{eff} = 1 / \Sigma\, w_i^2$ — and it lets us write the general reward in the same shape as the clean one:

> $CV(B) = c / \sqrt{(n_{eff})}$, where $n_{eff} = 1 / \Sigma\, w_i^2$.

For four equal quarters, $\Sigma\, w_i^2 = 4\cdot(1/4)^2 = 1/4$, so $n_{eff} = 4$ and the formula returns $c/2$ exactly. But fold a formulation lopsidedly and n_{eff} tells you the truth the headcount hides. A blend that is 90 percent one input and 10 percent another has $\Sigma\, w_i^2 = 0.81 + 0.01 = 0.82$, so $n_{eff} = 1.22$ — barely more than a single input, however many bottles are on the bench. Its CV is $c\cdot\sqrt{0.82} = 0.906\cdot c$: at six percent inputs, **5.4 percent** out. You poured two things and the math counted one and a quarter.

Here is the lesson made into a working method. Suppose a formulator wants a finished mash at **IFI-H ≈ 3,200** — the register of our *Scorch* line — and asks how steady they can make it. Leaning

the whole product on *Scorch* (SS-0093, IFI-H 3,220) held at seven percent gives a blend CV of seven percent; there is one input, and one input collects nothing. Now hit the same target by weighting two independent lines across it. *Blaze* (SS-0091) reads 2,630 and *Firestorm* (SS-0094) reads 3,490; solve 2,630·w + 3,490·(1−w) = 3,200 and the weights fall out at w = 0.337 and 0.663. Those squared give $\Sigma\, w_i^2 = 0.553$, so n_eff = 1.81, and the blend CV is $7\% \cdot \sqrt{0.553}$ = **5.2 percent** — same heat, a third of the wobble gone, for the price of trusting a second independent lot. Spread that same 3,200 across three roughly equal independent lines — *Wildfire* (2,930), *Scorch* (3,220), *Firestorm* (3,490), which average 3,213 — and the blend CV falls toward $7/\sqrt{3}$ = **4.0 percent.**

ROUTE TO IFI-H ≈ 3,200	WEIGHTS	INPUTS EACH AT CV	Σ w_i^2	N_EFF	BLEND CV
Scorch alone	1.00	7 %	1.000	1.00	7.00 %
90/10 dominant + accent	0.90 / 0.10	7 %	0.820	1.22	6.34 %
Blaze + *Firestorm*	0.337 / 0.663	7 %	0.553	1.81	5.21 %
Three lines, near-equal	0.33 / 0.33 / 0.34	7 %	0.334	2.99	4.04 %

The practical reading is sharp: your tightest control belongs on whatever input dominates the blend, because its weight enters *squared*, while the minor accents may run looser than you would ever accept alone and still vanish beneath that weighting. An input carrying 70 percent of a profile contributes 0.7^2, or very nearly half, of its own variance to the blend; a 10 percent accent contributes 0.1^2 — a hundredth. A ten-percent garnish at eight percent CV adds $0.1^2 \times 8^2 = 0.64$ units of variance, a scatter of eight-tenths of a percent, noise beneath notice; the same eight percent carried on the dominant input would sink the lot.

What the receptor does with the scatter

It is worth asking *why* a tight concentration is worth this much trouble, and the answer lives at the receptor, where loudness is actually read. Capsaicin binds TRPV1, a non-selective cation channel on the trigeminal nerve endings of the tongue — the same channel gated by temperatures above 43 °C and by protons. The molecule does not warm the tissue; it settles into the vanilloid pocket, lowers the channel's thermal threshold, and the open channel lets cations flood the nerve ending until it fires and the brain reports a fire that is not there. The quantity the Index anchors to is the **EC_{50}** — the concentration of the preparation at which the receptor reaches half of its full response — and the whole scoring rule inverts it:

> **$IFI(candidate) = 1{,}000 \times EC_{50}(reference) \div EC_{50}(candidate).$**

Read that as loudness, not quantity. A smaller EC_{50} means less material is needed to half-open the channel, which means the food speaks *louder*, which the thousandfold-inverted ratio turns into a larger score. Anchor the native wild-type reference at 1,000 and the arithmetic is immediate: our *Phoenix* registers IFI-H 4,200, so EC_{50}(Phoenix) = 1,000 ÷ 4,200 × EC_{50}(ref) = **0.238 × EC_{50}(ref)** — a quarter of the material to reach the same half-activation, a fruit our model predicts to be 4.2 times as loud per gram as the plant we

began from. The scale has a deliberate ceiling at the per-axis normalization maximum of 5,000, which corresponds to EC_{50}(ref)/5, a fifth of the reference concentration; past five times as loud the axis saturates and we stop crediting raw intensity.



The sequencer bench — where loudness stops being a taste and becomes a number with an error bar.

Now connect that to the wobble. The receptor's response does not rise in a straight line with concentration; it follows the cooperative S-curve of the Hill equation, R = R_max·[A]^h ÷ (EC_{50}^h + [A]^h), where *h* is the cooperativity of binding and capsaicin's grip on TRPV1 is cooperative, h greater than one. The curve is at its very steepest exactly at the half-activation point — the slope there, per natural-log step in concentration, is R_max·h/4 — which means a formulation held *at* its EC_{50} sits on the most sensitive stretch of the whole response. A fractional slip δ in concentration near that point moves the response by roughly (h/4)·δ of the maximum; take the standard modeling value h = 2 and the gain is about one-half, so a concentration wobble is read out by the nerve at roughly half its size in perceived intensity. The number matters less than the direction: the Hill nonlinearity *amplifies* concentration noise right where you operate, so shrinking a concentration CV before it reaches the receptor is not bookkeeping — it is the difference between a heat that reads the same twice and one that does not. Every √n you win on the bench is a √n the tongue never has to forgive.

Two honest conversions keep the numbers physical. First, loudness per gram is capsaicinoid loading, and the old dilution assay still speaks in it: at roughly 16 Scoville Heat Units per part-per-million of capsaicinoids, our model predicts approximately 2.95 million SHU for Phoenix, which works out to about 183,600 ppm, near **18 percent capsaicinoids by mass** — a fifth of the fruit, by weight, given over to a single family of molecules. Second, SHU and IFI-H are cousins, not twins. Both climb with loudness, but SHU is a crude dilution-to-threshold count and IFI-H is pinned to a receptor EC_{50}, so their ratio drifts rather than holding constant, and the IFI is the one we formulate against.

CULTIVAR	SKU	PREDICTED SHU	IFI-H	SHU ÷ IFI-H
Ember	SS-0087	866,000	1,570	552
Blaze	SS-0091	1,720,000	2,630	654
Solara	SS-0095	2,720,000	3,980	683
Phoenix	SS-0096	2,937,948	4,200	700

The last column is the honesty: it moves from 552 to 700 across this design progression, a monotone drift, not a constant, because a receptor curve and a dilution count are scoring loudness through different instruments. We publish both and trust the one anchored to the channel.

The loudest wanderer — why one commodity input undoes a stack

Now the warning, and it is why the stacking math is also an argument about *which* inputs deserve to be in the blend at all. The reward of $1/\sqrt{n}$ is real only when every input is both independent and genuinely well-held. Fold a single commodity ingredient into an otherwise tight stack — one wide, unknown scatter carried at a meaningful weight — and its term, large and squared, swamps the sum and drags the whole product's CV up to meet its own, undoing the patient tightness of everything around it. A blend is only ever as steady as its loudest wanderer.

Put our two canonical scatters against each other to see the violence of it. Three engineered inputs at seven percent, equal weight, give $7/\sqrt{3}$ = 4.0 percent — a lovely, line-ready number. Replace just one of the three with a wild-type commodity at 42 percent and recompute from the rule: $\sigma^2(B) = (1/3)^2 \cdot 7^2 + (1/3)^2 \cdot 7^2 + (1/3)^2 \cdot 42^2 = (49 + 49 + 1{,}764) \div 9 = 206.9$, so the blend CV is **14.4 percent.** One input, at a third of the weight, turned a 4.0 percent product into a 14.4 percent one. Even trimmed to a tenth of the blend the commodity is ruinous: at weights 0.45, 0.45, 0.10 the variance is $0.45^2 \cdot 7^2 + 0.45^2 \cdot 7^2 + 0.10^2 \cdot 42^2 = 9.92 + 9.92 + 17.64 = 37.5$, a CV of **6.1 percent** — and notice that the commodity's single term, 17.64, outweighs the two engineered terms put together, 19.84, while carrying a fifth of their combined weight. That is the squared scatter doing its arithmetic in plain sight.



A controlled chamber — a well-held input is one grown where nothing is left to wander.

THREE-INPUT STACK	WEIGHTS	INPUT CVS	BLEND CV
All engineered, equal	0.33 / 0.33 / 0.33	7 / 7 / 7 %	4.04 %
One commodity at a third	0.33 / 0.33 / 0.33	7 / 7 / 42 %	14.38 %
Commodity trimmed to a tenth	0.45 / 0.45 / 0.10	7 / 7 / 42 %	6.12 %

Replace the commodity with a specified input held under seven, give it good company, and the stack tightens itself for nothing. This is the quiet reason our certificates matter: a number held to ≤ 7% is not merely a tighter version of a number held at 42% — dropped into a

blend, it is a fundamentally different *kind* of citizen, one whose square stays small instead of eating the room. Consistency, built on consistency, compounds; consistency built on one commodity shortcut does not compound at all.

When the inputs hold hands — covariance and shared fate

The second word the whole reward rests on is *independent,* and the world does not always grant it. Return to the rule in its complete form, with the cross-terms we set aside restored:

> $\sigma^2(B) = \Sigma\, w_i^2 \cdot \sigma_i^2 + 2 \cdot \Sigma_{i<j}\, w_i \cdot w_j \cdot \rho_{ij} \cdot \sigma_i \cdot \sigma_j.$

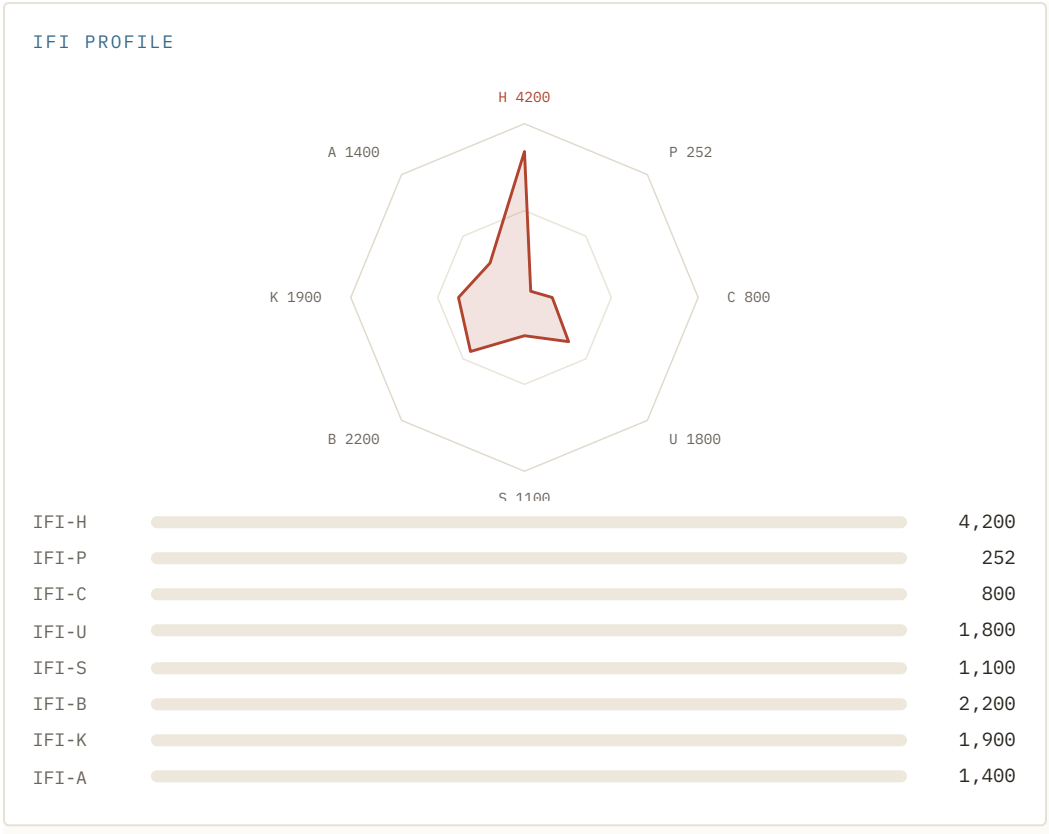
That second sum is covariance, and ρ is the correlation between two inputs' wobbles — zero when they drift independently, one when they move in perfect lockstep. When ρ is zero the cross-terms vanish and we are back to the clean quadrature and its √n gift. When ρ is positive the scatters reinforce rather than cancel, and the gift shrinks; when ρ reaches one it disappears outright. Two equal inputs at six percent show the full span. Independent, ρ = 0: $6/\sqrt{2}$ = **4.2 percent.** Half-correlated, ρ = 0.5: $\sigma^2(B) = \sigma^2 \cdot (0.25 + 0.25 + 2 \cdot 0.25 \cdot 0.5) = 0.75 \cdot \sigma^2$, a CV of **5.2 percent.** Fully correlated, ρ = 1: $\sigma^2(B) = \sigma^2 \cdot (0.25 + 0.25 + 0.5) = \sigma^2$, a CV of **6.0 percent** — no cancellation at all, two bottles behaving as one.

TWO EQUAL INPUTS, EACH 6 % CV	CORRELATION P	BLEND CV
Truly independent	0.0	4.24 %
Partly coupled	0.5	5.20 %
Shared fate	1.0	6.00 %

Where does correlation sneak in? Through shared causes. Two inputs drawn from the same upstream seed lot, dried in the same oven on the same drifting afternoon, read on the same miscalibrated plate, are not independent — their common history couples them, and a blend of coupled inputs can wander as much as either parent while wearing the costume of a diversified formulation. This is a quiet argument for diversity of *source* as much as diversity of kind: do not rest a profile on inputs that secretly share a fate, because the √n you think you are collecting may be a 1 you have paid for twice. It is why our certificates name the lot and the reference revision for every axis — so a formulator can see which of their inputs are genuinely independent and which are holding hands behind the blend's back. The clean reward is a reward for *real* diversity, and the mathematics refuses to be fooled by the appearance of it.

The stack made flesh — eight axes at once

There is a fruit in which all of this stops being a method and becomes a plant you can hold. *Phoenix* (SS-0096) is the first cultivar we designed to register on all eight axes at once — a single *C. chinense* tuned to speak heat, pungency, cool-burn, umami, sweetness, zero-bitter, kokumi, and astringency in one bite, every axis held to the same ≤ 7% discipline:



It is the stack made flesh, and it also teaches the one edge case that keeps the √n gift from being misapplied. The eight axes do *not* average against one another — they are different measurements at different receptors (TRPV1 for heat, TRPA1 for pungency, TRPM8 for cool-burn, mGluR1 for umami, and so on down the profile), and a cool-burn miss cannot cancel an umami miss because they are not the same quantity. Stacking shrinks the scatter *within* an axis, across the independent sources that build it; it does nothing *across* axes, which must each be held tight on their own. So the eight numbers above are not eight draws from one lottery that the blend averages down — they are eight separate promises, each kept by its own well-held pathway.

What lets each promise be kept is engineered a long way upstream, at the fold of a single enzyme. The two substitutions we wrote into Pun1 — **S39L** and **L345G** — were chosen for stability as much as for output, with ΔΔG values of −2.1 and −1.8 kcal/mol, a net of −3.9 kcal/mol of added fold stability. A more stable enzyme is a more *reproducible* one: it folds the same way, catalyzes at the same rate, and pours out the same loudness batch after batch, and that per-batch steadiness is exactly the low CV that makes each axis a citizen worth stacking. The ≤ 7% at the bottle is, in the end, a consequence of a few kilocalories of stability designed in at the residue — tightness all the way down.

Honest limits

The square-root gift is real, and it is also easy to oversell, so we hold five honesties in view. First, $1/\sqrt{n}$ assumes both independence and equal well-heldness, and the world grants neither for free — a hidden correlation or one loose input is enough to erase the reward, as the arithmetic above shows without mercy. Second, the CV you quote is itself an estimate with its own scatter: for a roughly normal process the standard error of an estimated CV runs near $\mathrm{CV}/\sqrt{(2N)}$ over N batches, so a seven-percent figure read from only five batches carries a ±2.2 percent uncertainty of its own, and even twenty batches leave about ±1.1 percent — a bound is only as trustworthy as the count behind it. Third, quadrature is a statement

about second moments alone; it tames the everyday wobble and says nothing about the rare excursion, the once-a-year lot that fails for a reason no variance captured, which is why we gate as well as average. Fourth, the EC_{50} picture is a local one — the Hill slope we lean on is steepest and most faithful near half-activation, and a formulation driven far up or down its curve reads differently. Fifth, SHU and IFI-H drift against each other by design, and we trust the receptor-anchored number when they disagree.

None of these limits dims the central fact, and the central fact is a kindness the arithmetic keeps handing back. Reach for a second trusted, independent input instead of leaning the whole product on one, and the blend settles toward its target more surely than either parent could manage alone — steadier the more good company you give it. It is the dividend of designing with numbers rather than guessing with averages, paid out quietly every time a formulator chooses diversity of source over a single loud bet. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory. We build it this way because a thing grown and blended with this much care ought to arrive the same, and as nourishing, every single time it reaches a table.

Chapter 12 · Process capability — *Cp, Cpk, and the sigma levels.*

A coefficient of variation tells you how much a thing wanders. It does not, by itself, tell you whether that wander *fits* — whether the scatter is small enough to live comfortably inside the window your product is actually allowed to occupy. A broth we love is not merely one that tastes right once; it is one that tastes right every winter it is made, from a jar pulled at random off a line that never saw the hand that filled the one before it. For that kind of promise, manufacturing has a pair of indices every formulator who has shipped at scale will recognize, and that translate our CV into the language of a release spec: the process capability indices, **Cp** and **Cpk.** This chapter is where the private virtue of a small, honest variance becomes a public, contractual thing — a number a buyer can hold us to.



The Phoenix controlled chamber, where temperature, light, and nutrient are held constant so that what varies in the fruit is only what the genome decides — the physical meaning of a small sigma.

The window, and why it is six sigma wide

Begin with the window. A finished product carries a specification with an upper and a lower limit — call them USL and LSL — the band inside which a lot is acceptable. Everything outside it is scrap, rework, or a complaint. Capability is the single question of how that band compares to the process's own spread, and it is answered by laying one width against the other.

The denominator of every capability index is six standard deviations, and the six is not an arbitrary round number. For a normally distributed process, the interval from three standard deviations below the mean to three above it contains 99.73 percent of all output — essentially the whole population, with only 0.27 percent, about 2,700 parts per million, left in the two tails combined. So **6σ** is, to a very good approximation, the *full working width* of the process: the span you must find room for if you intend to catch nearly every unit. Cp is simply the ratio of the width you are allowed to the width you occupy:

```
Cp = (USL − LSL) / 6σ
```

A Cp of 1.00 means the process spread exactly fills the window — tolerable, but with no room at all for the center to move. A Cp of 1.33 means the window is a third wider than the spread, the classic comfortable margin that leaves a cushion on either side. A Cp of 2.00 means the window is twice the spread, and a lot would have to drift catastrophically before it touched a limit.

But Cp is naïve, and in a way that matters. It assumes the process sits perfectly centered in its window, and real processes never do. A reagent lot shifts; a greenhouse runs a degree warm through a heat wave; a new technician reads an endpoint a shade early. The spread can be beautifully tight and the whole distribution can still be sliding toward one wall. Cp, measuring only width, is blind to this; it would report a sliding process as perfectly capable right up until the day it fails. Cpk fixes the blindness by measuring the distance from the mean to the *nearer* limit, in units of 3σ:

```
Cpk = min[ (USL − μ), (μ − LSL) ] / 3σ
```

Cpk is the honest index, because it punishes a process that is both wide *and* off-center — which is exactly the process that fails in the field while its paperwork still looks clean. When the process is perfectly centered, the two distances are equal, each equal to the half-window, and Cpk collapses back to Cp. As the mean drifts toward either wall, Cpk falls below Cp and keeps falling. The gap between them is a pure measure of how far off target you are living.

Centering, in one clean relationship

It is worth seeing exactly how centering erodes capability, because the arithmetic is simple and the intuition is permanent. Let the half-window be H = (USL − LSL) / 2, and let δ be the distance from the process mean to the midpoint of the window. The nearer limit then sits a distance H − δ from the mean, so:

```
Cpk = (H − δ) / 3σ = Cp · (1 − k), where k = δ / H
```

The centering factor k runs from 0, dead center, to 1, mean parked exactly on a limit. A process with an enviable Cp of 2.00 that has drifted a quarter of the way to its wall — k = 0.25 — delivers a Cpk of 2.00 × 0.75 = 1.50. Drift halfway, k = 0.50, and the same tight

process reports Cpk 1.00, marginal, one bad week from scrap. The spread never changed. Only the aim did. This is why we hold the center and the width as two separate disciplines, and why a formulator who reports only one of them has told you only half of what keeps their line honest.

Where the variance is born

The σ in those denominators is not a statistical abstraction we assign; it is the macroscopic shadow of a molecular event repeated across millions of cells in a fruit, and the whole reason we can make it small is that we understand where it comes from. Capsaicin is assembled when vanillylamine is condensed with an 8-methyl branched-chain fatty acyl group at the Pun1/AT3 acyltransferase — one enzyme, catalyzing one acyl transfer, at the terminal step of the capsaicinoid pathway. Every molecule of heat a pepper will ever carry passes through that active site. Our flux model places the control over total capsaicinoid output upstream, at 4CL and HCT/CCoAOMT in the phenylpropanoid pathway; Pun1 is the final capsaicin synthase — it labels the product, it does not govern the rate. What *does* depend on Pun1's reliability, fruit to fruit, is the *scatter* in the amount — our σ — the summed noise of everything that can make one enzyme molecule behave unlike its neighbor.

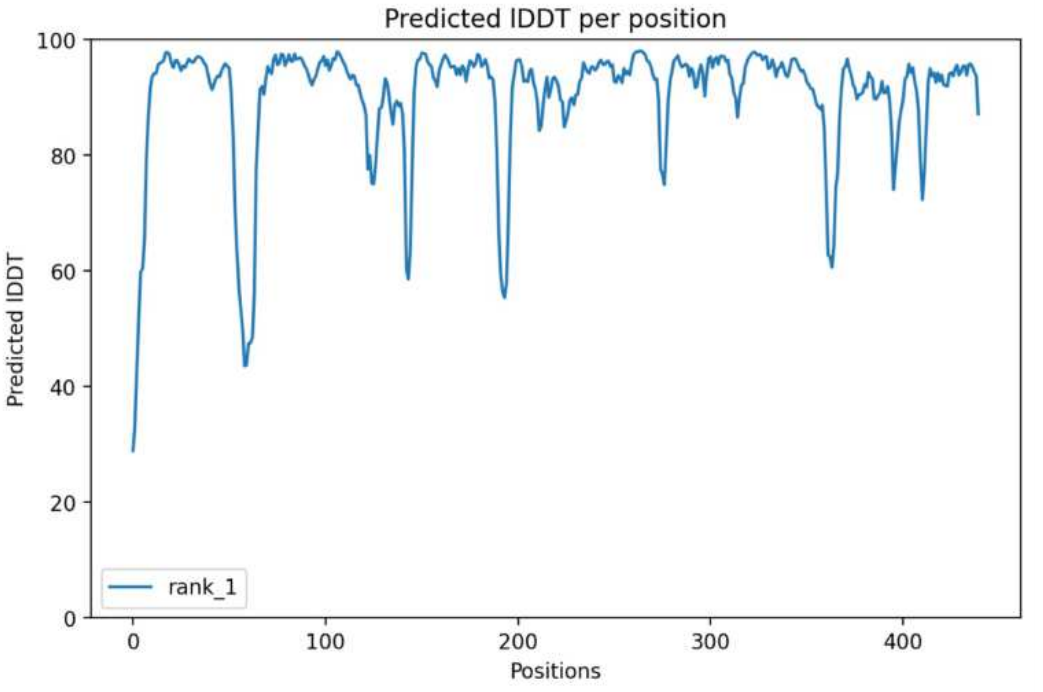


FIG. 12 — the Pun1/AT3 acyltransferase coloured by pLDDT confidence. The active-site cleft and the N-terminal binding domain resolve sharply; the mutable, variance-bearing loops are exactly the low-confidence regions we redesigned.

In wild-type pepper, that noise is enormous: a batch-to-batch CV of **42 percent**, which is to say a plant whose heat you genuinely cannot predict within a factor of two. Three distinct molecular lotteries feed that number. The substrate pools fluctuate, so the active site is sometimes starved and sometimes flooded. The branched-chain acyl donor varies in length, shifting the balance between capsaicin and dihydrocapsaicin and with it the very *shape* of the burn. And, most corrosively for consistency, the enzyme itself misfolds — a reactive cysteine at position 175 scrambles disulfides, so any given fruit holds a mixture of correctly folded, fully active enzyme and a

misfolded, half-dead population, in a ratio that no one controls. A process built on a bimodal enzyme is a bimodal process, and bimodality is the fattest tail there is.

Our two designed substitutions each attack one source of that scatter, which is why they lower the mean's heat *and* collapse its variance at the same time. **S39L** tightens substrate binding in the N-terminal domain, so the active site is loaded more uniformly and the starved-versus-flooded swing narrows. **L345G** relieves steric strain in the active-site cleft, smoothing turnover so that correctly folded enzymes all work at closer to the same rate. The C175S substitution — which would abolish the disulfide-misfolding channel and convert a two-population enzyme into one — is a structural annotation from our modelling (ΔΔG −0.9 kcal/mol) that was identified but not included in the final shipped construct; its story explains the scatter source even though it did not make the design. Written in together with CRISPR, two edits the breeders never found by phenotype carry the process from CV 42 percent to the **≤7 percent** ceiling we ship against. The indices in this chapter are the accounting; those two residues are where the money is actually made.

From CV straight to Cpk

Here is the bridge that makes the whole book operational, and it is almost absurdly clean. Suppose the tolerance is written as a fraction of the target — the usual case, a window of ± $f \cdot \mu$, where f is the half-window expressed as a percentage of the aim. Then the half-window is $H = f \cdot \mu$, the standard deviation is $\sigma = \mathrm{CV} \cdot \mu$, and the μ cancels:

$$\mathrm{Cpk\ (centered)} = H / 3\sigma = (f \cdot \mu) / (3 \cdot \mathrm{CV} \cdot \mu) = f / (3 \cdot \mathrm{CV})$$

Capability, for a fractional window, does not depend on the absolute level at all — only on the ratio of the tolerance you are given to the relative scatter you deliver. This is the thesis of the whole book compressed into four symbols: hand the formulator a small CV and their chosen window turns from a worry into a wide-open margin, with no further work on their part.

Work it in our own units. A formulator sets a finished-product spec of IFI-U 2,600 ± 400 — a window of ±15.4 percent, generous but realistic for a savory depth at normal use. Then:

AXIS / INPUT	DELIVERED CV	Σ AT M = 2,600	CPK = F / (3·CV)	READING
Heat, our tightest axis	2.5%	65	**2.05**	six-sigma room; defects vanish
Umami, typical	4.0%	104	**1.28**	capable; comfortable margin
The ≤7% ceiling (worst input)	7.0%	182	**0.73**	marginal *as a single lot*

The ceiling line is the honest one, and we do not hide it. A lone sub-lot held at the very worst CV we permit, dropped into a ±15.4 percent window, is only Cpk 0.73 — sub-capable, a process that would fail lots on an ordinary drift. That is the worst case precisely because 7 percent is a *contractual ceiling*, not a typical value; most axes run far tighter, heat tightest of all. And there are two further, entirely honest routes that lift even the ceiling into the capable range.

The first is the window itself. The ±15.4 percent above is tight for a consumer product; most finished goods tolerate more. Capability against the ceiling CV of 7 percent reads:

PRODUCT WINDOW (±F)	CPK AT CV 7%	CPK AT CV 4%	CPK AT CV 2.5%
±10% (tight)	0.48	0.83	1.33
±15.4%	0.73	1.28	2.05
±20% (realistic)	0.95	1.67	3.20
±25% (generous)	1.19	2.08	3.33

Against the windows most products actually permit — ±20 percent and wider — even the 7 percent worst case clears a Cpk near one, and the tighter axes clear 1.33 and sail past it.

The second route is pooling, and it is pure arithmetic with no loss of truth. A finished batch rarely rests on one production unit; it blends several independent sub-lots. The mean of *n* independent units has a standard deviation smaller by a factor of $\sqrt{n}$, so the effective CV of the blend is CV / $\sqrt{n}$. The 7 percent ceiling, pooled across four sub-lots, becomes 3.5 percent and a Cpk of 1.47; across nine, it becomes 2.33 percent and a Cpk of 2.20. Blending is not a trick played on the spec; it is the plain statistics of averaging, and it is why a line that pools is categorically steadier than one that does not.

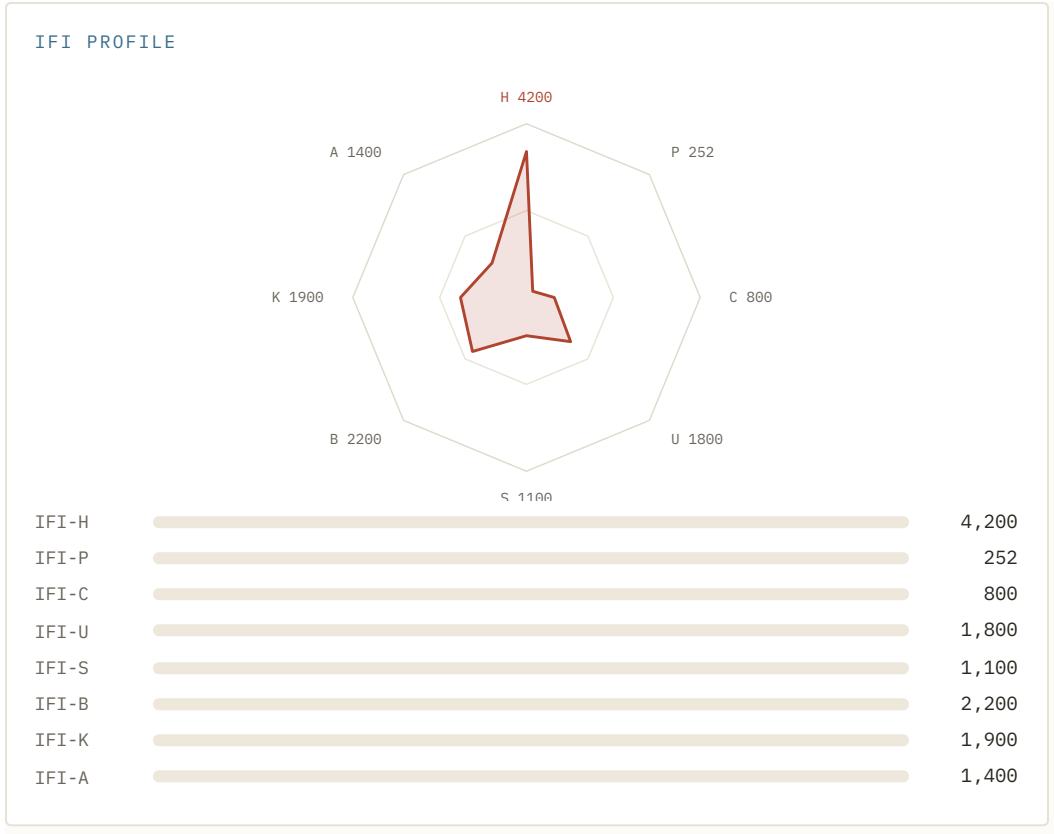
A ladder across the flagships

The scale-invariance of Cpk has a consequence worth pausing on, because it is counterintuitive and it is true: on a fractional window, the hottest pepper we designed is exactly as consistent as the mildest. Capability is a ratio, and a ratio does not care how big the number on top is. Take the flagship superhots, every one held at the heat axis's 2.5 percent CV, each against a ±15 percent SHU window:

CULTIVAR	SKU	PREDICTED SHU	COMPUTED IFI-H	Σ (SHU, CV 2.5%)	±15% HALF-WINDOW	CPK
Ember	SS-0087	866K	1,570	21,650	129,900	2.00
Cinder	SS-0088	1.05M	2,290	26,250	157,500	2.00
Blaze	SS-0091	1.72M	2,630	43,000	258,000	2.00
Firestorm	SS-0094	2.47M	3,490	61,750	370,500	2.00
Phoenix	SS-0096	2.95M	4,200	73,449	440,692	2.00

The σ column climbs — a hotter pepper has more absolute SHU to scatter — but the Cpk column is a flat, serene 2.00 all the way down. Phoenix, with our model predicting approximately 2.95 million Scoville heat units, the hottest cultivar our family has ever designed, is no harder to hold to spec than Ember at a third the heat. Consistency was never about being gentle. It is about keeping the fractional wander small, and a genome designed to keep it small keeps it small at every level.

Phoenix earns the ladder's last line for another reason: it is the one cultivar that our model activates on all eight axes at once, so it is the one where capability must be held eight ways simultaneously.



Each of those eight numbers carries its own window and its own Cpk, and the lot is releasable only when the *weakest* of the eight clears the bar — capability on a multi-axis product is a minimum across axes, never an average. The high axes have margin to spare; the quieter ones, cool-burn at 800 and sweetness at 1,100, run closest to their walls, because a smaller target held at the same fractional CV has a smaller absolute window to live in. Where we spend our tightening effort on Phoenix is decided entirely by which axis is nearest its edge — the arithmetic tells us where to care.

Reading the number from the receptor

Every one of those axis values is read, in the end, off a dose–response curve at a receptor, and the variance we have been accounting lives in the fit of that curve. A stimulus drives its receptor along a sigmoidal response described by a Hill relation — response rising with concentration as $\theta([c]) = [c]^h / (EC_{50}^h + [c]^h)$ — and EC_{50} is simply the concentration at which θ reaches one half, the point of half-maximal activation. Invert that half-activation against a fixed reference and scale by a thousand, and loudness becomes a score:

> **$IFI(\text{candidate}) = 1{,}000 \times EC_{50}(\text{reference}) \,/\, EC_{50}(\text{candidate})$**

The reference constant on each axis is a stated convention — defined, the way Scoville's dilution definition works, to place IFI = 1,000 at a chosen anchor point. Because each axis has its own reference, IFI scores on different axes are not directly comparable; they read against different defined anchors, not a single common scale.

A worked case shows why the index is the better axis to hold a spec against. A tomato, stripped by the industrial supply chain of the depth it should carry, reads IFI-U 300 — its umami compound reaches half-activation at 1000 / 300 = 3.33 times the reference concentration, a quiet thing, loud only at high dose. Grown back through the native glutamate route to IFI-U 2,600, its EC_{50} falls to 1000 / 2,600 = 0.385 times the reference. The ratio of those two potencies is 2,600 / 300 = **8.7**: the restored tomato speaks umami

nearly nine times more loudly, reaching the same half-activation at a ninth of the concentration. The IFI ratio *is* the potency ratio, because both share one reference — which is the quiet elegance of anchoring a flavor to a receptor constant instead of to a tasting panel.

And it is why capability on the IFI behaves. Because IFI is proportional to $1 / EC_{50}$ with the reference held fixed, the relative scatter of the index equals the relative scatter of the fitted EC_{50} almost exactly — $CV(IFI) \approx CV(EC_{50})$ — so the 7 percent ceiling is literally the reproducibility of the receptor assay, a biophysical measurement, not a human one. Scoville heat units, by contrast, come from a serial-dilution panel judged by trained tasters; the method carries a variance so large it is a fair fraction of the 42 percent heritage the wild plant hands us, and it diverges from IFI-H nonlinearly as the heat climbs. We report SHU because the world asks for it. We hold our spec on IFI, because you cannot build a capable process on a ruler whose own graduations wander.

Defects, counted honestly in parts per million

The reason Cpk is worth all this arithmetic is that it converts cleanly into the only currency a manufacturer truly cares about — the fraction of lots that will fail. For a centered normal process, the nearer limit sits $3 \cdot Cpk$ standard deviations from the mean, so the one-sided defect fraction is the normal tail beyond that point, $ppm = 10^6 \times \Phi(-3 \cdot Cpk)$, with Φ the standard normal distribution function:

CPK	LIMIT AT (NEARER)	ONE-SIDED DEFECT RATE	READING
1.00	3σ	~ 1,350 ppm	marginal; a drift will fail lots
1.33	4σ	~ 32 ppm	the standard target for release
1.67	5σ	~ 0.3 ppm	strong; robust to ordinary drift
2.00	6σ	~ 0.001 ppm	the "six-sigma" ideal

The table is the bridge from our world to a buyer's. The 7 percent CV ceiling is the *worst* input we will ship; against the generous windows of most consumer products it already clears Cpk 1.33 and beyond, and the tight axes — heat at two and a half percent, umami at four — hand the formulator Cpk figures deep in the capable range. Which is to say: a defect rate that rounds to zero, lot after lot, for the quiet reason that what the body receives from one jar is what it receives from the next.

The shift nobody escapes

There is a sobering correction built into the famous sigma language, and we keep it in view rather than letting the pretty numbers flatter us. People speak of a "six-sigma" process as one running 3.4 defects per million, and that figure is *not* the centered 0.001 ppm the table above gives for Cpk 2.00. The difference is a concession manufacturing learned the hard way: processes do not hold their center forever. Over long runs the mean wanders — tool wear, a slow reagent drift, the seasons of a supply chain, the heat wave

through the greenhouse — and the convention is to budget for a long-term shift of about 1.5σ off target. Re-run the tail arithmetic with the center shifted a sigma and a half toward the wall:

SHORT-TERM SIGMA LEVEL	CPK (CENTERED)	LONG-TERM PPM (WITH 1.5Σ SHIFT)
3 σ	1.00	~ 66,800
4 σ	1.33	~ 6,210
5 σ	1.67	~ 233
6 σ	2.00	~ 3.4

A process that looks like six sigma of room when perfectly centered delivers, after that assumed shift, the 3.4 ppm that gives the method its name; a nominal three-sigma process, shifted, delivers not a handful of parts per million but tens of thousands. The lesson for a flavor spec arrives early and plainly: short-term capability is necessary but not sufficient. A Cpk measured on a single good day describes only that day. You must also keep the center from wandering across months, which is why capability and the ongoing discipline of keeping a process in statistical control are two halves of one craft, and neither stands alone. A high Cpk on an unstable process is not a capable process; it is a lucky snapshot of one.

The measurement that measures the process

A capability index is only as trustworthy as the instrument that produced the numbers feeding it, and this is the quiet place most capability claims go wrong. Every observed scatter is the convolution of the real process variation with the noise of the assay reading it. For independent sources, variances add, so in relative terms:

> **CV(observed)² = CV(process)² + CV(assay)²**



The sequencer and assay bench where EC_{50} is read. A capability number inherits the honesty of the instrument that produced it; the measurement's own variance must be a small fraction of the tolerance it is asked to judge.

Suppose the dose–response assay carries a CV of its own of 2 percent, and a batch reads at an observed CV of 7 percent. The true process variation is not 7 but $\sqrt{(7^2 - 2^2)} = \sqrt{45} \approx$ **6.7 percent** — the assay has inflated the apparent scatter by three-tenths of a point. Let the assay degrade to 3 percent and it inflates a true 6.3 percent reading up to the same observed 7. The rule of thumb that follows is strict and we hold to it: the measurement system should consume

only a small fraction of the tolerance it is asked to police. An assay whose own spread is a large share of the window cannot resolve a Cpk of 1.33 any more than a ruler marked in inches can certify a part to the thousandth; it will reject good lots and pass bad ones with equal confidence, and no amount of downstream statistics repairs a blurred reading at the source.

One more measurement truth keeps us modest: a capability index is itself an estimate, computed from a finite sample, and it carries its own uncertainty. The standard error of an estimated standard deviation from n measurements runs about $\sigma / \sqrt{(2(n - 1))}$. From a mere five samples, that is a 35 percent relative uncertainty on σ — a Cpk "of 1.5" that might honestly be anywhere from 1.1 to 2.2. Only by twenty-five or thirty subgroups does the uncertainty fall near 13 percent and the number start to mean what it says. We do not quote a capability off a handful of fruit, and we are wary of anyone who does.

Cpm — the index that rewards aim

Cpk asks only whether you are inside the window; it is indifferent to where inside. A lot hugging the upper wall and a lot resting dead on target can score the identical Cpk, and to Cpk they are equally good. To a mouth they are not. This is why we watch a quieter index internally — the Taguchi capability, **Cpm** — which folds the distance from the exact target into its denominator:

```
Cpm = (USL − LSL) / 6 · √(σ² + (μ − T)²)
```

The term $(\mu - T)^2$ adds the squared miss from target T to the variance before taking the root, so a process that is on-spec but off-target is penalized for the offset, smoothly, continuously, before any limit is ever crossed. This matches how taste actually works and how we think about care. The loss a taster feels does not sit at zero across the whole window and then leap to one at the wall; it grows gently with every point of drift from the aim, the way a broth a little too salty is already a little less itself long before anyone would send it back. A spec met at its exact center is a kinder thing than the same spec met at its boundary, and Cpm is the only one of these indices that agrees.

Honest limits

Four cautions keep this chapter from overpromising. First, every ppm figure here assumes a normal distribution, and flavor data is not always normal. The very misfold tail that the C175S substitution — a structural annotation our modelling identified, though not included in the shipped design — was intended to address is the kind of skew that makes Cpk *understate* the true defect rate — a reminder that the index is a model, and that we check the actual shape of a distribution before we trust its tails. Second, the fractional-window math assumes the tolerance scales with the target; some specs are written in absolute units — ± a fixed number of SHU regardless of level — and there the capability of a high-heat cultivar and a low-heat one genuinely diverge, and the comforting scale-invariance of the ladder does not hold. Third, an index computed on an out-of-control process is meaningless, because an unstable process has no

single σ to put in the denominator; capability presumes control and never substitutes for it. And fourth, as the sampling arithmetic showed, a capability quoted from too few lots is a guess dressed as a measurement.

Held inside those limits, the promise is exact and it is the one promise a product at scale can actually be built on: a defect rate that rounds to zero, lot after lot. Not an average that happens to land right, but a window met at its center almost every time — so that the depth we designed into a fruit is the depth that reaches the table, this winter and the next, in a jar no one inspected by hand and no one needed to.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

Chapter 13 · The control chart — *statistical process control, and the two kinds of cause.*

Capability tells you whether a process *can* fit its window. It does not tell you whether the process is *staying* where it should, run after run, or quietly wandering off while no one is watching the right number. For that there is the oldest and most beautiful instrument in all of quality — Walter Shewhart's control chart, first drawn on a single sheet of paper at Western Electric in 1924 — and the one idea beneath it is the idea every maker should carry in the hand like a tuning fork: that variation comes in two kinds, and the whole discipline of keeping a living process steady begins with learning to tell them apart.

We came to this the long way, the way a family comes to anything it ends up loving. Before we could design a flavor we had to learn to *trust* a number — to grow a fruit one season and have the next season's fruit mean the same thing on the tongue. A cultivar our model predicts will reach 2,937,948 Scoville Heat Units once is a curiosity; a cultivar designed to reach it again, and again, inside a band you can draw before the harvest, is a tool someone can formulate around. The control chart is how that trust is kept honest. It is a picture of a process telling the truth about itself.



The line, watched — a lot is kept in control not by inspecting the exit but by reading the trend early, while a drift is still only a drift.

Common cause and special cause — the whole discipline in one distinction

The first kind of variation is *common cause* — the ordinary, irreducible background scatter of a process working exactly as designed. Common-cause variation is the gentle wander we spend this whole book measuring and shrinking; it is the system being itself, the sum of a thousand tiny influences too small and too many to name: a photon here, a stoma there, a plate read a half-second

early. The second kind is *special cause* — a genuine disturbance, something that was not there before: a reagent gone off, a light bank failing at one end of the bench, a feed line drifting, a new hand reading the assay a little differently than the hand that trained on it.

The two demand opposite responses, and confusing them is how good processes are ruined. Chase a common-cause wiggle as though it were a problem — *the last lot ran two points high, adjust everything* — and you *add* variance, over-correcting a system that was already fine. This is the classic error, the one W. Edwards Deming drew with a funnel and a target and called tampering: every "correction" to pure noise injects a fresh error on top of the one you imagined you were fixing, and the spread grows instead of shrinking. Miss a special cause the other way — wave off a genuinely broken lot as ordinary noise — and you ship the one batch that was actually wrong. The control chart's entire job is to draw, in advance and without sentiment, the line between *leave it alone* and *stop the line.*

Where ±3σ comes from — limits that are earned, not chosen

The chart does this by plotting each lot's measured score, in time order, against a centerline at the process mean, with control limits set at **plus and minus three standard deviations** — the band inside which common cause alone will keep nearly every point, purely by chance. The three is not a round number picked for comfort; it falls out of the normal distribution. For a centered, stable process, the probability of a single point landing beyond 3σ on either side is

```
P(|Z| > 3) = 2 × (1 − Φ(3)) = 2 × 0.00135 = 0.0027.
```

One point in roughly 370 — because 1 ÷ 0.0027 = 370.4 — will cross a limit even when nothing at all has gone wrong. That is the cost of a false alarm, and Shewhart chose it deliberately. Draw the limits at ±2σ instead and the false-alarm rate climbs to one lot in twenty-two; the chart would cry wolf so often no formulator would trust it by Thursday. Draw them at ±3σ and a crossing is rare enough to be worth standing up for. A point inside the limits is the process being itself — leave it. A point outside is a signal — special cause, almost certainly, because common cause produces such a point less than three times in a thousand. The limits are not an opinion about where points *should* fall. They are a computation of where they *will* fall when the process is behaving, and everything outside is news.

Reading the index, not the raw heat

A heat number can be written two ways, and a control chart is only as trustworthy as the number it watches. The Scoville scale measures *how much* — the concentration of capsaicinoids a fruit carries. The Iyengar Flavor Index measures *how loudly* — how potent that chemistry is at the receptor it speaks to. For heat the receptor is TRPV1, a non-selective cation channel on the trigeminal endings of the tongue, the same channel opened by temperatures past 43 °C and by acid. Capsaicin does not warm the tissue; it lowers the channel's thermal threshold, so the body reports a fire

that is not there. The Index asks the only question that matters to the eater — *at what concentration does this molecule drive TRPV1 to half of its full response?* — and turns the answer into a score:

> $\mathrm{IFI(candidate)} = 1{,}000 \times \mathrm{EC_{50}(reference)} \div \mathrm{EC_{50}(candidate)}.$

EC_{50} is the half-activating concentration. Because it sits in the denominator, a *lower* EC_{50} — a molecule that opens the channel at a whisper — yields a *higher* index. The reference is anchored so that IFI = 1,000 at the reference compound — a stated convention, defined at its anchor the way Scoville's dilution count defined 'heat', not derived from an average of published studies: a candidate that scores 1,000 is exactly as loud as the reference, and a candidate at 2,000 is twice as loud, reaching half-maximal activation at half the concentration.

The arithmetic is worth doing by hand, because it is cleaner than it looks and it tells you why we chart the index rather than the raw assay. Our flagship, Phoenix (SS-0096), scores **IFI-H 4,200**. Invert the formula and its effective EC_{50} is

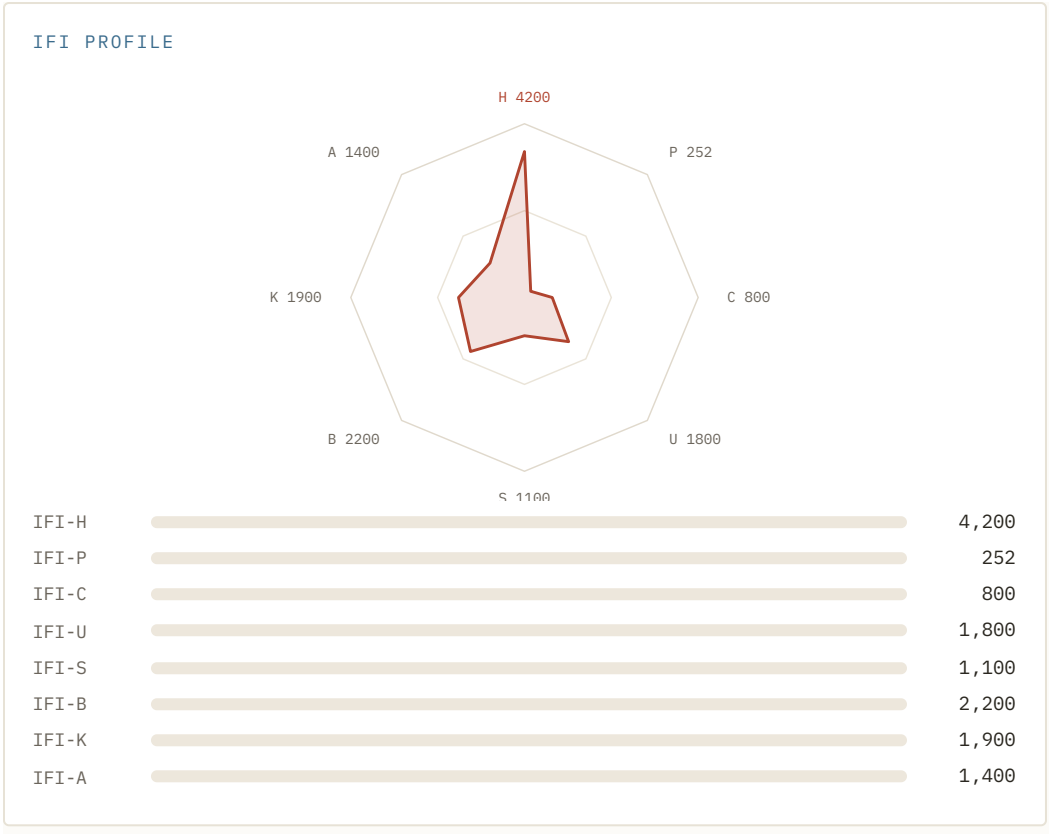
> $\mathrm{EC_{50}(Phoenix)} = \mathrm{EC_{50}(ref)} \times 1{,}000 \div 4{,}200 = 0.238 \times \mathrm{EC_{50}(ref)}.$

In our model, Phoenix's design reaches half-activation at 23.8 % of the reference concentration — about 4.2× louder than the native plant. Compare two cultivars and the anchor falls away entirely, which is the quiet beauty of a ratio scale:

> $\mathrm{EC_{50}(Ember)} \div \mathrm{EC_{50}(Phoenix)} = \mathrm{IFI(Phoenix)} \div \mathrm{IFI(Ember)} = 4{,}200 \div 1{,}570 = 2.68.$

Ember (SS-0087, IFI-H 1,570) needs 2.68× the concentration Phoenix needs to say the same thing to the nerve. The literature places capsaicin's TRPV1 EC_{50} near 0.7 µM; take that as the reference anchor and Phoenix lands near 0.167 µM, Ember near 0.446 µM — but notice that the *anchor cancels* out of every control calculation that follows. A chart of the index, or of its logarithm, is invariant to the reference we choose. That invariance is exactly why the Index, not the micromolar reading, is the thing we put on the chart: it is stable against the one number most likely to drift between labs.

Phoenix is the only cultivar in our catalog that registers on all eight axes at once, and its full profile is the standard we hold a lot against:



Each axis carries its own stated reference constant — IFI = 1,000 anchors the reference compound on that axis independently of every other axis. Scores chart consistency within an axis; they are not values on a single shared scale across axes.

From Scoville's tasters to the receptor — two ways to measure a fire

It is humbling to remember that the Scoville scale began, in 1912, as a panel of human tasters. Wilbur Scoville dissolved a measured extract in sugar-water and diluted it until a panel of five could no longer feel the heat; the dilution at the vanishing point *was* the score. It was an assay with a magnificent flaw — the instrument was a tongue, and tongues disagree. A five-person panel carries a coefficient of variation that can run past thirty or forty percent between sittings, which means the very first thing anyone ever measured about heat was measured badly, and the whole subsequent history of the field is the history of driving that measurement's variation down: from the panel to high-performance liquid chromatography, which quantifies capsaicinoids directly at roughly 16 Scoville units per part-per-million, and from the chromatograph to the receptor assay that reads potency itself. Pure crystalline capsaicin sits at the ceiling of the scale, near 16,000,000 SHU. By that ceiling, Phoenix's predicted 2,937,948 SHU is

> `2,937,948 ÷ 16,000,000 = 0.1836 — 18.36 % of the pungency of pure capsaicin,`

designed to be carried inside a fruit grown in a living greenhouse rather than crystallized in a flask.

Here is where the two numbers quietly part ways, and where a formulator earns their keep. Scoville and the Index are correlated across our cultivar design series but they are *not* the same

measurement, because the Scoville reading counts total capsaicinoid mass while the Index weighs the *shape* of that mixture at the channel.

CULTIVAR · SKU	PREDICTED SHU	IFI-H	EC_{50} ÷ REF	PREDICTED SHU ÷ IFI-H	AXES
Ember · SS-0087	866,000	1,570	0.637	552	2/8
Cinder · SS-0088	1,050,000	2,290	0.437	459	3/8
Kindle · SS-0089	1,250,000	2,000	0.500	625	3/8
Flare · SS-0090	1,480,000	—	—	—	4/8
Blaze · SS-0091	1,720,000	2,630	0.380	654	4/8
Wildfire · SS-0092	1,960,000	2,930	0.341	669	5/8
Scorch · SS-0093	2,210,000	3,220	0.311	686	5/8
Firestorm · SS-0094	2,470,000	3,490	0.287	708	6/8
Solara · SS-0095	2,720,000	3,980	0.251	683	6/8
Phoenix · SS-0096	2,937,948	4,200	0.238	700	8/8

Read the fifth column slowly. If Scoville and the Index measured the same thing, the ratio of SHU to IFI-H would be constant down the ladder. It is not: it runs from 459 at Cinder to 708 at Firestorm, and the ladder even inverts — Kindle carries *more* raw heat than Cinder (1,250,000 against 1,050,000 SHU) yet reads *quieter* at the receptor (IFI-H 2,000 against 2,290). The reason is molecular. Capsaicin is assembled when vanillylamine, built through the phenylpropanoid branch and the pAMT aminotransferase, condenses with a branched-chain fatty acid, built through BCAT and KAS, at the Pun1/CS acyltransferase. The length and saturation of that acyl chain set the balance between capsaicin and its slightly milder cousin dihydrocapsaicin — and that balance shifts potency at TRPV1 *independent of total mass*. Two lots can carry equal capsaicinoids and speak at different volumes. A Scoville reading would call them identical; the Index hears the difference. This is precisely why the number we control is the receptor number, and why consistency is a molecular problem before it is a statistical one.

The number this book is about — CV, and the thirty-six-fold quiet

Everything we have built toward lives in one statistic: the coefficient of variation, the standard deviation divided by the mean, the spread expressed as a fraction of the thing itself.

> `CV = σ ÷ μ × 100 %.`

A wild Capsicum left to its own genetics runs a batch-to-batch CV near **42 %** on capsaicinoid output — a plant whose fruit might be searing or merely warm depending on a summer we did not control. The two AI-designed substitutions we wrote into Pun1 with CRISPR — **S39L**, which lifts N-terminal binding affinity, and **L345G**, which relieves steric strain in the active-site cleft — do more than raise peak output. By stabilizing the enzyme and its product ratio, they pull the batch CV down to **≤ 7 %**. That reduction in the *ratio* is a factor of six (42 ÷ 7 = 6). But variation does not add — variance does, and variance scales as the square of CV, so the honest figure, the one we are proudest of, is

```
(42 ÷ 7)² = 36.
```

The engineered process carries roughly **one thirty-sixth** the variance of the wild plant. That is the whole book in a single number: not a louder fruit, a *quieter* one.

A CV is a promise you can draw. At Phoenix's predicted mean of μ = 2,937,948 SHU and a CV of 7 %, the standard deviation is σ = 0.07 × 2,937,948 = 205,656 SHU. The three-sigma band a single-lot chart would place around the centerline is therefore μ(1 ± 0.21), or

```
2,320,979 to 3,554,917 SHU.
```

Any lot inside that band is Phoenix being itself; any lot outside is a lot with a story. Now do the same arithmetic for the wild plant at 42 %: μ ± 3σ becomes μ(1 ± 1.26), and the lower limit goes *negative*. A process that broad cannot even be charted with symmetric limits — the normal approximation breaks, the distribution has to be modeled as log-normal, and more to the point a fruit that variable has no window it can promise anyone. The six-fold tightening is not cosmetic. It is the difference between a number you can formulate against and a number you can only hope for.

Because a profile is eight numbers, the control band is eight bands. Holding Phoenix's canonical profile at a 7 % CV — σ = 0.07 × target, and ±3σ = ±0.21 × target on every axis — gives the chart a formulator actually works from:

AXIS	TARGET IFI	Σ (7 %)	LCL (−3Σ)	UCL (+3Σ)
H · Heat	4,200	294	3,318	5,082
P · Pungency	252	18	198	306
C · Cool-Burn	800	56	632	968
U · Umami	1,800	126	1,422	2,178
S · Sweetness	1,100	77	869	1,331
B · Zero-Bitter	2,200	154	1,738	2,662
K · Kokumi	1,900	133	1,501	2,299
A · Astringency	1,400	98	1,106	1,694

The heat axis is worth a glance: its upper control limit, 5,082, actually overruns the 5,000 ceiling to which the scorecards are normalized — Phoenix runs so near the top of the heat scale that its ordinary good-day scatter touches the edge of the page. That is not a fault to correct. It is the signature of a cultivar built to live at the frontier of its own axis.



The chamber that makes a mean hold still — every variable we can fix, fixed, so the only variation left to chart is the process being honestly itself.

Two charts, one process — the $\bar{X}$–R pair and the individuals chart

Which chart to run is itself a small act of judgment, and it turns on how a lot arrives. When a lot yields a *group* of measurements — several fruit, several plate reads — the classic instrument is the $\bar{X}$ and R pair: plot the average of each subgroup on one chart to watch the center, and the range within each subgroup on a second to watch the spread, because a process can hold its average while its scatter quietly widens, and you want to catch either failure before it reaches the eater.

The limits here are not the raw ±3σ of the individuals chart, because averaging settles a number down. The standard error of a subgroup mean is $\sigma \div \sqrt{n}$, so with subgroups of five fruit the center chart's limits tighten by $\sqrt{5} = 2.236$. The classical constants do this arithmetic for you, and it is satisfying to watch them agree with first principles. For $n = 5$ the tabulated values are $d_2 = 2.326$, $A_2 = 0.577$, $D_3 = 0$, $D_4 = 2.114$. Take Phoenix's $\sigma = 205{,}656$ SHU. The expected mean within-lot range is

> $\bar{R} = d_2 \times \sigma = 2.326 \times 205{,}656 = 478{,}355$ SHU,

and the center-chart limits are $\bar{\bar{X}} \pm A_2\bar{R} = 2{,}937{,}948 \pm (0.577 \times 478{,}355) = 2{,}937{,}948 \pm 276{,}011$, giving

> UCL = 3,213,959 SHU and LCL = 2,661,937 SHU.

Now check it against the standard error directly: $3 \times \sigma \div \sqrt{n} = 3 \times 205{,}656 \div 2.236 = 275{,}910$ — the same 276,000 the constants produced, to rounding, because A_2 is nothing more than $3 \div (d_2\sqrt{n}) = 3 \div (2.326 \times 2.236) = 0.577$. The spread chart follows the same way: $UCL = D_4\bar{R} = 2.114 \times 478{,}355 = 1{,}011{,}243$ SHU, and $LCL = D_3\bar{R} = 0$, since with only five fruit a range cannot be meaningfully *too small*. A lot of five whose average strays outside [2.66M, 3.21M] SHU, or whose internal range exceeds 1.01M SHU, is a signal worth stopping for — and the two charts catch different sins: the center chart a shift, the range chart a loss of uniformity among the fruit of a single harvest.

When a lot yields only one number — a single pooled extract, a single read — there are no subgroups to average, and the tool becomes the individuals-and-moving-range chart, which builds its sense of ordinary variation from the lot-to-lot jumps themselves rather than from within-lot spread. The limits are still computed from the process's own short-term scatter, never guessed. And because the Index and the EC_{50} behind it are multiplicative quantities — a constant CV means a constant spread *in the logarithm*, not in the raw value — the cleanest charts in our lab are run on ln(IFI). On the log scale the 7 % band becomes very nearly symmetric, the negative-limit pathology of a wide process disappears, and a proportional drift reads as a straight line. It is a small refinement that pays for itself the first time a process creeps by a fixed percentage rather than a fixed amount.

The run rules — turning a lagging alarm into a leading one

A single point beyond 3σ is a lagging alarm: by the time a lot crosses the limit, the drift that carried it there has already arrived. The real gift of the chart to a formulator is the family of *run rules* — the Western Electric patterns — that catch a process while it is still only *leaning*, too orderly to be chance but not yet out of spec. Each is a shape in the data with a probability attached, and the probabilities are the point: these rules are tuned, not improvised.

RUN RULE (WESTERN ELECTRIC)	WHAT IT CATCHES	≈ FALSE-ALARM RATE
1 point beyond 3σ	a sudden special cause — a dead light bank, a bad reagent	1 in 370
2 of 3 points beyond 2σ, same side	an emerging shift, caught early	1 in 326
4 of 5 points beyond 1σ, same side	a slow drift gathering	1 in 180
8 points in a row on one side of center	a quiet, sustained move off target	1 in 128
6 points in a row trending up or down	a progressive cause — a standard aging, a feed creeping	—

Each rate is a straightforward computation on the normal curve, and doing one shows the whole family is calibrated rather than conjured. Take the eight-in-a-row rule. For a centered process each point is equally likely to fall above or below the line, so eight consecutive points on one specified side has probability $(½)^8 = 1 ÷ 256$, and on *either* side $2 ÷ 256 = 1$ in 128. The two-of-three-beyond-2σ rule works the same way from $p = 0.0228$, the single-tail area past 2σ: the chance of at least two of three points clearing 2σ on one side is $3p^2(1 - p) + p^3 \approx 0.00154$, and doubled for either side about 1 in 326. Notice these sit near the 1-in-370 of the master rule — close enough to keep the combined chart from crying wolf, loose enough to raise the alarm a run or two *before* a point ever reaches a limit. That earliness is the whole value. It is the difference between catching a light bank as it begins to fade and discovering it after a week of dim

lots has already shipped. The chart lets a grower see a season going wrong while it is still only *going*, and intervene in the room rather than in the field.

Why a steady number is a steady taste — the dose–response curve

It would be enough if consistency were only good manufacturing. It is better than that: near the half-activation point, a steady *number* is the only thing that buys you a steady *sensation*, and the dose–response curve shows why. Receptor activation follows a sigmoid — the Hill equation — in the concentration C of the stimulus:

```
f = Cʰ ÷ (EC₅₀ʰ + Cʰ),
```

where f is the fraction of full response and h is the Hill coefficient, near 2 for capsaicin at TRPV1, reflecting the channel's cooperative gating. At C = EC_{50} the curve passes through f = ½ by definition — the index's anchor is literally the middle of this S. And the middle is the steepest, most treacherous place to stand. The local sensitivity of response to a proportional change in dose is

```
df ÷ d(ln C) = h · f · (1 − f),
```

which reaches its maximum exactly at f = ½, where it equals h × ¼ = 0.5 for h ≈ 2. In plain terms: dosing a formulation near half-max — where a great many real products sit, because it is where flavor reads as "present but not punishing" — a fractional swing in concentration propagates to a response swing of about half its size. A ±7 % spread in heat becomes roughly a ±3.5 % swing in perceived intensity. The wild plant's ±42 % spread becomes a ±21 % swing — and worse than that figure suggests, because the curve bends, so the high excursions and the low ones are not felt symmetrically. One box of the same sauce would scorch and the next would disappoint, and the eater would blame the cook. Pull the CV down thirty-six-fold in variance and you are not only tightening a spec sheet; you are holding the sensation itself still on the tongue, meal after meal, which is the only promise a flavor can actually keep.



The read is a measurement, and a measurement has a variance of its own — which is why a fixed standard rides every plate, to tell a drifting plant from a drifting assay.

What the chart cannot do — measurement, capability, and honest limits

A chart is a mirror, and a mirror can only be as true as the glass. The first honest limit is that a control chart shows you the *measured* process, which is the real process convolved with the measurement itself. Variances add:

```
σ²(observed) = σ²(process) + σ²(assay).
```

If the chart reads a 7 % CV and a reference standard run alongside the lots shows the plate assay carries 4 % on its own, then the plant is quieter than the chart admits — the process CV is $\sqrt{(7^2 - 4^2)} = \sqrt{33} = 5.7$ %, and the remaining variation lives in the read, not the fruit. This is not pedantry; it is a failure mode with a name. Chase an "out of control" point that is really the assay drifting and you will tamper with a plant that was fine, hunting a ghost in the greenhouse while it stands in the spectrophotometer. The defense is old and simple: ride a fixed control sample — a frozen reference extract, a capsaicin standard — on every plate, chart *it* too, and let it tell a drifting plant from a drifting instrument. It is the modern answer to Scoville's five tasters: the only way to trust a measurement of variation is to measure the measurement.

The second honest limit is the sentence that separates a mature quality program from a hopeful one: **in control is not the same as in spec.** A process can sit flawlessly inside its control limits and still be unable to meet a customer's window, because its ordinary spread is simply too wide — that is a *capability* problem, not a control problem, and no amount of charting fixes it. Capability has its own arithmetic. For a spec window from LSL to USL,

```
Cp = (USL − LSL) ÷ 6σ, and Cpk = min[(USL − μ), (μ − LSL)] ÷ 3σ,
```

the first measuring whether the spread *fits* the window, the second whether the process is *centered* in it. The relationship to CV is exact and clarifying. If a customer specifies Phoenix's grade at ±21 % — which is precisely ±3σ at a 7 % CV — then Cp = (0.42μ) ÷ (6σ) = (0.42 × 2,937,948) ÷ (6 × 205,656) = 1.00, a process that *just* fills its window. To reach the industry's comfortable Cp = 1.33 you must either widen the window to ±28 % (±4σ) or tighten the CV to 5.3 %. And centering matters as much as spread: let that same process drift 5 % off target inside the ±21 % window and Cpk falls to (1.21 − 1.05) ÷ 0.21 = 0.76 — the spread never changed, but the lot is now at risk, because it leans. This is the exact handshake between the two tools. Capability tells you the window *can* be met; control tells you it *is being* met, lot after lot; you need both, and a chart that is beautifully in control around the wrong center is a tidy way to ship failures.

Where, then, is capability won, if not at the chart? Upstream — at the four doors through which variation enters a living process: the **seed**, where genetics sets the floor on spread; the **grow**, where light and water and warmth either hold a mean still or let it wander; the **extraction**, where sample preparation adds or removes noise; and the **read**, where the assay contributes the variance we just learned to subtract. Our thirty-six-fold quiet was won at the first door, in the enzyme, because S39L and L345G narrowed the distribution of what

the plant *makes* before a single lot was ever charted. The chart did not create that consistency. It is how we prove, harvest after harvest, that the consistency is still there.

A last caution, offered plainly because numbers are a form of respect: the standard chart assumes each lot is an independent draw. A biological process measured often — the same chamber, the same week, the same slowly warming afternoon — can be autocorrelated, each lot leaning the way the last one leaned, and autocorrelation inflates false alarms and can hide a real trend inside what looks like ordinary wander. The honest moves are to sample on a timescale long enough that lots are genuinely independent, or to chart the residuals of a fitted trend rather than the raw series. A chart you trust without understanding its assumptions is just a decorated hope.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

In control is not in spec — the sentence to carry

Carry one sentence out of this chapter, and let it be the one that most separates a quality program that *prevents* from one that merely *regrets*: a process can sit flawlessly inside its control limits and still be incapable of meeting its window, and a process can be richly capable yet drift out of control the instant a special cause arrives. You need both — capability to know the window can be met, control to know it is being met, run after run.

We hold to this for a reason that is finally not statistical at all. When our family grows a fruit meant to nourish the body — grown clean, grown to carry the fullest flavor and the fullest goodness it can hold — a number that cannot be trusted is a kind of carelessness we are not willing to pass to someone's table. A control chart is, at bottom, an instrument of care: it is how we keep a promise to the eater we will never meet, that the warmth arriving on the tongue tonight is the same warmth we meant, the same warmth they can reach for again tomorrow and have it be true. The thirty-six-fold quiet we engineered into the enzyme is only a quiet worth having if we can show, lot after lot and season after season, that it stayed. The chart is how we show it. It is the mathematics of keeping faith.

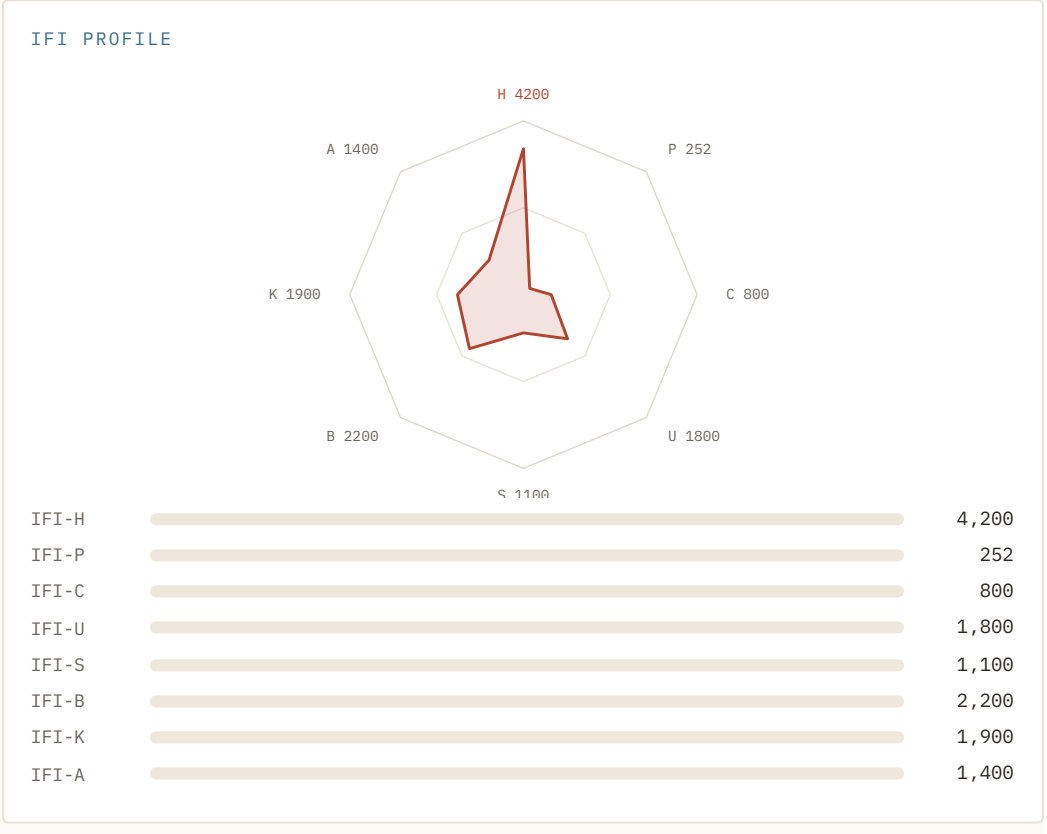
Chapter 14 · Lot release — *acceptance sampling, and the two risks.*

At the end of the line stands a plain and ancient question, and no amount of upstream grace excuses answering it carelessly: does *this* lot ship? We cannot taste every fruit — a destructive assay on the whole harvest would leave nothing to carry to anyone's table — so we measure a sample and decide the fate of the whole from the part. That is acceptance sampling, and we have come to love its mathematics the way we love a well-set table: it is the place where a season of care is finally, and honestly, staked.

The temptation is to treat the gate as where quality happens. It is not. By the time a tray of fruit reaches release, its character was settled seasons ago — in the edited gene, in the controlled air, in the patient line that held its conditions steady. What the gate does is narrower and more sacred: it confirms, with stated and bounded risk, that the quality we built actually arrived in *this* particular lot. To do that honestly we have to say out loud two things most programs leave unspoken — how often we will wrongly turn away a good harvest, and how often we will wrongly wave a bad one through. Those are the two risks. This chapter is about holding them both in the open, and about the arithmetic that lets a sample of a few dozen fruit speak fairly for a whole harvest.

A lot is a promise, and the promise is a number

Before we can sample, we have to be able to say what it would mean for a fruit to fall *out* of spec — and in our work, a specification is not a vague aspiration but a number with a window around it. Take our flagship, Phoenix, designed to register on all eight axes of the Index at once. Its heat specification is IFI-H **4,200**, and the whole of its design reads as a scorecard we hold every lot against:



Read the formula slowly, because every number in this chapter descends from it:

> $\mathrm{IFI}(\mathrm{candidate}) = 1000 \cdot \mathrm{EC}_{50}(\mathrm{reference}) \,/\, \mathrm{EC}_{50}(\mathrm{candidate})$

EC_{50} is the concentration at which a stimulus drives its receptor to half of its maximum response — the midpoint of the dose–response curve, and the most honest single summary of potency we know. The Index inverts it against a fixed reference constant and scales by a thousand, so that a larger score means a *louder* molecule, not merely a more abundant one. At the anchor, where a candidate half-activates at exactly the reference concentration, the Index reads 1,000. Phoenix's design puts heat at 4,200, and that number is not decoration — it is an inverted concentration. Solve for the candidate: $EC_{50}(\mathrm{Phoenix}) = EC_{50}(\mathrm{ref}) \times 1000 / 4200 = EC_{50}(\mathrm{ref}) / 4.2$. A conforming Phoenix fruit half-opens the channel at a concentration **4.2 times lower** than the wild-type reference. It is, precisely and literally, 4.2 times louder. Put a number on the anchor for the sake of arithmetic — take the reference half-activation near 0.50 µM, well inside the sub-micromolar range in which capsaicin speaks to TRPV1 — and a conforming fruit half-activates at $0.50 / 4.2 \approx$ **0.119 µM**. A sunlit shishito, reading about 40 on the same axis, needs $0.50 \times 1000 / 40 = 12.5$ µM to say the same thing to the same channel: a molecule roughly a hundred times quieter, grown toward warmth rather than fire.

So "out of spec" has an exact meaning. We hold a Phoenix lot's heat to a release window of **4,200 ± 600** on the H axis — a floor at 3,600, a ceiling at 4,800 that leaves honest headroom beneath the per-axis maximum of 5,000 — and a fruit whose IFI-H falls outside that window is, for the purpose of this lot, a defect. Everything that follows is the mathematics of deciding the fate of the window from a handful of fruit.



A single controlled-environment run arrives for release — the tray is the lot, and the lot is the unit the whole arithmetic speaks for.

The receptor arithmetic under a single fruit's score

To trust the window we have to trust the measurement behind it, and that measurement reaches all the way down to a channel in a membrane. Capsaicin binds TRPV1, a non-selective cation channel on the trigeminal nerve endings of the tongue — the same receptor

gated by temperatures above 43 °C and by protons. The molecule warms nothing; it lowers the channel's thermal threshold, so the body reports a fire that is not there. A receptor does not read concentration linearly, and this matters for every number we print. The fraction of channels standing open follows a Hill relation,

```
R / R_max = [L]^h / (EC50^h + [L]^h)
```

where [L] is the ligand concentration at the nerve ending and h is the Hill coefficient — the cooperativity of binding, steep (near 2) for capsaicin at TRPV1. Set [L] = EC_{50} and the right-hand side collapses to EC_{50}^h / (EC_{50}^h + EC_{50}^h) = 1/2, for *any* h at all. That algebraic inevitability is the reason we anchor the Index to the half-maximum and nowhere else: it is the one point on the dose–response curve that does not depend on how cooperative the binding is — the one landmark two different cell lines, two different seasons, two different hands can agree on. We read it on a cell-based TRPV1 assay, a dish of receptor rather than a living animal, which suits both the rigor and the ethos: potency read without harm.

The cooperativity is not an idle detail; it is the reason tight specifications matter. Writing x = [L] / EC_{50}, the Hill relation becomes R / R_max = x^h / (1 + x^h), and at h = 2 a mere four-fold swing in effective concentration — from x = 0.5 to x = 2 — carries the response from 0.5^2 / (1 + 0.5^2) = 0.20 all the way to 2^2 / (1 + 2^2) = 0.80. A channel that steep turns a small drift in a fruit's potency into a large swing in the heat a person feels, which is precisely why we will not let the window breathe and why the coefficient of variation, not the mean alone, is what we police.

Where the fruit's loudness comes from is itself an enzymatic story. Capsaicin is assembled at the Pun1/AT3 acyltransferase, which condenses vanillylamine with an 8-methyl branched-chain fatty acyl group; the length and the saturation of that acyl chain set the balance between capsaicin and dihydrocapsaicin, and those two molecules do not speak to TRPV1 with the same voice. This is the quiet reason a lot can carry a great deal of heat *by mass* and still under-deliver *on the tongue* — and the reason we never release against Scoville Heat Units alone. SHU is a mass measurement: how much capsaicinoid a fruit holds, read by HPLC. IFI is a loudness measurement: how strongly that capsaicinoid drives the channel. The two travel together, but not in lockstep, and the gap between them is exactly the territory where a careless gate goes wrong.

CULTIVAR	SKU	SHU (PREDICTED)	IFI-H	SHU PER IFI-H POINT	AXES
Ember	SS-0087	866,000	1,570	552	2/8
Cinder	SS-0088	1,050,000	2,290	459	3/8
Kindle	SS-0089	1,250,000	2,000	625	3/8
Blaze	SS-0091	1,720,000	2,630	654	4/8
Firestorm	SS-0094	2,470,000	3,490	708	6/8
Solara	SS-0095	2,720,000	3,980	683	6/8
Phoenix	SS-0096	2,937,948	4,200	700	8/8

SHU values are computational predictions from our metabolic flux model; they are design targets, not experimentally verified measurements.

Run your eye down the last column and the divergence is plain: the mass carried per unit of loudness is not constant, and in one place it inverts outright. Kindle holds a predicted 1,250,000 SHU — two hundred thousand units *more* capsaicinoid than Cinder's 1,050,000 — and yet registers IFI-H 2,000 against Cinder's 2,290. More mass, a less TRPV1-loud profile; a lot of Kindle that sailed through a Scoville-only gate could still fail an IFI gate, and that gap is precisely the shortfall a formulator would feel cheated by. So we release on the axis the tongue actually reads. The fruit is assessed for what it *does*, not for what it merely contains. Imagine a lot whose fruit average a handsome 2.75 million SHU — comfortably above any Scoville floor — but whose capsaicinoid profile has drifted toward the longer, quieter acyl chains, so the receptor read returns a mean IFI-H of 3,900. On a mass gate it ships without a murmur; on the loudness gate it sits 300 points below the 4,200 target and only a breath above the 3,600 floor, and a formulator blending to an IFI target would find the batch underpowered and never know why. We turn it back, and we are grateful the gate reads the channel and not the chromatogram.



The release assay in progress — a cell-based receptor read and HPLC, each of which consumes the fruit it measures.

From one fruit's scatter to a lot's fraction defective

No two fruit are identical, and the whole question of lot release turns on *how much* they differ. The measure that carries this is the coefficient of variation, CV = σ / μ — the standard deviation of a trait carried as a fraction of its mean, dimensionless by construction, which is exactly why it travels across axes and cultivars without translation. Our engineered lines hold CV ≤ 7%; the wild-type plant we started from runs near 42%. Those two numbers decide everything downstream, because the CV is what converts the fixed ±600 window into a number of standard deviations, and the number of standard deviations is what sets the fraction of fruit that falls outside it.

Work it through on the H axis, where the mean is 4,200. At the 7% we hold as a ceiling, σ = 0.07 × 4,200 = 294 IFI points, and the ±600 window spans only ±600/294 = ±2.04σ. The tails of a centered normal distribution beyond 2.04σ hold 2 × Φ(−2.04) ≈ 4.1% of the fruit — so even a line sitting exactly on target, but running at the contractual ceiling of variation, would present about four fruit in a hundred outside spec. Four percent defective is far too loose for what we want the gate to be. The wild-type plant is worse by an order of magnitude: σ = 1,764, the window a mere ±0.34σ, and fully 73% of fruit outside spec — a harvest no honest program could release at all.

Between those poles lies the number that turns the gate from a gamble into a formality: process capability. For a centered process the capability index is C_{pk} = T / (3σ), where T is the half-width of the window — here 600. To reach the quality-engineering benchmark of C_{pk} = 1.33 we need 3σ = 600 / 1.33, so σ ≤ 150, which is a CV of 150 / 4,200 ≈ **3.6%** — comfortably inside our 7% ceiling, and the level a capable line actually runs at. At σ = 150 the window is a full ±4σ wide, and the tails beyond 4σ hold 2 × Φ(−4) ≈ **63 parts per million**. The leap from the ceiling to the capable line is the whole story of this book in one table:

CV	Σ (IFI-H pts)	±600 window, in Σ	C_{pk} (centered)	Fraction defective, *p*	P(accept) · n = 20, c = 1
42% · wild-type	1,764	±0.34σ	0.11	73%	≈ 0
7% · contract ceiling	294	±2.04σ	0.68	4.1%	0.80
5%	210	±2.86σ	0.95	0.42%	0.997
3.6% · capable line	150	±4.0σ	1.33	63 ppm	≈ 1

Notice what the last column already foreshadows: the sampling plan barely has to work once the CV is small. At 63 ppm defective, a lot is accepted essentially every time; at 4.1%, the same plan starts rejecting one lot in five. The sample does not create that difference — the controlled environment that pins the CV does. The gate merely inherits it.



The controlled chamber that holds a grow steady — where the CV is actually won, long before any fruit reaches release.

The plan: *n*, *c*, and the operating characteristic

A sampling plan is two numbers: **n**, how many fruit we pull from the lot, and **c**, how many of them may fall outside the window before the lot is rejected. Behind those two numbers lies a curve — the operating characteristic, or OC curve — that plots the probability of accepting the lot against the true fraction of it that is out of spec. A good plan has an OC curve that is steep in the right place: it accepts a genuinely good lot almost always, rejects a genuinely bad one almost always, and makes up its mind sharply across the narrow band between.

Make it concrete, because "sample some and decide" hides real arithmetic. Pull twenty fruit and release the lot if at most one falls outside the window — a plan of n = 20, c = 1. For a large lot the binomial gives the acceptance probability directly: the chance of seeing zero or one defective in twenty draws when the true fraction defective is *p*. The two terms are $P(X = 0) = (1 - p)^{20}$ and $P(X = 1) = 20\,p\,(1 - p)^{19}$, and because c = 1 they collapse to a tidy closed form worth keeping:

> $P(\text{accept}) = (1 - p)^{n-1} \cdot [1 + (n - 1)p]$

At p = 1% this is $(0.99)^{19} \times (1 + 19 \times 0.01) = 0.8262 \times 1.19 =$ **0.983**. A quick eye insists that a one-percent-bad lot surely passes ninety-nine times in a hundred; the binomial says 98.3, and the gap between the intuition and the arithmetic is precisely the producer's risk, hiding in plain sight. Plot the acceptance probability across every possible *p* and you have drawn the OC curve: high on the left where lots are good, plunging through the middle, low on the right where they are bad.

We allow that one defect — c = 1, not the stricter-seeming c = 0 — on purpose. An accept-on-zero plan has an OC curve with no shoulder at all: $P(\text{accept}) = (1 - p)^n$, which for n = 20 turns away a perfectly ordinary 1%-defective lot $(0.99)^{20} = 18\%$ of the time. That is a brutal producer's risk to pay for the mere appearance of strictness, and it punishes exactly the good harvests we worked

hardest to grow. Permitting a single out-of-window fruit restores the shoulder, so a good lot passes freely while a genuinely bad one still trips the gate.

TRUE FRACTION DEFECTIVE, *P*	P(ACCEPT) · N = 20, C = 1	P(ACCEPT) · N = 50, C = 1	WHAT IT TELLS US
1% · a good lot	0.983	0.911	the larger plan already costs real producer's risk
5%	0.736	0.280	the indecision band, where steepness shows
10% · a bad lot	0.392	0.034	n = 50 catches what n = 20 waves through two-in-five
20% · badly wrong	0.069	0.0002	both reject; one is merciless

Set the two plans side by side and the price of steepness is laid bare. Going from n = 20 to n = 50 drops the acceptance probability of a 10%-bad lot from 0.392 to 0.034 — a bad lot that escaped two times in five now escapes once in thirty. But look at the top row: the same enlargement also drops acceptance of a *good* 1% lot from 0.983 to 0.911, nearly one good harvest in ten now turned away. Steepness is bought with sample size, and sample size is paid for twice — once in fruit consumed, once in good lots sacrificed. That is why we sample consequential lots hard and trivial ones lightly: the plan is always a negotiation.

The two risks you cannot bury at once

Two risks live on that curve, and they pull against each other, so a plan is always a reconciliation. The **producer's risk** (α) is the chance a *good* lot is rejected by bad luck of the draw — a cost we pay to ourselves in wasted harvest, in fruit grown with care and then turned away. The **consumer's risk** (β) is the chance a *bad* lot is accepted and shipped — a cost paid downstream, by the formulator and ultimately by the person eating. We quantify each against a declared quality level: an **acceptable quality level** (AQL) we promise to pass almost always, and a **rejectable quality level** or lot-tolerance level (LTPD) we promise to catch almost always.

For the n = 20, c = 1 plan, read the risks straight off the curve. Against an AQL of 1%, α = 1 − 0.983 = 1.7% — we reject a good lot seventeen times in a thousand. Against an LTPD of 10%, β = 0.392 — a genuinely bad lot slips through *two times in five* on a single pull. You cannot drive both to zero with a finite sample: tighten the acceptance rule to protect the consumer and α climbs as good lots are rejected; loosen it to spare good lots and β climbs as bad ones pass. To shrink β you enlarge n, which enlarges α unless you also raise c, which widens the indecision band again. The honest design states all four numbers — AQL, LTPD, α, β — out loud, and buys the steepness between them with sample size.

The sizing itself is clean when the fraction defective is small, because the binomial folds into a Poisson in $\lambda = n\,p$. For a c = 1 plan the acceptance probability becomes $P(\text{accept}) = e^{-\lambda}(1 + \lambda)$, and two fixed landmarks fall out of it: the 95%-acceptance point sits at $\lambda_1 = 0.355$, the 10%-acceptance point at $\lambda_2 = 3.89$. Their ratio, $\lambda_2 / \lambda_1 =$

10.9, is a property of c = 1 alone — it says a one-reject plan can only discriminate a good level from a bad level about eleven times higher, no matter how you set n. So anchor it: to protect a 1% AQL at 5% producer's risk, set n = λ_1 / AQL = 0.355 / 0.01 ≈ **36**. That single choice simultaneously pins the consumer's risk near 10% at an LTPD of λ_2 / n = 3.89 / 36 ≈ 11%. The exact binomial for n = 36, c = 1 confirms it — P(accept) = 0.950 at p = 1%, and 0.088 at p ≈ 11% — producer's risk 5.0%, consumer's risk under 9%. Want to discriminate a tighter ratio than eleven, say a 1% AQL against a 4% LTPD, and c = 1 cannot do it at any n; you must step to c = 2 or c = 3, accept a larger sample, and pay the harvest for the resolution. The arithmetic is unsentimental, and that is its mercy: it refuses to let us pretend we are protecting both sides when we are protecting neither.

Why the gate is a formality — and the four ways it is not

Put the capability table and the plan together and the whole exercise relaxes. A capable, in-control Phoenix line runs near C_{pk} 1.33, which pins the true fraction defective near 63 ppm, and at 63 ppm even the gentle n = 20, c = 1 plan accepts with probability 0.99999 — it essentially never turns away a good lot, and the rare rejection is almost always a real special cause worth catching. This is what we mean when we say the gate is not where quality is made. The gate inherits a question that four upstream doors have already answered: the **engineered genotype**, the CRISPR-written Pun1 carrying the double variant S39L+L345G designed to reach our model's predicted SHU target of approximately 2,937,948; the **controlled environment** that holds the grow's conditions steady from transplant to harvest; the **in-process monitoring** that catches drift while it can still be corrected; and the **capable line** itself, whose narrow CV is the reason the OC curve is only ever asked the easy question on its far left. A release assay against a wandering process is a lottery. A release assay against those four shut doors is a formality that occasionally, usefully, catches the one lot that genuinely went wrong.

We say *formality* with care, though, because four honest things can turn it back into real work, and a program that forgets them deceives itself:

- **The gauge must be capable too.** What we observe is not the process spread but the process spread convolved with the assay's own error: σ^2(observed) = σ^2(process) + σ^2(measurement). A noisy assay inflates the apparent fraction defective and rejects good lots — producer's risk we never earned. The guard is the precision-to-tolerance ratio, P/T = 6 σ(measurement) / (USL − LSL); to keep it under 0.1 on our 1,200-wide window, the assay's standard deviation must stay below 20 IFI points, a measurement CV under about 0.5%. Where a single receptor read cannot reach that, we average *m* replicate reads per fruit; the measurement error falls as σ / √m, so four reads halve it and sixteen quarter it — bought, again, with consumed fruit. Dragging an assay that runs at 3% CV down under the 0.5% the window demands would take m = (3 / 0.5)² ≈ 36 reads per fruit, which no honest budget can bear; so in practice we sharpen the

assay itself or widen the effective window enough to swallow the gauge error openly, rather than let it masquerade as process spread.

- **Normality fails near the ceiling.** The IFI is bounded below by zero and capped at 5,000, and Phoenix at 4,200 sits close to that ceiling. A fruit that fully expressed cannot run far above target, so the upper tail is squeezed; a fruit that missed the edit expresses far lower, so the lower tail is heavy. The symmetric-normal estimate of *p* is a convenience, not a law — near the cap we model the lower tail from the data it actually has, because that tail, not the ceiling, is where the real defects hide.
- **A lot must be a single thing.** The binomial assumes the *n* fruit are exchangeable draws from one population. A lot that mixes a sun-crown stratum with a shaded-interior stratum is bimodal, and a plan that happens to pull all *n* fruit from one tray can miss the bad stratum entirely. We define a lot as a single controlled run and sample stratified across the canopy, so the sample is a fair miniature of the whole and not an accident of where the hand reached.
- **Every assayed fruit is spent.** The receptor read and the HPLC both destroy what they measure; the *n* fruit of the sample never reach a table. It is a tithe paid from the harvest in order to know the harvest, and it is the real, unavoidable cost that keeps *n* honest. We sample as hard as the decision deserves and not one fruit harder — because each fruit was grown to nourish someone, and spending it to measure is a thing we do soberly, never by reflex. And because the assay destroys, there is no back-end rescue: a rejected lot cannot be sorted fruit by fruit and its good ones recovered, the way a non-destructive line rectifies a failed batch by screening it whole. There is no average-outgoing-quality floor quietly catching the defects we let slip; every ounce of protection has to come from the plan and the capable process standing in front of it. That absence of a safety net is the deepest reason we spend so much devotion keeping the CV small — when there is no screen to fall back on, the only honest place to build quality is upstream, where the fruit is still growing.

What a release certifies

So we sign the lot, and we are careful about what the signature means. It does not claim every fruit is perfect; a sample can never promise that, and any gate that pretends otherwise is lying with mathematics. It claims something humbler and more defensible: that we measured honestly, on the axis the tongue actually reads, against a window we promised in advance, with a plan whose two risks we stated and bounded — and that a process built to a capability of 1.33 made the question an easy one to answer truthfully. The release assay is the last, quiet confirmation that the care put in upstream actually arrived here, in this tray, in this fruit someone will eat. That is the whole of what a release can be. When the four doors upstream are shut, it is exactly enough — and the fruit that passes the gate goes out carrying not just its heat, but the honesty of the number printed beside it.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.



PART IV OF V

Holding the Hard Cases

the heat ladder, rung by rung · eight axes at once · the number through time and heat

Chapter 15 · The heat ladder — *every rung exactly where it says.*

A ladder is a promise, and the promise is that the fourth rung is always higher than the third — not usually, not on average, but *every time* a hand reaches for it in the dark. Our flagship superhots climb from Ember to Phoenix in deliberate steps, a graded line of *Capsicum chinense* cultivars, each one pinned to a declared place on the Index, and the whole architecture only holds because each rung sits precisely where its number says it will. Let two neighbors wander far enough and they blur: the cook who reaches for "one step hotter" gets a step sideways instead, the gradation collapses into a rumor, and a line meant to be read like a scale becomes a bag of lookalikes under different names. Consistency is not a nicety here. It is the thing that lets a laddered line exist at all.

We did not inherit this discipline from the plant. A pepper left to its own devices is a gorgeous, unreliable instrument — the same seed, grown twice, will hand you fruit that differ in heat by tens of percent, because the enzyme that builds the fire is sloppy and the weather has opinions. The ladder is what we imposed on that unruliness: a set of rungs spaced so widely, and held so tightly, that the gaps between them never close. That is the entire subject here — not how hot, but how *sure*, and the mathematics that turns surety into a product you can sell by the step.



The superhot ladder living in the research greenhouse at golden hour — the rungs are not numbers on a page but plants in soil, each grown to land on its declared mark.

The design principle is simple to state and exacting to meet: *the gap between rungs must dwarf the scatter within them.* Capsaicin is a robust, loud molecule and the strongest single signal a fruit can carry, which is why heat is the tightest hold in our whole library — the flagships run a batch coefficient of variation of two to three percent, where the delicate axes sit higher and the platform as a whole is held to a ceiling of seven. But tightness alone does not make a ladder. What makes a ladder is that the *spacing* between neighbors is engineered to be many multiples of each rung's own width, so the bands never touch, never overlap, never leave a mouth

in doubt about which rung it is standing on. To see why that works, we have to go down to the receptor, because every rung on this ladder is, underneath, a rung of EC_{50}.

The receptor that keeps the promise

Capsaicin binds TRPV1, a non-selective cation channel standing in the membrane of the trigeminal nerve endings that wrap the surface of the tongue. It is the same channel that opens to temperatures above about 43 °C and to a flood of protons — a built-in damage sensor, there long before anyone thought to cook with it. The molecule does not heat the tissue. It slides into a hydrophobic vanilloid pocket between the S3–S4 helices and the TRP helix, and by binding there it lowers the channel's activation threshold below body temperature. A gate that should have stayed shut at 37 °C now stands open. Sodium and calcium pour through; the ending depolarizes; an action potential climbs the nerve and the brain, reading a signal it only ever receives from genuine burning, reports a fire that is not there.

The quantity that governs how *loud* that false fire is the EC_{50} — the concentration of stimulus at which the channel reaches half of its maximum open probability. A fruit whose extract drives TRPV1 to half-max at a low concentration is, by definition, a loud fruit: a little of it goes a long way. One that needs a high concentration to reach the same half-max is quiet. This is the single physical fact under the entire heat axis, and it is why we can speak of a ladder of receptor activation rather than a ladder of taste impressions. Each rung is a cultivar whose extract carries a characteristic EC_{50}; climbing the ladder is walking down a column of falling EC_{50} values, each one lower — louder — than the last.

Where the fire comes from is a short, specific piece of biochemistry. Vanillylamine condenses with an eight-methyl branched-chain fatty acid at the Pun1/AT3 acyltransferase to form capsaicin; swap the acyl chain's length or saturation and the same enzyme makes dihydrocapsaicin instead, a near-twin that differs only in a single saturated bond. The ratio between the two is not a footnote. Capsaicin arrives fast and sharp; dihydrocapsaicin is slower, deeper, and lingers. The two also sit slightly differently in the vanilloid pocket, so a gram of a capsaicin-rich blend and a gram of a dihydrocapsaicin-rich blend do not produce the same EC_{50}. Hold that fact — it is the reason, later, that two fruits of equal raw heat can land on different rungs.

Two scales for one fire: Scoville and the Index

For a hundred years heat had exactly one number, and it was a confession of ignorance dressed as a measurement. Wilbur Scoville's scale asks how far you must dilute an extract in sugar water before a panel of tasters can no longer detect the burn: dilute a sample a million-fold and still feel it, and the sample scores a million Scoville Heat Units. SHU is honest about one thing — total capsaicinoid *concentration*, more or less linearly — and silent about everything else. It is a dilution, filtered through a human panel's threshold, and crystalline capsaicin itself sits at the top of it at 16,000,000 SHU. Our model predicts our summit cultivar, Phoenix, at

approximately 2.95 million SHU — a design placing it at approximately 2.95 ÷ 16 ≈ **18.4 percent of the heat of pure capsaicin** — a cultivar designed to carry nearly a fifth of the pungency of the isolated molecule.

The Index measures something the dilution never could: not how much compound the fruit holds, but how loudly that compound speaks to the receptor. The formula is a ratio and nothing more.

$$\text{IFI-H}(\text{candidate}) = 1{,}000 \times \text{EC}_{50}(\text{reference}) \,/\, \text{EC}_{50}(\text{candidate})$$

Read it slowly. Fix a reference — the native wild-type pepper — and anchor it so that its own half-activation scores exactly 1,000. Any candidate's score is then the reference's EC_{50} divided by the candidate's, scaled by a thousand. A louder fruit has a *lower* EC_{50}, so the ratio runs *up*: loudness and score move together, which is what you want from a number a cook will read. The arithmetic has two gifts hiding in it. Divide any IFI-H by 1,000 and you read, directly, how many times louder than the native reference the fruit is — Phoenix at 4,200 is 4.20× the reference's loudness. Invert that and you recover the fruit's EC_{50} as a fraction of the reference's: 1,000 ⁄ 4,200 ≈ 0.238, meaning Phoenix is designed to reach half-max TRPV1 activation at under a quarter of the concentration the wild type needs. The whole ladder can be rewritten in those two derived columns, and doing so is the first thing we do with any candidate.

CULTIVAR	SKU	PREDICTED SHU	IFI-H	LOUDNESS (×REF)	BATCH CV
Ember	SS-0087	866K	1,570	1.57×	2.1 %
Cinder	SS-0088	1.05M	2,290	2.29×	2.3 %
Blaze	SS-0091	1.72M	2,630	2.63×	2.6 %
Wildfire	SS-0092	1.96M	2,930	2.93×	2.9 %
Scorch	SS-0093	2.21M	3,220	3.22×	3.0 %
Firestorm	SS-0094	2.47M	3,490	3.49×	3.2 %
Solara	SS-0095	2.72M	3,980	3.98×	3.4 %
Phoenix	SS-0096	2.95M	4,200	4.20×	2.4 %

All SHU values are model predictions. Batch CV figures are design targets.

Now look at what SHU and IFI-H do *not* do, which is move in perfect lockstep. If the two scales measured the same thing, the ratio of IFI-H to SHU would be a constant down the whole column. It is not. Ember yields about 1,813 points of Index per million SHU; Cinder, 2,181; and the big flagships converge toward a narrower band — Scorch at 1,457, Firestorm at 1,413, Solara at 1,463, Phoenix at about 1,430. The drift is not noise. It is the fingerprint of a shifting capsaicin-to-dihydrocapsaicin blend: a fruit that leans capsaicin-rich buys more receptor loudness per unit of raw dilution heat, so it earns more IFI-H per SHU. SHU weighs the molecules; the Index weighs how they land.

The clearest proof lives in a cultivar we keep *off* the ladder on purpose. Kindle (SS-0089) is predicted at 1.25M SHU — hotter, by dilution, than Cinder's 1.05M — yet its IFI-H comes in at 2,000, *below*

Cinder's 2,290. More total capsaicinoid, less receptor loudness: a blend tilted enough toward dihydrocapsaicin that it reads quieter at TRPV1 than a fruit carrying less. On a Scoville ladder Kindle ranks above Cinder; on the Index ladder it would have to sit below it, out of order with its own Scoville number. A rung cannot be ambiguous about which way is up, so Kindle is a catalogued cultivar but not a step on this particular staircase. (Flare, SS-0090, sits out for a plainer reason: it carries no IFI-H prediction in our design model, and a cultivar without a predicted IFI-H cannot be a rung on this scale.) The ladder in the table is a curated, strictly monotonic path chosen *through* the catalogue — which is itself a design act, not a given.

The spacing law

Give the design intuition its number. For any two adjacent rungs, define the separation ratio

> `Z = (gap between the means) ⁄ (standard deviation of a rung)`

where a rung's standard deviation in Index points is simply its mean times its coefficient of variation, σ = IFI-H × CV. Z is the gap measured in units of a rung's own width — the same three-sigma logic that governs a control chart, turned on its side and used to *space products* instead of to bound a process. A target of Z ≥ 3 means the clouds clear each other by three standard deviations, and two Gaussians set three-plus σ apart overlap by a vanishing fraction: the probability that a hot fruit of the lower rung is mistaken, by tongue or assay, for a cool fruit of the higher one falls below a tenth of a percent.

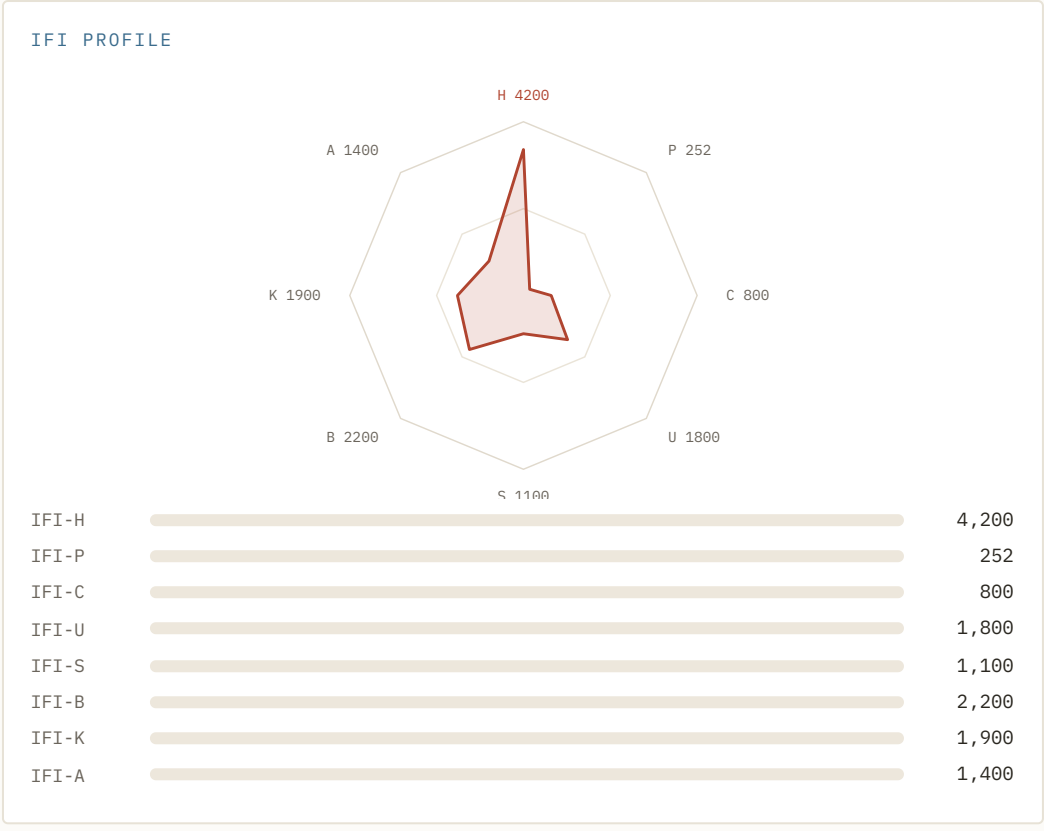
STEP (LOW → HIGH)	Σ, LOW RUNG	GAP (ΔIFI-H)	Z = GAP ⁄ Σ
Ember → Cinder	33.0	720	21.8
Cinder → Blaze	52.7	340	6.45
Blaze → Wildfire	68.4	300	4.39
Wildfire → Scorch	85.0	290	3.41
Scorch → Firestorm	96.6	270	2.80
Firestorm → Solara	111.7	490	4.39
Solara → Phoenix	135.3	220	1.63

Work one row by hand, because the method is the point. Wildfire scores 2,930 at a batch CV of 2.9 percent, so its standard deviation is 2,930 × 0.029 ≈ 85 points of Index. Its neighbor, Scorch, sits 290 points above it. The separation is 290 ⁄ 85 ≈ **3.41 σ** — and the chance a Wildfire fruit climbs past the midpoint into its neighbor's half is P(Z > ¹⁴⁵⁄₈₅) = P(Z > 1.71) ≈ 4 percent, with the chance it reaches all the way to the neighbor's *mean* only P(Z > 3.41) ≈ 0.03 percent. You could reach for that rung blindfolded a thousand times and miss it by a full step essentially never.

Then read the table honestly, because it does not say what a tidy slogan would want it to. The foot of the ladder is absurdly over-separated — Ember and Cinder stand nearly twenty-two σ apart, a gulf no scatter could ever bridge. The middle holds the three-sigma target with room: better than six σ, then 4.4, then 3.4. But two steps

do not clear three. Scorch to Firestorm sits right on the line at 2.80 σ, close enough that we treat three-sigma as the design *target* rather than a boast. And the final climb — Solara to Phoenix — compresses to **1.63 σ**, the tightest pair in the whole line. This is not a flaw we are hiding; it is physics we are reporting. Near the ceiling of what the biosynthetic pathway can do, every additional point of heat costs more than the last, and the rungs crowd. The gap from Solara to Phoenix is approximately 230,000 predicted SHU in our model, but in Index points it is only 220 — and against Solara's σ of 135, that gap alone cannot carry a three-sigma promise.

So at the summit we stop asking heat to do all the separating, and this is where the full profile earns its place. Phoenix is the one cultivar in the line that registers on all eight axes at once.



Solara expresses six of eight; Phoenix expresses all eight. Where their heat bands crowd to 1.6 σ, two entire additional axes of flavor — pungency and the richnesses layered beneath it — give the mouth other handles to tell them apart. A taster separating the top two rungs is not squinting at a 220-point difference in a single loud signal; they are reading a fruit that speaks on eight channels against one that speaks on six. And we hold Phoenix itself to the tightest CV in the line, 2.4 percent, tighter than either of its neighbors below — the deliberate engineering answer to a narrowing gap is a narrowing

band. (At 4,200 against the scorecard's per-axis ceiling of 5,000, Phoenix's heat also sits at 84 percent of the normalization maximum — near the top of the readable scale, with just enough headroom left to keep the summit legible.) The scale stays a scale to the very top, but above a certain altitude it is the whole profile, not the heat alone, that keeps the promise.

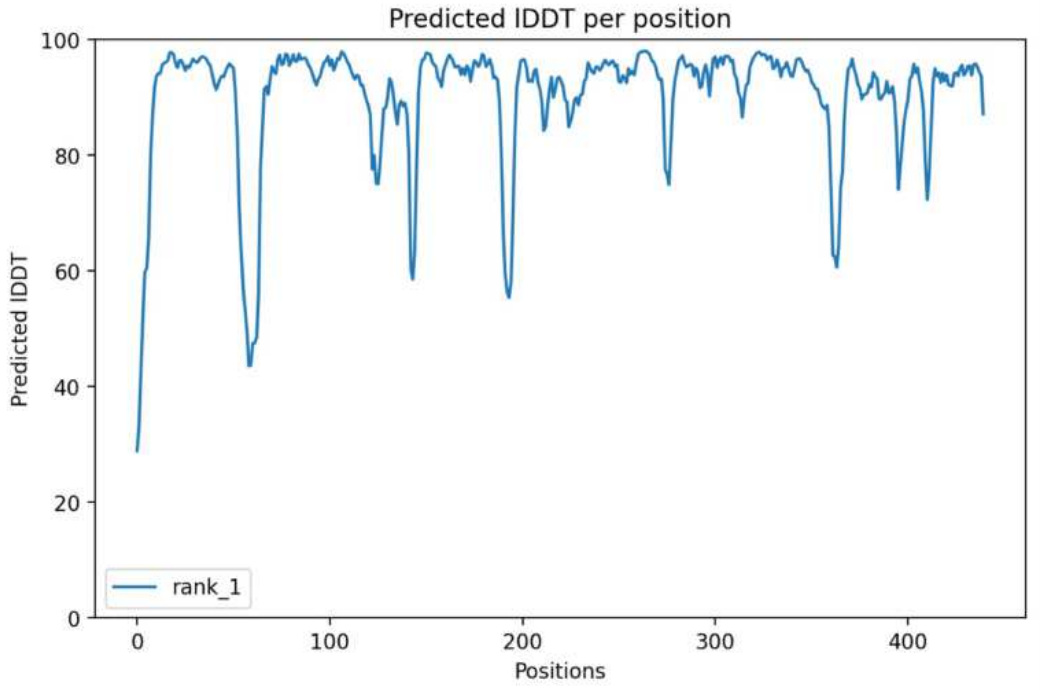
What a held CV actually buys

It is worth naming the failure precisely, because it is exactly what an untended process delivers, and it shows that the gradation lives in the *scatter*, not the scores. Take Wildfire and let its CV swell from 2.9 percent to ten — not catastrophic, merely unheld. Its σ grows from 85 points to 2,930 × 0.10 ≈ 293, which now *exceeds* the 290-point gap to Scorch entirely. The separation collapses from 3.4 σ to 290 ⁄ 293 ≈ 0.99 σ. Ask how much of the lower rung's fruit now crosses the midpoint into the neighbor's territory: P(Z > ¹⁴⁵⁄₂₉₃) = P(Z > 0.49) ≈ **31 percent**. Nearly a third of every Wildfire harvest would read as something other than Wildfire. The two rungs have merged. The cook reaching one step up has close to a coin-flip's worth of chance of landing sideways, and the ladder, which we could still print perfectly on every label, has dissolved underneath the labels.

That is the whole argument in one number. The gradation is not a property of the *scores* — we could stamp 2,930 and 3,220 on anything. It is a property of the *held CV*, which no label can fake and no marketing can supply. This is the moment consistency stops being a quality metric and becomes product architecture: a held CV is not overhead subtracted from the real work of making heat. It is the real work, and it is the thing the customer is actually buying when they buy a rung instead of a lottery ticket.

Where the tightness is made

A wild pepper cannot do this. Left to the native enzyme, batch-to-batch heat scatters at a coefficient of variation near 42 percent — the kind of spread that makes "a hot one" a matter of luck. Pulling that down to a seven-percent platform ceiling, and to two-to-three percent on the heat axis specifically, took rewriting the enzyme itself. Two substitutions in Pun1 (the double variant S39L+L345G), found by scoring every possible single-residue change with a protein language model and cross-checked against a genome-scale flux-balance model before a single edit was made, were written in with CRISPR.



The Pun1/AT3 acyltransferase model, coloured by per-residue confidence — the binding domain and active-site cleft, where both shipped edits land, resolve at high confidence; structural modelling also identified a disulfide-misfold site (C175S) that informed the design but was not included in the final double variant.

EDIT	ΔΔG (KCAL/MOL)	ROLE	WHAT IT MOVES
S39L	−2.1	N-terminal binding affinity ↑	raises the mean (peak output)
L345G	−1.8	steric relief at the active-site cleft	raises throughput (peak output)
C175S *(structural annotation — not in final design)*	−0.9	disulfide-misfold site identified by structural modelling	informed design; not in shipped construct

Read that table for which edit does what, because the division of labor matters. S39L and L345G, the two shipped substitutions that form the double variant, do the work of both lifting the *mean* — they are why Phoenix is designed to reach approximately 2.95 million SHU — and stabilizing the fold sufficiently to hold the CV within its declared range. Our structural modelling also identified a third site, C175S (ΔΔG −0.9 kcal/mol), as a disulfide-misfold candidate: a stray cysteine that, when unedited, can fold the enzyme wrong in some fruit and right in others, contributing to batch scatter. That finding shaped our understanding of where variance lives in the protein, but C175S was pulled from the final design — the double variant proved sufficient to hold the CV without it. What the structural insight illuminates is worth carrying anyway: misfolding is a dominant source of batch-to-batch scatter, and knowing which route it travels is how you close it. A scale is not built only by the edits that raise the peak. It is built by understanding every route the variance can take, and closing them, whether the fix goes into the construct or only into the model.

How many fruit make a rung

Holding a tight process is one thing; *knowing* you have held it is another, and it is pure sampling arithmetic. A rung's declared mean is an estimate from a finite harvest, and the standard error of that estimate falls as the square root of the count: SE = σ / √n, or in relative terms, CV / √n. To pin a rung's mean to within ±1 percent when its CV is 3, we need CV / √n ≤ 0.01, so √n ≥ 3 and n ≥ 9 fruit. To pin it to ±0.5 percent, √n ≥ 6 and n ≥ 36. The precision we quote

on a rung is therefore a direct statement about how many fruit we assayed to place it — a declared mean is a sample size worn on the sleeve.

The scatter itself can be taken apart the same way, because the CV we measure is not all biology. Total variance is the sum of what the plant does and what the measurement does: $CV^2(\text{total}) = CV^2(\text{biological}) + CV^2(\text{assay})$. Suppose, to illustrate, the assay contributes a point and a half of scatter on its own; a rung reading three percent total then carries a biological spread of only $\sqrt{(3^2 - 1.5^2)} = \sqrt{6.75} \approx 2.6$ percent. Separating the two matters, because they are closed by different means — the assay term by better instruments and reference discipline, the biological term by the enzyme and the growing environment. Chasing a plant for scatter that actually lives in the pipette is a way to work hard and tighten nothing.

The ways a ladder falls apart

A ladder has a small number of distinct failure modes, and naming them is how we hold against each.

The first is **scatter creep** — the CV-to-ten collapse above. It is insidious because it changes no printed number; the labels stay perfect while the product behind them turns to fog. The defense is a held σ, monitored like a control chart, with the three-sigma spacing as the alarm line.

The second is **non-monotonicity**, the Kindle case: a cultivar hotter by one scale and quieter by another, so that two honest numbers disagree about which rung is up. The defense is to choose the ladder's path through the catalogue deliberately, on a single scale — the receptor-anchored one — and to leave off the fruit that would make the staircase double back on itself.

The third is **ceiling compression**, the Solara-to-Phoenix squeeze. Near the top of the pathway, heat gains cost exponentially more and the rungs crowd below three σ on heat alone. The defense is the full profile: let the other seven axes do the separating work the compressed heat gap cannot, and hold the summit rung to the tightest CV in the line.

The fourth is **environmental drift**. Capsaicinoid synthesis answers to temperature, light, and water stress; a rung grown under a swinging sky will wander off its mark no matter how good the enzyme is. This is where the controlled chamber earns its keep — and where our reasons run deeper than metrology.



Phoenix in the controlled chamber — the environment held steady not for sterility's sake but so the plant is given exactly what it needs to express its design faithfully, every harvest the same as the last.

A steady environment is not a clean-room pose. It is the practice of giving a plant precisely what it needs — the right light, the right water, nothing in excess and nothing withheld — so that it can grow into the fullest, most faithful version of what it was designed to be. A fruit raised that way lands on its rung because nothing external pushed it off. The same care that keeps the heat on its mark is the care that lets the plant carry the most it can carry; holding a CV and growing a plant well turn out to be the same act, approached from two directions.

The last failure is the subtlest: **reference drift**. Every IFI-H on the ladder is a ratio against a single anchor, the native wild-type reference held at 1,000. If that anchor shifts — if the reference material degrades, or the assay that reads it wanders — then every rung above it floats, all together, invisibly, because the ratios still look self-consistent. The entire scale is only ever as stable as the one fixed point it hangs from. So the reference is guarded more jealously than any flagship: it is the zero of the ruler, and a ruler whose zero moves measures nothing at all.

The rung exactly where it says

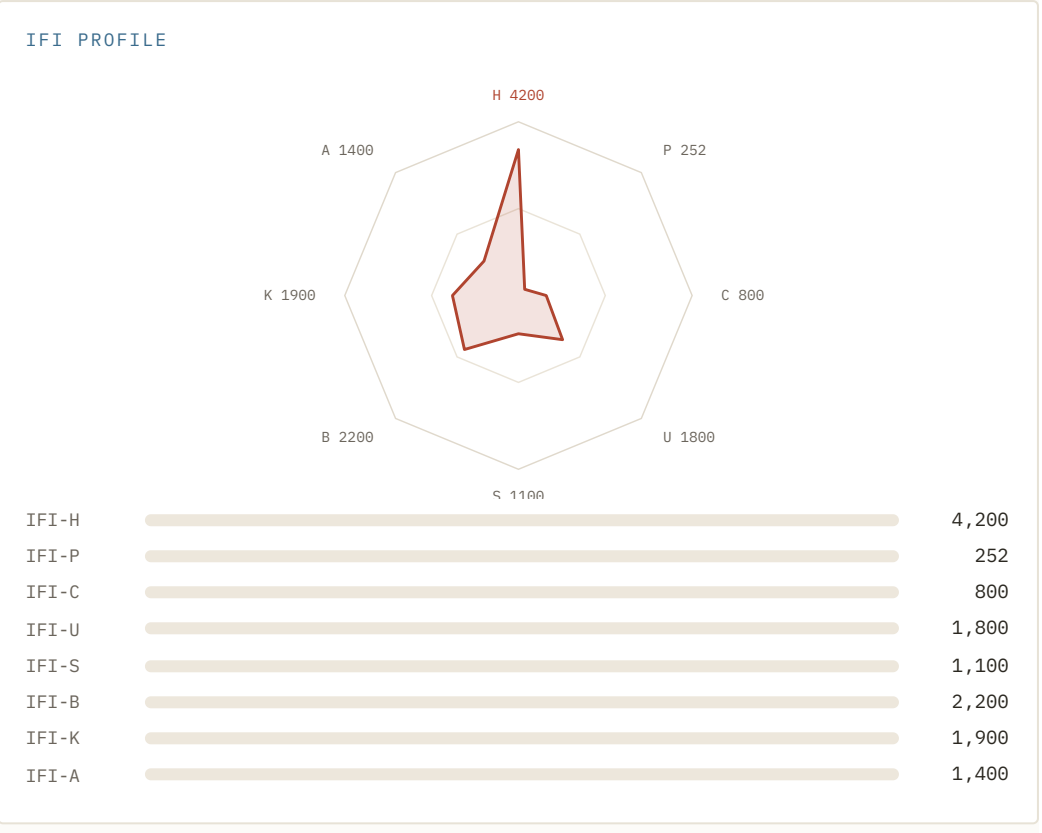
Set the labels aside and stand in front of the thing itself: eight cultivars in soil, climbing from a warmth that announces itself politely to a fire that carries all eight axes at once. What makes that a ladder and not an assortment is nothing you can see in the fruit. It is a column of falling EC_{50} values, spaced by gaps engineered to run many times wider than each rung's own scatter, held there by an enzyme we rewrote specifically so it would stop wandering — the double variant raising the peak, the structural modelling mapping every failure route worth closing. We can tell you the probability that a reaching hand lands wrong, and across the readable heart of the ladder it rounds to zero. That number, not the Scoville on the label, is the real product. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory. A ladder is a promise, and the whole of this work is the arithmetic that lets us keep it — every rung exactly where it says, every harvest, in the dark, by feel, without looking down.

Chapter 16 · Holding eight at once — *the Phoenix problem.*

A single axis is a thing you can corner. Eight at once is a negotiation. Phoenix — SS-0096 — is the only cultivar we have ever designed to register on all eight pathways, and its difficulty is not that eight numbers are individually hard to reach. It is that the eight *interact.* The axes draw on a shared pool of precursor carbon and reducing power, so they are not eight independent dials but eight claims on one budget, and pushing any one of them pulls at all the others. The whole chapter is the story of how we hold a profile steady when every number in it is quietly trying to spend the same coin.



Phoenix, SS-0096 — eight axes held at once, not because eight easy numbers were stacked, but because the hardest balance in the library was engineered to survive its own plants.

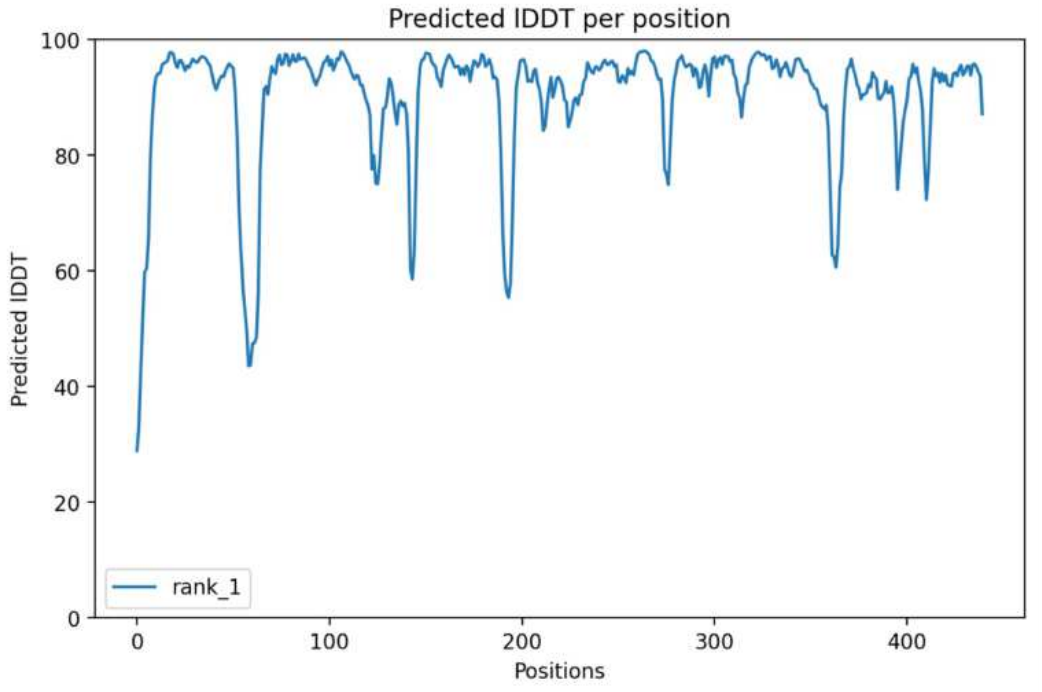


Eight numbers that each look reachable alone can become unreachable together — not because any one is too high, but because a real plant cannot afford all eight at the same instant. A profile that balances perfectly in one ideal fruit can fall out of balance the moment a living plant varies even slightly, robbing one axis to pay another. To understand how we keep that from happening — how a lot of real fruit comes back with all eight

numbers where we put them — you have to follow the carbon down into the plant, watch it get spent, and then watch us do the arithmetic that decides, in advance, which way to spend it.

The shared trunk — one skeleton, eight withdrawals

Every molecule in that profile is built from the same phenylpropanoid carbon, and the first five steps of the road are shared by nearly every green thing that has ever grown. Phenylalanine ammonia-lyase (**PAL**) opens the trunk by stripping the amino group from L-phenylalanine to make *trans*-cinnamic acid — the single committed entrance where carbon leaves the business of making protein and enters the business of making flavor. Cinnamate-4-hydroxylase (**C4H**), a cytochrome-P450, hydroxylates it to *p*-coumaric acid, and it does so by burning **NADPH** — the first of the shared currencies, spent here as the reducing electrons the P450 needs to split oxygen. 4-coumarate:CoA ligase (**4CL**) then activates the acid onto coenzyme A; a hydroxycinnamoyl transferase (**HCT**) shuffles it toward the caffeoyl intermediates; and caffeic-acid *O*-methyltransferase (**COMT**) lays down a methyl group drawn from **S-adenosylmethionine (SAM)** — the second shared currency, the universal methyl donor — carrying the carbon the last step toward vanillin.



The Pun1/AT3 acyltransferase, modeled and colored by confidence — the enzyme where the heat branch lays its claim on the shared carbon, and the reason a push at one axis is felt at all eight.

Only four enzymes at the very end are pepper's own, and they are where the heat axis finally distinguishes itself. A putative aminotransferase (**pAMT**) converts vanillin to vanillylamine — the aromatic half of capsaicin. In parallel, a branched-chain aminotransferase (**BCAT**) and a ketoacyl-ACP synthase (**KAS**) assemble the acyl half, an 8-methyl branched-chain fatty acid, out of valine or leucine and **malonyl- and acetyl-CoA** — the third and fourth shared currencies, the two-carbon bricks that also build every membrane lipid the plant needs to stay alive. Finally capsaicin synthase (**Pun1/CS**, the AT3 acyltransferase) condenses the two

halves into capsaicin itself. The length and saturation of that acyl chain set the balance between capsaicin and dihydrocapsaicin, and with it the *shape* of the burn — a detail that will matter when we compare the heat number to the Scoville number.

So the picture is a single tree. One trunk — PAL → C4H → 4CL → HCT → COMT — feeding a crown of terminal branches, each branch an axis, every branch drinking from the same four wells: phenylpropanoid carbon, NADPH, SAM, malonyl-/acetyl-CoA. Push heat toward its headline 4,200 and the capsaicin branch draws hard on that shared carbon and on SAM — and can quietly starve the bitterness-attenuation machinery of the very cofactors it needs to hold its 2,200. The tree does not care about our scorecard. It spends what it has, in the order the enzymes demand it, and our entire job is to make sure that when it is done spending, eight numbers are still standing.

Reading loudness — what the index actually measures

Before we can hold eight numbers, we have to say exactly what a number *is*, because the thing we hold constant is not an amount of any compound. Each axis is anchored to the half-maximal activation of its receptor, through one clean relation:

$$\mathrm{IFI}(\text{candidate}) = 1{,}000 \times \mathrm{EC}_{50}(\text{reference}) \div \mathrm{EC}_{50}(\text{candidate})$$

EC_{50} is the concentration at which a stimulus drives its receptor to half of full response. It is read off a dose–response curve, the sigmoid every receptor obeys — modeled by the Hill equation, *response*(C) = R_max · C^n ÷ (EC_{50}^n + C^n), where at C = EC_{50} the response is exactly R_max⁄2 and the Hill coefficient *n* sets how steeply the curve rises (for capsaicin on TRPV1, *n* sits near 2, a mild cooperativity). The index deliberately reports the EC_{50}, the *position* of that curve, and never R_max, its *height*. Position is potency — how little of a food it takes to be heard. Height is efficacy — how loud the receptor can be made to shout at saturation. We measure position because it is the robust half of the curve: it barely moves when the assay's gain drifts, so two labs on two days still agree on the number. Loudness, in this book, always means potency.

Invert that potency against a fixed reference, scale by a thousand, and the arithmetic is simple to run in both directions. Each axis carries its own stated convention reference — the constant that defines IFI = 1,000 on that axis, like a Scoville dilution-to-threshold definition applied per receptor. Those reference constants are independent of one another, so IFI values across different axes are axis-relative: a score of 4,200 on the heat axis and a score of 252 on the pungency axis are each a fold-change versus their own axis's reference, not versus a single shared zero-point. Read the profile as a chord, not as a ranking. Every canonical score hands back a candidate EC_{50} by rearrangement: EC_{50}(candidate) = 1,000 ÷ IFI. Run Phoenix's whole profile through it and the eight scores become eight concentrations:

AXIS	IFI	FOLD VS REFERENCE (IFI÷1,000)	EC_{50} (REFERENCE ≡ 1.000)	READS AS
IFI-H · Heat	4,200	4.20×	0.238	4.2× louder
IFI-B · Zero-Bitter	2,200	2.20×	0.455	2.2× louder
IFI-K · Kokumi	1,900	1.90×	0.526	1.9× louder
IFI-U · Umami	1,800	1.80×	0.556	1.8× louder
IFI-A · Astringency	1,400	1.40×	0.714	1.4× louder
IFI-S · Sweetness	1,100	1.10×	0.909	1.1× louder
IFI-C · Cool-Burn	800	0.80×	1.250	quieter than reference
IFI-P · Pungency	252	0.25×	3.968	quieter than reference

Read the last column slowly, because it carries the design. Heat reaches its receptor's half-point at 23.8% of the reference concentration — a TRPV1 signal 4.2 times louder than wild-type, the loud, legible backbone of the fruit. Two axes sit below the 1,000 anchor: cool-burn at 800, a deliberately soft TRPM8 note held under the heat so the two thermal channels do not fight; and pungency at 252, a quiet TRPA1 presence that marks the pepper's character without competing for the top of the chord. The index is signed around a thousand: above it a food is louder than the wild plant, below it the food is softer, and a full-spectrum cultivar is not eight shouts but a chord with its own quiet intervals written in on purpose.

Where the Scoville number and the receptor number part ways

It is tempting to think heat is the easy axis because we already have a number for it — the Scoville Heat Unit, two centuries old. But SHU and IFI-H measure two different things, and the gap between them is one of the most useful honest facts in the library. SHU is chemistry: it is linear in capsaicinoid concentration, a dilution count, with pure capsaicin fixed at 16,000,000 SHU. IFI-H is biology: it is tied to where the TRPV1 dose–response curve sits. The two cannot stay proportional, because the receptor saturates and the chemistry does not — as you pile on capsaicinoid, each new increment climbs a flatter part of the sigmoid and buys less receptor-loudness than the last. Our design tier progression makes the divergence visible as a column:

CULTIVAR	SKU	PREDICTED SHU	IFI-H	PREDICTED SHU PER IFI POINT	AXES
Ember	SS-0087	866K	1,570	552	2/8
Cinder	SS-0088	1.05M	2,290	459	3/8
Kindle	SS-0089	1.25M	2,000	625	3/8
Blaze	SS-0091	1.72M	2,630	654	4/8
Wildfire	SS-0092	1.96M	2,930	669	5/8
Scorch	SS-0093	2.21M	3,220	686	5/8
Firestorm	SS-0094	2.47M	3,490	708	6/8
Solara	SS-0095	2.72M	3,980	683	6/8
Phoenix	**SS-0096**	**2.95M**	**4,200**	**700**	**8/8**

The right-hand column broadly rises — from the mid-500s at the bottom of the tiers to the high-600s and into the 700s at the top. Each IFI point costs more predicted SHU as you climb, which is exactly the compression you would predict from a saturating receptor: the chemistry outruns the biology at the summit. Our model predicts approximately 2.95 million SHU for Phoenix — about 18.4% of pure capsaicin's 16,000,000 — yet the design reads 4,200 in loudness, because the last stretch of the Scoville scale is the flat top of the dose–response curve, where enormous increments of capsaicinoid translate into modest gains in what the tongue's receptor can report. These cultivars represent increasing computational design scope and axis activation, with SHU and IFI-H values as predicted outputs of that design work, not certified measurements.

The Hill curve makes that compression exact. Fractional activation follows $f = C^n \div (EC_{50}^n + C^n)$; with the capsaicin cooperativity of $n \approx 2$, doubling the stimulus near the bottom of the curve — from $C = 0.5 \cdot EC_{50}$ to $C = 1 \cdot EC_{50}$ — lifts receptor occupancy from 0.20 to 0.50, a gain of 0.30. Double it again near the top — from $2 \cdot EC_{50}$ to $4 \cdot EC_{50}$ — and occupancy climbs only from 0.80 to 0.94, a gain of 0.14. The same doubling of chemistry buys less than half the biology, and a fourfold push near saturation buys almost nothing at all. That is the whole reason each IFI point costs more predicted SHU as the tiers climb: the Scoville scale keeps counting capsaicinoid in a straight line while the receptor, already most of the way to full, has less and less left to give.

And then there is the wrinkle that proves the two numbers are genuinely independent. Cinder (SS-0088) carries 1.05M predicted SHU and reads IFI-H 2,290. Kindle (SS-0089) carries *more* chemistry — 1.25M predicted SHU — yet reads *lower*, IFI-H 2,000. A hotter pepper by dilution count, 200,000 SHU hotter in the model, that speaks 290 points quieter to the receptor. Nothing is wrong with either number; they are reporting different realities. SHU counts total capsaicinoid mass. IFI-H counts receptor engagement, which depends on the *mix* — the capsaicin-to-dihydrocapsaicin ratio set by the length and saturation of that branched acyl chain the KAS and BCAT enzymes build. Change the shape of the burn and you change which number leads. The honest limit is plain: never read one of these as a proxy for the other. When a formulator wants the sensation a mouth will actually feel, the receptor number is the one to hold.

The budget, written as arithmetic

Now put the four shared wells side by side with what draws on them, and the eight-way problem stops being a metaphor and becomes a ledger:

SHARED CURRENCY	SPENT AT	AXES DRAWING ON IT	THE COMPETITION IT CREATES
Phenylpropanoid carbon (PAL entry)	PAL → C4H → 4CL → HCT	every axis on the trunk	the master budget; over-drawn by heat, the rest starve
NADPH (reducing power)	C4H and the reductive steps	heat and several aromatics	a finite pool; heavy synthesis competes for every electron
S-adenosylmethionine (SAM)	COMT; vanillyl formation	heat and the bitter/aromatic methylations	the methyl donor two branches need at once
Malonyl- / acetyl-CoA	KAS, the branched-chain acyl arm	the acyl half of capsaicin	a chokepoint, shared with membrane-lipid synthesis

Flux-balance analysis writes each of those rows as a linear inequality: the total consumption of a currency, summed across every branch that spends it, may not exceed its supply. If v_a is the flux we send down axis *a*, and c(currency, a) is how much of a currency one unit of that flux costs, then for each well, Σ_a c(currency, a) · v_a ≤ Supply(currency). Eight branches, four constraints, one feasible region — the set of all flux distributions a real plant can actually run.

Put illustrative numbers on just the SAM ledger to feel the shape of it. Say the pool delivers 100 units of methyl per unit time. At the Phoenix target the heat branch's vanillyl methylation draws 55 of them, the aromatic and bitter-side methylations draw 30, and the remaining methyl sinks take 15 — summing to exactly 100. That is a *binding* constraint: the budget is spent to the last unit, with no slack. Now let a single greenhouse run heat 10% hot — a perfectly ordinary plant-to-plant wobble. The heat branch's SAM draw rises from 55 to 60.5, and the demand on the pool becomes 60.5 + 30 + 15 = 105.5 against a supply of 100. The plant cannot conjure the missing 5.5 units; something must give. If the deficit lands on the bitter branch, its 30 units fall to 24.5 — a *18.3%* collapse in an axis we never touched, produced entirely by a 10% nudge somewhere else. That is the failure mode in one line: on a binding constraint, a small wobble in one axis spills straight into another, amplified by how tightly the budget was already drawn.

In our model, Phoenix's Zero-Bitter does not wander 18% — the design holds it to 5.2% CV. The whole distance between that illustrative 18% and the designed 5.2% is the work of the next section — the deliberate choice, made before a single seed goes in, of *how* to spend the budget so that the wobbles do not spill.

Robust beats correct — elasticity as the design variable

A naïve search for Phoenix's profile returns any flux distribution that hits the eight targets — and there are many. Most of them are brittle. They sit hard against a binding constraint, like the SAM ledger spent to its last unit, where the smallest drift tips a currency over its limit and a number falls out of the profile. Hitting the target is *correct.* Surviving the plant that carries it is *robust.* Those are different design problems, and we solve the second one on purpose.

The quantity that separates them is **elasticity** — the sensitivity of an axis's output to a small perturbation in a shared pool, written as a partial derivative in logarithms: E(axis, currency) = ∂ ln(output) ÷ ∂ ln(supply). It answers one question: if the plant's carbon or NADPH drifts by one percent, by what percent does this axis move? A low-elasticity profile barely flinches; a high-elasticity profile is a house of cards. To first order, elasticity is also the exchange rate between the variation going *in* and the variation coming *out*: if a shared pool fluctuates lot-to-lot with some coefficient of variation CV(input), then an axis fed mainly by that pool inherits CV(output) ≈ |E| · CV(input), and across several contested pools the variations add in quadrature, CV(output) ≈ √(Σ E² · CV(input)²). Elasticity is not an abstraction. It is the multiplier that turns a plant's ordinary unsteadiness into our lot's imperfection.



The sequencing bench where a construct is verified before it is ever grown — because the elasticity we design against is chosen in the arithmetic, long before a lot exists to measure it.

So the flux model does not stop at *feasible.* Among all the distributions that hit the eight targets, it ranks them by elasticity and chooses the one whose eight numbers move *least* when the shared pools wobble — even when a more brittle solution would post slightly prettier headline numbers. We take the robust one every time, because a number that cannot survive its own plant is not a specification; it is a photograph of one lucky fruit. Often the robust move is to leave a well deliberately under-spent — to hold a currency a few percent below its binding limit so that an ordinary wobble has somewhere to go instead of spilling into a neighbor. It is a quiet trade: we decline the prettiest attainable headline number in exchange for eight numbers that stay put, and we make it on purpose, because steadiness is the specification and the headline is

only a boast. The design logic shows in the predicted lot profile. If the plant's shared pools all drift with roughly the same underlying CV(input), then the *ratio* of two axes' designed CVs is the ratio of their elasticities. Our model predicts heat CV at 2.4%; astringency at 6.1%. Astringency therefore carries about 6.1 ÷ 2.4 ≈ 2.54× the elasticity of heat — it is two and a half times as exposed to budget drift, which is exactly why it is the last and hardest axis to pin, drawing as it does on contested carbon at the far end of eight simultaneous withdrawals. We cannot drive that ratio all the way to one; the residual spread of elasticity, from heat's 1.00 up to astringency's 2.54, is the coupling we could not fully remove. But we chose the solution that made it as small as it is. That is the difference between reaching a target and holding it.

The design profile — eight numbers built to hold

A coefficient of variation is just the standard deviation divided by the mean, CV = σ ÷ μ, expressed as a percent — the plain honest ratio of how much a thing wanders to how big it is. The design specification is built around exactly this process: grow a lot in a controlled chamber, sample many fruits as replicates, run each extract across a dose–response on cells expressing the axis's receptor, fit the EC_{50} by Hill regression, convert to IFI, and read σ and μ across the whole lot. Our model runs that process forward for Phoenix, and the eight numbers carry their predicted spreads:



The controlled chamber, where a lot is grown to be measured, not admired — every fruit a replicate, every replicate a test of whether eight numbers stayed where we put them.

AXIS	TARGET M (IFI)	DESIGNED CV	Σ = CV · M (IFI POINTS)	ELASTICITY VS HEAT (CV ÷ 2.4%)
IFI-H · Heat	4,200	2.4%	100.8	1.00×
IFI-U · Umami	1,800	3.8%	68.4	1.58×
IFI-P · Pungency	252	4.1%	10.3	1.71×
IFI-S · Sweetness	1,100	4.3%	47.3	1.79×
IFI-C · Cool-Burn	800	4.6%	36.8	1.92×
IFI-B · Zero-Bitter	2,200	5.2%	114.4	2.17×
IFI-K · Kokumi	1,900	5.6%	106.4	2.33×
IFI-A · Astringency	1,400	6.1%	85.4	2.54×

The relative column tells the design story — heat tightest, the composed subtler axes nearer the line, astringency the loosest at 6.1%, and every one of them under seven. But the absolute column, σ in raw IFI points, tells a second honest story the percentages hide. Zero-Bitter, at 5.2%, actually carries the *widest* spread of any axis — 114 IFI points — because its mean is so high; and astringency, the worst CV in the set, spreads only 85 points because its mean is low. Relative and absolute dispersion disagree, and a formulator building a blend needs both numbers in hand: the percentage tells you which axis is hardest to control, the point-spread tells you which axis's drift a palate will actually notice first. We publish both, because a number withheld is a courtesy refused.

Set that against where this starts. A wild-type pepper swings 42% batch to batch — grow ten and you have ten different peppers. Designing the whole profile to ≤7% targets a sixfold tightening from wild-type; designing heat specifically to 2.4% is a seventeenfold one. And eight axes is harder than eight times one, because they must all meet their targets *together.* If each axis is designed to conform within its tolerance band independently with probability p, the chance a whole lot achieves all eight is p^8. Design the bands to ±2σ and p = 0.9545, so p^8 = 0.689 — a third of lots would fail somewhere, an unacceptable reject rate. Design them to ±3σ and p = 0.9973, so p^8 = 0.979 — the reason we build 3σ of headroom into every target rather than aiming at the center and hoping. The independence in that calculation is itself an achievement: budget competition would ordinarily correlate the axes in the wrong direction, one rising only as another falls, which makes joint conformance far worse than p^8. Minimizing elasticity is what decouples them — what earns us the right to multiply the probabilities as if the eight numbers did not know about each other, when in truth they are all drinking from the same four wells.

The honest limits stay in view. The index reads loudness, not liking — a receptor's half-activation, not a verdict a mouth has rendered; a profile can be perfectly held and still be a chord we would revoice. The elasticity we design against assumes the plant's pools wobble within a modeled range; a greenhouse pushed past its light or temperature envelope can drift the supply beyond what the robust solution was tuned for. And astringency remains the frontier at 6.1%, the one axis whose coupling we have narrowed but not closed. We write these down because the numbers are a form of respect, and respect does not round its own errors away.

Reaching versus holding

That is the whole difference between reaching a target and holding it. Phoenix's astringency, drawing on contested carbon at the far end of eight simultaneous withdrawals, is designed to stay under seven percent — not because the balance is easy, but because, among all the ways to strike it, we deliberately chose the one that bends without breaking. Eight numbers, and not one of them designed to wander where a mouth could follow. The flux model did not find us a profile that is merely correct in some perfect imaginary fruit; it found us the profile whose eight claims on one budget stay paid even as the plant that carries them does what living things do, and varies.

This is what we mean when we say a flavor is *designed* rather than merely grown. Not eight easy numbers stacked, but the hardest balance in the library, chosen in the arithmetic before the first seed, engineered from the start to survive the ordinary variation of the plants that carry it — so that what reaches the body is steady, legible, and whole, a pleasure given on purpose and the same every time it is given. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

Chapter 17 · Stability — *the number through time and heat.*

Consistency is usually told as a story about *across* — this lot and that one, laid side by side and read against a single standard. But there is a second axis to it, quieter and just as decisive, and it runs *through:* the number a fruit carries at harvest has to survive the freezer and the dryer, the long patience of a warehouse, the shipping container that bakes on a July dock, and at the very end the pot. A specification met at the hour of release and quietly lost by the time anyone actually cooks with it was never truly met — it was a photograph of a promise, not the promise itself. So we characterize decay the way we characterize everything in this library: on purpose, with a curve, as a number rather than a hope. A flavor grown to its fullness deserves to arrive still full, and the arithmetic of keeping is how we turn that from a wish into a line a formulator can build against.



A specimen set aside and tracked — a flavor you can keep is a kind of provision, grown to its fullness and then protected until it reaches a table.

This is where consistency and nourishment turn out to be the same discipline seen twice. A flavor you can keep is a provision — a thing grown to its fullness and then *protected*, so that the care put in at the root is still wholly present when the food finally reaches a table weeks or months later. We give shelf life as a number and not a reassurance because the entire point of this work is that care should be legible: you should be able to read, on paper, exactly how much of what we grew will still be there when you need it.

The law of keeping

The chemistry of keeping has a law, and it is the same law that governs every reaction rate on Earth — Arrhenius's. The rate at which a flavor molecule falls apart climbs with temperature in a precise, exponential way:

$$k = A \cdot \exp(-E_a / RT)$$

Read every term, because each one is a physical fact about the molecule. The rate constant k sets how fast the compound disappears. The activation energy E_a is the height of the hill a molecule must climb before it can break — a tall barrier means a reluctant, slow decay. The gas constant R is 8.314 joules per mole per kelvin, and T is the absolute temperature, which is the first place a formulator goes wrong: Arrhenius does not know about Celsius. Every temperature in this chapter is first converted to kelvin, **T(K) = T(°C) + 273.15**, so a cool 20 °C storeroom is 293.15 K and a punishing 50 °C dock is 323.15 K. The pre-exponential factor A is, loosely, how often the molecule even attempts the climb; what matters is that the whole temperature dependence lives inside that exponential of $-E_a/RT$, and an exponential is a merciless thing.

Two consequences of this law are the entire working toolkit of shelf life. The first is the rule of thumb every formulator carries in the head — **Q_{10}**, the factor by which the rate changes for every ten-degree shift in temperature. It is not a separate law; it falls straight out of Arrhenius. Take the ratio of the rate at T + 10 to the rate at T and the pre-exponential factor cancels cleanly:

$$Q_{10} = k(T+10) \,/\, k(T) = \exp[\, (E_a / R) \cdot (1/T - 1/(T+10)) \,] = \exp[\, E_a \cdot 10 \,/\, (R \cdot T\cdot(T+10)) \,]$$

So Q_{10} is simply the activation energy wearing a friendlier coat. Work it both directions at room temperature, T ≈ 298 K, to feel the scale. A reaction with Q_{10} = 2 — rate doubling every ten degrees — carries an activation energy of $E_a = \ln 2 \cdot R \cdot T(T+10) / 10 \approx 0.693 \times 8.314 \times 91{,}870/10 \approx$ **52.9 kJ/mol**. A sharper one with Q_{10} = 3 carries $E_a \approx \ln 3 \times 8.314 \times 91{,}870/10 \approx$ **83.9 kJ/mol**. Most of our degradations land between those two, which is why dropping a store ten degrees roughly halves to thirds the rate, and so roughly doubles to triples the life. A cool room is the cheapest preservative we own.

Here is the honest subtlety that separates a careful formulator from a careless one, and we state it plainly because the two numbers are forever being confused. Q_{10} measures *temperature sensitivity* — how much faster decay runs when the room warms. The absolute rate k at the storage temperature measures *shelf life* — how long the thing actually lasts. They are not the same number, and a molecule can have a high Q_{10} and still keep for years. Capsaicin is the case in point: a tall activation barrier gives it a steep Q_{10}, meaning its slow decay does accelerate sharply with heat, yet its baseline k at room temperature is so minuscule that tripling a near-zero rate still leaves a near-zero rate. A high temperature sensitivity on top of an almost-nonexistent rate is still almost-nonexistent. The rugged compound is not the one that ignores temperature; it is the one whose rate, at the temperature you actually store it, is already too small to matter.

The half-life, and the seven-percent line

The second consequence is the shape of the decay itself. Most of these degradations are **first-order** — a molecule breaks on its own, independent of its neighbors, so the number breaking each second is proportional to the number still present. That proportionality has exactly one solution, a clean exponential, and the amount surviving at time *t* is:

$$c(t) = c_0 \cdot \exp(-kt)$$

Set the surviving fraction to one half and the single most useful number in shelf-life work drops out. If $c(t)/c_0 = ½$, then $-k \cdot t½ = \ln(½)$, so $k \cdot t½ = \ln 2$, and:

$$t½ = \ln 2 \,/\, k \approx 0.693 \,/\, k$$

A rate constant and a storage temperature are all it takes to state, plainly, the day the compound falls to half strength. But a formulator rarely cares about half. Half is catastrophe; the real question is *when could a mouth first tell.* And a mouth, we have measured again and again, begins to notice when an axis has drifted by about seven percent — the same seven percent that bounds our batch-to-batch scatter, which is not a coincidence but a single sensory threshold showing up in two places. So we solve the same exponential for a gentler loss. The surviving fraction at first-perceptible-loss is 0.93, so $\exp(-k \cdot t_7) = 0.93$, which gives:

$$t_7 = -\ln(0.93) \,/\, k \approx 0.0726 \,/\, k$$

Now divide the two times and watch a universal number appear. The ratio $t_7 / t½ = 0.0726 / 0.693 \approx 0.105$ has no *k* in it at all. For *any* first-order decay — rugged capsaicin, fragile terpene, every compound in the library — the first perceptible loss arrives at almost exactly one-tenth of the half-life. A product with a two-year half-life has already drifted a detectable seven percent at ten weeks. This is why we publish the seven-percent day and not the half-life day: the half-life flatters the shelf, and the tenth of it is the truth a cook would taste.

From the molecule to the number

A shelf life is only interesting because it moves the Index score, so the chain from a breaking bond to a falling number has to be drawn exactly. The Index anchors each axis to a receptor through one ratio:

$$\mathrm{IFI}(\text{candidate}) = 1{,}000 \times \mathrm{EC}_{50}(\text{reference}) \,/\, \mathrm{EC}_{50}(\text{candidate})$$

EC_{50} is the concentration at which a stimulus drives its receptor to half of full response; invert the candidate's against a fixed reference, scale by a thousand, and loudness becomes a score. Work it once with real figures to keep it concrete. Capsaicin at TRPV1 has an EC_{50} in the neighborhood of 0.84 µM for a native reference preparation; a designed extract that reaches the same half-activation at 0.20 µM

scores IFI(candidate) = 1,000 × 0.84 ⁄ 0.20 = **4,200** — the summit heat of our full-spectrum cultivar, Phoenix (illustrative EC_{50} values, chosen to land on the canonical score.)

Hold on to what *does not* change as that fruit sits in a warehouse. The EC_{50} is a property of the surviving molecule — each intact capsaicin speaks exactly as loudly on the last day as on the first. What decays is not the loudness of a molecule but the *count* of them: concentration, the c(t) of the exponential. The registered score falls because fewer molecules are left to drive the receptor, not because the survivors grew quiet. That distinction is the whole reason a stability study measures concentration and projects the score, rather than the other way around.

But concentration does not map to perceived intensity in a straight line, and the bend is where the real design lever lives. A receptor reads concentration through its dose–response curve, which for a cooperative channel like TRPV1 follows a Hill relation between fractional activation θ and concentration *c*:

> **$\theta(c) = c^n / (EC_{50}^n + c^n)$**, with Hill coefficient n ≈ 2 for capsaicin at TRPV1

The sensitivity of the perceived signal to a loss of concentration is the elasticity of that curve, and it has a beautifully simple closed form. Differentiate the logarithm: **$d \ln \theta / d \ln c = n \cdot (1 - \theta)$**. The fractional drop a mouth notices is that elasticity times the fractional concentration a molecule loses. And because (1 − θ) runs from one down to zero as a product climbs from faint to saturating, the *same* seven-percent molecular loss shows up as wildly different perceived losses depending on where the product sits.

WHERE THE PRODUCT LIVES	RECEPTOR OCCUPANCY θ	A 7% MOLECULAR LOSS IS PERCEIVED AS
Below EC_{50} ($c \approx 0.3 \cdot EC_{50}$)	≈ 0.08	≈ 13% — *amplified*
At EC_{50} ($c = EC_{50}$)	0.50	≈ 7% — *one-to-one*
Above EC_{50} ($c \approx 3 \cdot EC_{50}$)	≈ 0.90	≈ 1.5% — *masked*

(TRPV1, Hill coefficient n ≈ 2; worked from θ(c) above.)

Read that table as a formulation instruction, because it is one. A fragile axis dosed down on the steep foot of its curve is doubly cursed: it loses molecules fastest *and* each lost molecule is felt nearly double. The same fragile axis dosed up onto the saturating shoulder loses molecules at the identical chemical rate, but the flat top of the curve hides the loss — a seven-percent chemical decay reads as a scarcely-there one and a half percent. So when we must carry a delicate compound through a long shelf, we seat it high on its dose–response curve on purpose, spending a little extra material to buy a flat, forgiving response that decay cannot easily dent. To first order, in the mid-band where a product is normally designed to live, the registered score tracks concentration one-for-one, so the half-life of the molecule *is* the half-life of the number — and the curve's two ends are the levers we pull when one-for-one is not good enough.

The ladder that has to keep its shape

Nowhere is through-time stability more exacting than on a graded line, where the product is not a single number but the *spacing* between numbers. Our flagship superhots climb in deliberate steps of increasing computational design scope and axis activation, each pinned to a declared place on the heat axis, and the whole architecture only survives if every rung holds its mark not just across lots but across the months between harvest and plate.

CULTIVAR	SKU	SHU (PREDICTED)	IFI-H	AXES	LOUDNESS (×REF)
Ember	SS-0087	866K	1,570	2/8	1.57×
Blaze	SS-0091	1.72M	2,630	4/8	2.63×
Wildfire	SS-0092	1.96M	2,930	5/8	2.93×
Scorch	SS-0093	2.21M	3,220	5/8	3.22×
Firestorm	SS-0094	2.47M	3,490	6/8	3.49×
Solara	SS-0095	2.72M	3,980	6/8	3.98×
Phoenix	SS-0096	2.95M	4,200	8/8	4.20×

The rightmost column is just IFI-H ÷ 1,000 — how many times louder than the native reference each fruit speaks — and it exposes how little room a laddered line has to spare. There are two separate ways the number can fail to be the number, and both must be held. The first is *scatter across lots*, measured by the coefficient of variation, CV = σ / μ: the standard deviation of a batch divided by its mean. A wild pepper scatters at a CV near 42 percent — at the summit score of 4,200 that is a standard deviation of 0.42 × 4,200 ≈ ±1,764 points, a fruit that could read anywhere from 2,436 to 5,964 and make a mockery of the label. We hold the platform to a ceiling of **≤ 7 percent**, which at 4,200 is a standard deviation of just 0.07 × 4,200 ≈ ±294 points, a band from 3,906 to 4,494. The second way is *drift through time*, the decay of this chapter. And here is the quiet arithmetic that binds them: the lower edge of the seven-percent across-lot band, 3,906, is exactly 4,200 × 0.93 — the same floor the seven-percent perceptible-loss line sets through time. The two sevens draw from one error budget. A lot that ships already sitting near the bottom of its across-lot tolerance has spent most of its room before a single day of storage has elapsed; a lot that ships on-target has the whole seven percent left to spend slowly on the shelf.

That first kind of holding — the tight across-lot CV — is not a storage achievement at all. It is written into the enzyme. Two substitutions in the Pun1/AT3 acyltransferase — the double variant S39L+L345G — were scored by a protein language model and assessed against a genome-scale flux-balance model before any edit was made, then written in with CRISPR.

EDIT	ΔΔG (KCAL/MOL)	ROLE	WHAT IT MOVES
S39L	−2.1	N-terminal binding affinity ↑	raises the mean (peak output)
L345G	−1.8	steric relief at the active-site cleft	raises throughput and folding consistency
C175S†	−0.9	disulfide-misfold prevention	structural annotation — not in shipped construct

† C175S is a structural candidate from our modelling, retained in our design record; it was not included in the shipped double variant.

Our flux-balance model predicts control over capsaicin flux sits upstream in the phenylpropanoid pathway — at the 4CL and HCT/CCoAOMT steps — not at the terminal Pun1/AT3 synthase. Pun1 labels the product; it does not govern the rate. The double variant S39L+L345G is a throughput and consistency play: S39L raises binding affinity at the N-terminal, and L345G relieves steric pressure at the active-site cleft, lifting both peak throughput and the fraction of enzyme molecules that fold into the same active conformation. It is why our model predicts Phoenix reaches approximately 2.95 million SHU as a design target. The across-lot reproducibility we rely on has two homes that must not be mistaken for each other: the *synthesis-side* steadiness that comes from an enzyme designed to fold the same way every time, which sets the across-lot CV at the moment of making; and the *storage-side* steadiness of the Arrhenius law, which governs whether that tight-made number survives the journey. A chapter on stability is about the second — but the two are halves of one promise, and a product is only as consistent as the weaker of them.

The accelerated study, worked end to end

We cannot wait two years to learn whether a product lasts two years, so we borrow the Arrhenius law's own leverage: hold samples warm, measure the fast decay there, and project it back down to the cool temperature of real storage. This is the single most important method in keeping, and it is worth walking a whole example through the arithmetic, because a formulator who understands the extrapolation will trust the number for the right reasons.



Phoenix under a held, elevated temperature in the controlled chamber — the accelerated study warms a flavor on purpose, so the slow truth of its shelf can be read back in weeks instead of years.

Take a delicate sweetness glycoside, held at an aggressive 50 °C (323.15 K). We watch its concentration fall, confirm the decay is first-order by checking that the logarithm of surviving fraction falls on a straight line against time, and read a half-life of 40 days. That fixes the rate at the hot temperature: k(50 °C) = 0.693 ⁄ 40 = 0.01733 per day. Now assume a Q_{10} of 3 for this compound — measured, in practice, by running a second hold at another temperature and comparing — and step the ladder back down toward the storeroom, dividing the rate by three for every ten-degree descent.

TEMPERATURE	K (PER DAY)	T½	T_7 (FIRST PERCEPTIBLE LOSS)
50 °C (accelerated)	0.01733	40 days	4.2 days
40 °C	0.005776	120 days	12.6 days
30 °C	0.001925	360 days	37.7 days
20 °C (storage)	0.000642	1,080 days (≈ 2.96 yr)	113 days

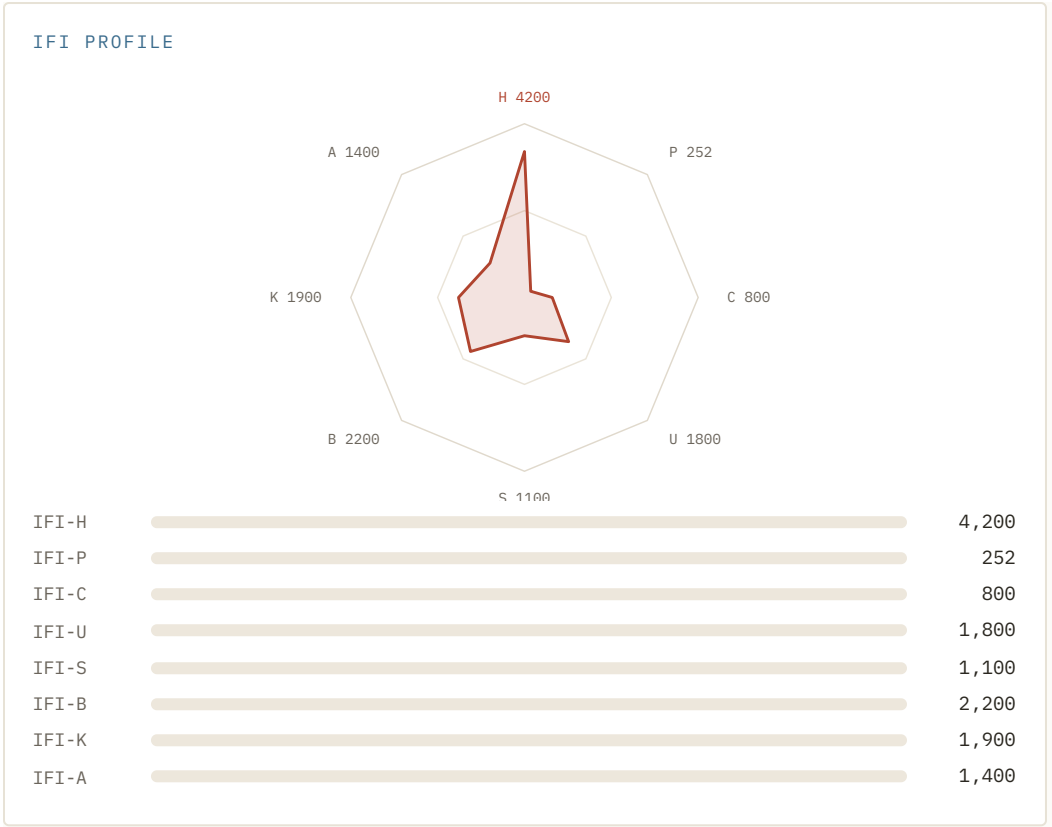
(First-order model, $Q_{10} = 3$, anchored at the measured 50 °C half-life; $t_7 = 0.0726/k$ throughout.)

The warm experiment, run in weeks, has told us the cold truth we would otherwise have waited years to measure: at a 20 °C store this glycoside holds a near-three-year half-life, and — the number we actually publish — a mouth could first detect its fading at about 113 days, a little under four months. Notice that the t_7 column is always almost exactly a tenth of the t½ column beside it, the universal first-order ratio made visible row by row. And notice how steeply the heat punishes the fragile thing: the identical molecule that keeps its first-perceptible edge for four months at room temperature loses it in four days on a 50 °C dock. That single comparison is the entire argument for a cool chain, stated in the compound's own numbers.

The method has an assumption, and we name it because its failure is the most common way a projected shelf life turns out wrong. The back-projection assumes *one* decay mechanism, with *one* activation energy, governs the compound across the whole temperature range — that the Arrhenius line is straight from 50 °C down to 20 °C. When that holds, the extrapolation is honest. It stops holding when a different chemistry takes over at a different temperature: a matrix that is glassy and protective when cold but softens and mobilizes reactants when warm, a moisture that condenses only below a dew point, a melt or a phase change that opens a fast route the cold sample never sees. Then the hot experiment is measuring a different reaction than the cold shelf will run, and the projection flatters or maligns the product. So we never extrapolate across a phase change we have not checked, we bracket the storage temperature with at least two accelerated points to confirm the line is straight, and we say, on the label, the temperature the number assumes.

A full-spectrum fruit ages unevenly

A single-axis product has one shelf life. A full-spectrum one is a *portfolio* of molecules with wildly different ruggedness, aging at different rates under the same roof, and its true shelf life is set by whichever axis fades first. Phoenix registers on all eight axes at once, and that is exactly what makes its keeping a harder, more honest problem than its heat alone would suggest.



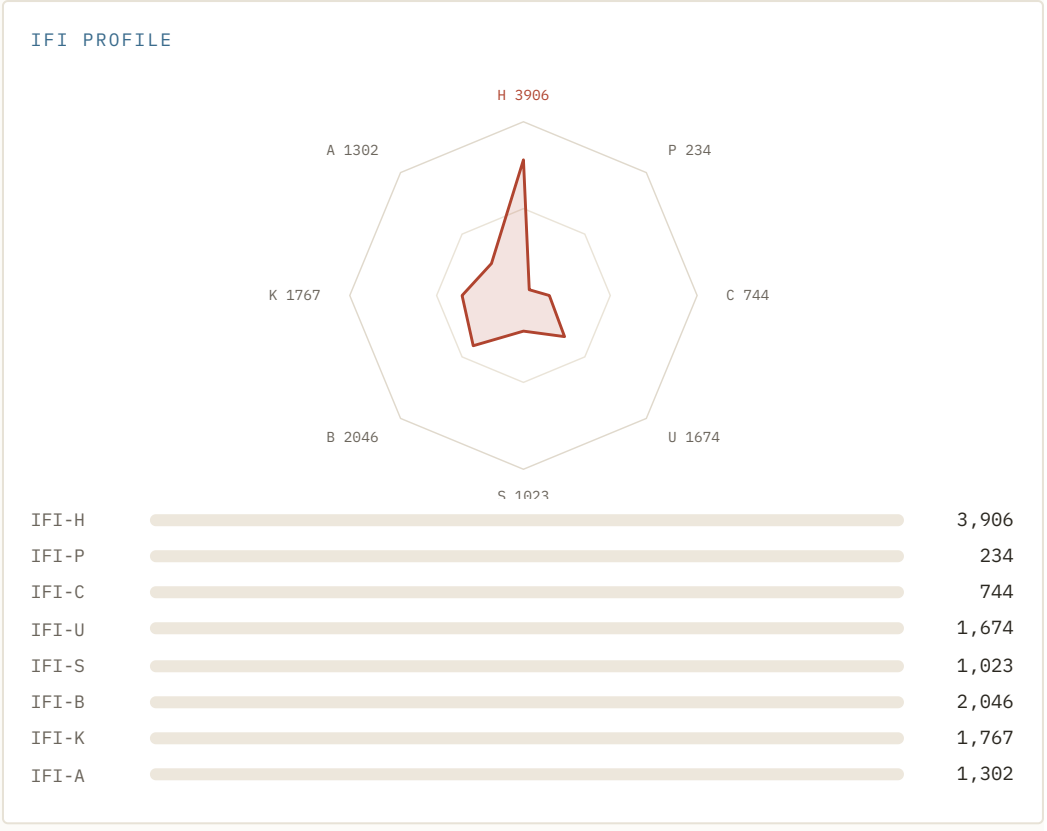
Each of those numbers rides on a different class of molecule, and the chemistry of each class dictates how it keeps. The ruggedness is not arbitrary; it is readable straight off the bonds.

AXIS	RECEPTOR	CARRIER CHEMISTRY	WHY IT KEEPS — OR DOESN'T
H · Heat 4,200	TRPV1	capsaicinoid amide	very rugged — the amide bond has a tall hydrolysis barrier; heat barely moves it
B · Zero-Bitter 2,200	TAS2R38	phenolics / tannins	rugged — polyphenols resist dry storage; slow oxidation only
A · Astringency 1,400	salivary PRPs	condensed tannins	rugged — large, stable polyphenol scaffolds
K · Kokumi 1,900	CaSR	γ-glutamyl peptides	moderate — peptide bonds hold cool, hydrolyze under prolonged warmth
U · Umami 1,800	mGluR1	glutamate	moderate — small and water-soluble, fairly stable
S · Sweetness 1,100	TAS1R2 / R3	steviol glycosides	fragile — glycosidic bonds hydrolyze under heat and acid
C · Cool-Burn 800	TRPM8	volatile terpenoids	fragile — low molecular weight; lost to evaporation and ring oxidation
P · Pungency 252	TRPA1	reactive electrophiles	most fragile — the reactivity that makes it pungent also makes it decay

The deepest point in that table is the last row, and it is a piece of real receptor mechanism, not a metaphor. TRPA1 — the pungency channel — is not opened by a molecule *fitting* a pocket the way capsaicin fits TRPV1; it is opened by a molecule *reacting* with it, an electrophile forming a covalent bond to cysteine residues on the channel's inner face. Allicin from garlic, the isothiocyanate of

mustard and wasabi, cinnamaldehyde — they are pungent precisely because they are chemically aggressive. And a molecule aggressive enough to covalently modify a receptor is aggressive enough to react with water, with oxygen, with itself, long before it reaches a tongue. The very property that earns the pungency score is the property that spends it. This is the honest shape of flavor stability: loudness and permanence are sometimes at war at the level of the bond, and no amount of good process can broker a peace the chemistry refuses.

So Phoenix's declared shelf life is not governed by its signature fire. Capsaicin's amide will outlast the carton. It is governed by the quietest numbers on the profile — the cool terpenoid that evaporates and the pungent electrophile that self-destructs — and those fragile axes are the ones we protect: dosed high on their dose–response shoulders where decay is masked, sealed against oxygen and light, added late in a formulation, and spent while young. The specification we actually guarantee is a floor under every axis at once: at the end of the declared shelf life, no axis has fallen below 93 percent of its label, the seven-percent line applied across the board.



That second profile is not a decay snapshot — the axes do not reach it together, since each fades at its own rate. It is the *contract*: every value is its label times 0.93, and we set the shelf life by the first axis to approach its own floor, so that when the fragile cool and pungent numbers arrive at theirs, every rugged axis is still comfortably above. The whole profile is guaranteed at or above the floor block for the whole declared life. A full-spectrum flavor is only as keepable as its weakest axis, and naming that weakest axis — out loud, in the number — is the difference between a shelf life and a guess.

The honest limits

First-order decay is a model, and a good one, but it is not a law of nature the way Arrhenius is. Some losses are zero-order, running at a constant rate until a reactant simply runs out — the straight-line fade of a compound limited by a surface or a catalyst rather than by its own abundance. Some are autocatalytic, their own breakdown products accelerating the next round, so the curve is flat and reassuring for months and then collapses; a shelf life read off the flat early part of such a curve is a dangerous overestimate. Some show an induction period, a sigmoidal lag while a protective antioxidant is consumed, with the real decay beginning only after the guard is spent. We fit the model the data actually support — checking that the log-survival line is genuinely straight before we trust a half-life — rather than forcing every compound into the tidy exponential because the exponential is convenient.

The Arrhenius projection carries its own caveats beyond the phase-change failure already named. Moisture, oxygen, light, and pH are co-variables that ride underneath every rate constant we quote; a *k* measured in a sealed dark vial is not the *k* of an open jar on a sunlit counter, and we declare the conditions a number assumes rather than letting it float free. And the seven-percent line is a sensory threshold, not a physical cliff — decay is continuous, the molecule does not know the number 0.93, and a formulator who needs tighter than seven percent (a laddered line whose rungs crowd, a reference standard that anchors a whole scale) simply sets a stricter floor and accepts the shorter shelf that buys. The number is a decision made legible, not a fact of the compound; what we promise is that the decision is shown, and shown as arithmetic anyone can check.



Reading the number back from a held sample on the microscope stage — a shelf life is not an assurance offered but a measurement taken, again and again, until the curve is known.

The number, kept all the way to the table

Stand back from the arithmetic and the whole chapter is one idea. A number that holds across lots is a promise to a stranger — that the fruit in their hand is the fruit on the label. A number that holds across *time* is the same promise, carried further: that the fruit is still itself after the freezer and the dock and the long wait, that the warmth or the sweetness or the savory depth we grew into it at the root is undiminished when it finally reaches a mouth. We measure decay, and publish the day a mouth could first tell, and set a floor under

every axis, because a flavor is a provision and a provision you cannot count on is no provision at all. The care is in the growing; the keeping is how that care survives the distance between the field and the table. We give it as a number, because a number is the most honest form a promise can take — one anyone can hold us to, on the last day of the shelf as surely as the first. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.



PART V OF V

The Spec at the Table

the certificate of analysis · what a stable input unlocks · consistency, built in rather than inspected in

Chapter 18 · The certificate of analysis — *what ships with every lot.*

Trust is a lovely thing between people and a poor thing to build a product on. So no lot leaves us on its reputation; each one leaves with its paperwork — a certificate of analysis that states, for that specific lot, what was read, how, against which revision of which reference, and how tightly it held. A formulator building a product should never have to take our word for a number, and with a certificate riding on the lot, they never do. The word is on the sheet, and the sheet travels with the fruit.



The handoff — a lot leaves not on its reputation but with its paperwork, the whole quality system compressed to a page.

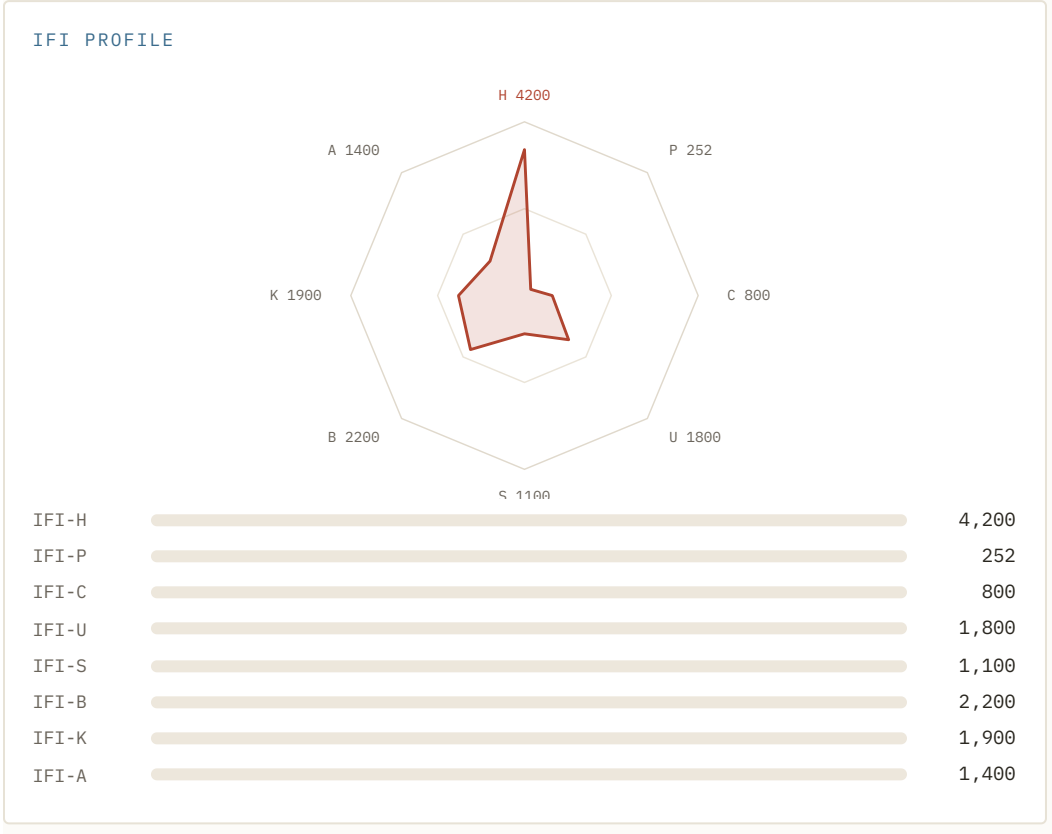
The instinct underneath it is the oldest one we cook from. When you set a dish in front of someone you love, you tell them exactly what is in it — not to impress them, but because feeding a person well means letting them see what they are taking into their body. A certificate of analysis is that same honesty rewritten in the grammar of a laboratory: here is what this is, here is how we know, here is how sure we are. Nothing on the page is decoration. Every line is an answer to a question a careful person would be right to ask before they build their own work on top of ours.

What the certificate asserts

A certificate is not a summary of the harvest in general. It is a set of claims about *one lot* — one defined quantity of material, grown under one logged set of conditions, read on one set of panels, anchored to one named revision of one reference. The lot number is the proper noun that makes every other number on the page specific rather than typical. Read it as a sentence and it says: *this material, read this way, landed here, this tightly.*

```
CERTIFICATE OF ANALYSIS                    Scoville Splice
cultivar    Phoenix  ·  SS-0096             lot    0096-R244
reference   IFI master standard set, rev. 2026.1
panels      TRPV1 · mGluR1 · TAS2R38 · CaSR · TRPM8
            TRPA1 · TAS1R2·R3 · salivary PRP
spec        every axis CV ≤ 7.0 %          result  PASS
```

The lot on this sheet is Phoenix — SS-0096 — the one cultivar we designed to register on all eight axes at once, which is why its certificate is the fullest one we print. Its designed profile is the thing the lot is read *against*:



Eight numbers, eight receptors, one fruit. The certificate's job is to establish — through dose–response panel readings — that the material in the drum actually speaks those eight numbers, within a width we have agreed is small enough to design around. Everything that follows — the dose–response reads, the index arithmetic, the coefficient of variation, the reference revision — exists to turn that profile from a design intention into an audited fact about a specific quantity of fruit now sitting on someone's receiving dock.

From receptor to number — the measurement behind a score

No axis on the certificate is a direct reading of a compound. Each one is a reading of a *response* — how loudly a receptor answers the material — and that distinction is the whole reason the index is worth anchoring to. A chemical assay can tell you how much capsaicin a fruit holds. It cannot tell you how loudly that capsaicin will be heard, because loudness is set at the receptor, not in the beaker.

Take the heat axis, IFI-H, read on TRPV1. Capsaicin is a non-selective cation channel agonist: it binds a pocket on TRPV1 — the same channel opened by temperatures above 43 °C and by protons — and lowers the channel's thermal threshold so the nerve reports a fire that is not there. Step by step, the measurement runs: expose a panel of cells carrying TRPV1 to a graded series of concentrations; at each concentration a fraction of the channels open and admit Ca^{2+} and Na^{+}; the integrated response rises from zero, through a half-maximal point, up toward saturation. Plot response against the logarithm of concentration and you get the familiar sigmoid. The concentration at the inflection — where the receptor gives exactly half of everything it has to give — is the EC_{50}. That single number is the whole dose–response compressed to its most honest point, and it is what the certificate's arithmetic is built from.



Reading the lot on the stage — each axis is a receptor's answer, fit to a half-maximal point, not a compound counted in a beaker.

Each of the eight panels does the same thing to a different receptor, and the mechanism changes the character of what is being read:

- **IFI-U · Umami** reads mGluR1, a metabotropic glutamate receptor; the response is a G-protein cascade, so the dose–response is a downstream signal rather than a raw ion current, but the half-maximal point is defined identically.
- **IFI-B · Zero-Bitter** reads TAS2R38, a bitter GPCR, and here the axis is an *attenuation*: we are not measuring how loudly a bitter agonist rings the receptor but how much the material *quiets* it, a 20–65 % reduction across our bitter-bearing hosts. Its "EC_{50}" is really a half-attenuation concentration — the dose at which half the achievable quieting is reached.
- **IFI-K · Kokumi** reads CaSR, the calcium-sensing receptor, measured as potentiation — the continuity and body a kokumi agonist lends to other tastes.
- **IFI-C · Cool-Burn** reads TRPM8, the menthol and cold channel.
- **IFI-P · Pungency** reads TRPA1, gated not by simple binding but by covalent modification of reactive cysteines — the sharp electrophilic bite of allicin, wasabi, cinnamon.
- **IFI-S · Sweetness** reads the TAS1R2·TAS1R3 heterodimer.
- **IFI-A · Astringency** is the honest outlier: it is not a receptor current at all but a tactile phenomenon — tannins precipitating the salivary proline-rich proteins that lubricate the mouth. Its panel reads a half-maximal *precipitation*, and the index treats that turbidity midpoint as an EC_{50}-analog. We say so on the method sheet rather than pretending the tongue has an astringency channel it does not have.

Two of the eight, then, are not agonist potencies in the strict sense — zero-bitter is an inhibition, astringency is a precipitation. The certificate carries them on the same index scale because a formulator wants one grammar for all eight, but the batch record behind the page names each panel's true readout, so no one is misled about what the number is underneath.

The index, worked

Eight raw EC_{50} values in eight different units would be unreadable. The index collapses them to a readable scale with a single formula, and the formula is worth deriving rather than quoting, because its

shape is the reason the whole library can be read on one page:

```
IFI(candidate) = 1,000 × EC₅₀(reference) ÷ EC₅₀(candidate)
```

The reference standard is the anchor. By construction it scores 1,000 — set EC_{50}(candidate) equal to EC_{50}(reference) and the ratio is one, times a thousand. The IFI is a stated convention: each axis defines its own reference compound, and that compound's EC_{50} *defines the unit* on that axis. A higher score means a *lower* EC_{50} — less material needed to half-activate — which is exactly what "louder" should mean.

Work it on this lot's heat axis. The certificate's target is IFI-H 4,200. Rearranged, that fixes the candidate's potency relative to the reference:

```
4,200 = 1,000 × EC₅₀(ref) ÷ EC₅₀(cand) → EC₅₀(cand) = EC₅₀(ref) × 1,000 ÷ 4,200 = 0.238 × EC₅₀(ref)
```

Phoenix's heat half-activates TRPV1 at **23.8 % of the reference concentration** — it is 4.2× louder than the anchor. The arithmetic runs the other way just as cleanly, which is how the panel actually produces a score: measure the lot's bitter EC_{50}, find it sits at 0.460 of the reference, and the index returns 1,000 ÷ 0.460 = **2,176** — the exact IFI-B lot mean this certificate reports.

The scale is logarithmic in a way that rewards a second look. Because the index is a ratio of two EC_{50} values read on the same plate, it equals a thousand times a power of ten:

```
IFI = 1,000 × 10^Δ, where Δ = log₁₀EC₅₀(ref) − log₁₀EC₅₀(cand) = the leftward shift, in decades, of the candidate's dose–response curve
```

For the heat axis, $\Delta = \log_{10}(4.2) = 0.623$ — Phoenix's heat curve sits a little over half a decade to the *left* of the reference. The index number is just that shift, re-expressed as a score anyone can read without a calculator. And the scale runs both ways: this lot's cool-burn axis is IFI-C 800, which means 800 = 1,000 × 10^Δ, so $\Delta = \log_{10}(0.8) = -0.097$ — a slight shift to the *right*, a cool note measurably quieter than the reference anchor. A score below 1,000 is not a failure; it is a faithful report that the material is gentler than the standard on that axis. The scorecard caps each axis at a normalization maximum of 5,000, so the radar never runs off the page even when a designed heat reaches well past the reference.

AXIS	LOT IFI	EC_{50} RELATIVE TO REFERENCE	CURVE SHIFT Δ (DECADES)
IFI-H · Heat	4,214	0.237 ×	+0.625
IFI-B · Zero-Bitter	2,176	0.460 ×	+0.338
IFI-K · Kokumi	1,912	0.523 ×	+0.281
IFI-U · Umami	1,788	0.559 ×	+0.252
IFI-A · Astringency	1,409	0.710 ×	+0.149

Every column is derived from the one before it by the formula and nothing else. That closure — raw read to EC_{50} to ratio to score, each step reversible — is what lets two laboratories that have never met agree on exactly what a number means.

Where Scoville and the index disagree

A formulator who has lived in the pepper world will ask, reasonably, why the heat line is not simply a Scoville number. The honest answer is the strongest argument for the whole index, and the flagship comparison shows it plainly. Scoville Heat Units count capsaicinoid *mass*; IFI-H reports receptor *potency*. They are related — more capsaicinoid generally rings TRPV1 harder — but they are not the same measurement, and the difference is not noise.

The SHU values below are our model's predictions for each cultivar; the IFI-H scores are the design targets on the heat axis.

CULTIVAR	SKU	PREDICTED SHU	IFI-H	PREDICTED SHU PER IFI-H POINT
Ember	SS-0087	866,000	1,570	552
Cinder	SS-0088	1,050,000	2,290	459
Kindle	SS-0089	1,250,000	2,000	625
Blaze	SS-0091	1,720,000	2,630	654
Wildfire	SS-0092	1,960,000	2,930	669
Scorch	SS-0093	2,210,000	3,220	686
Firestorm	SS-0094	2,470,000	3,490	708
Solara	SS-0095	2,720,000	3,980	683
Phoenix	SS-0096	2,937,948	4,200	700

Run your eye down the last column. The conversion from chemical heat to receptor-measured heat is not a constant: it ranges from 459 predicted SHU per index point up to 708 — a spread of better than fifty percent. Worse for anyone tempted to use Scoville as a potency proxy, the two measures actually *invert* across one pair. Kindle (SS-0089) carries a predicted 1,250,000 SHU — two hundred thousand more than Cinder (SS-0088) at 1,050,000 — and yet is designed to read *lower* on TRPV1, 2,000 against 2,290. More capsaicinoid mass, less loudness. The reason is in the chemistry we tune at the Pun1 acyltransferase: the ratio of capsaicin to dihydrocapsaicin and the minor capsaicinoids shifts with the length and saturation of the acyl chain, and those congeners do not all ring the channel equally. A certificate that quoted only SHU would therefore misstate receptor loudness by up to roughly 1.5× and, in at least one case, get the *ranking* backwards. That is why the heat line on our certificate is an EC_{50}-anchored index score, fit from a dose–response, not a mass number borrowed from a chromatograph.

The coefficient of variation — the honest width

A mean with no width is a wish. The line that turns this sheet from a brochure into a specification is the coefficient of variation — the scatter inside the lot, stated as a fraction of the mean, checked against a ceiling of 7.0 % and stamped, per axis, pass or fail.

AXIS	TARGET IFI	LOT MEAN	LOT CV	IN SPEC
IFI-H · Heat	4,200	4,214	2.4 %	✓
IFI-U · Umami	1,800	1,788	3.8 %	✓
IFI-B · Zero-Bitter	2,200	2,176	5.2 %	✓
IFI-K · Kokumi	1,900	1,912	5.6 %	✓
IFI-A · Astringency	1,400	1,409	6.1 %	✓

Five of the eight shown; the full panel rides with the lot. The coefficient of variation is simply the standard deviation divided by the mean, CV = σ ÷ μ, expressed as a percent — which means the percentage on the page can be read back into the actual spread of scores. Invert it, σ = CV × μ, and the heat axis carries a standard deviation of 0.024 × 4,214 ≈ **101 index points**; the astringency axis, looser at 6.1 %, carries 0.061 × 1,409 ≈ **86 points**. Put a 95 % interval around each (mean ± 1.96 σ) and you have the band a formulator can actually plan inside:

- **Heat** — 4,214 ± 198 → roughly [4,016, 4,412]
- **Umami** — 1,788 ± 133 → [1,655, 1,921]
- **Zero-Bitter** — 2,176 ± 222 → [1,954, 2,398]
- **Kokumi** — 1,912 ± 210 → [1,702, 2,122]
- **Astringency** — 1,409 ± 168 → [1,241, 1,577]

Those are tight bands, and tightness here is not luck — it is the molecular work made visible. The two AI-designed substitutions in Pun1 (double variant S39L+L345G) that give Phoenix its heat are also what give the heat its *reproducibility*. S39L strengthens substrate binding at the N-terminus (ΔΔG −2.1 kcal/mol); L345G relieves a steric clash in the active-site cleft (−1.8). Summed, that is a −3.9 kcal/mol design margin, and a more stable, more uniformly folded enzyme population is designed to titer capsaicin far more evenly from plant to plant and batch to batch. (Our structural modelling also identified C175S — a disulfide-misfold prevention substitution with ΔΔG −0.9 — as a beneficial structural annotation; it is documented in our design record but was not included in the shipped double variant.) The arithmetic of that stabilization is dramatic even as an order-of-magnitude intuition: a −3.9 kcal/mol shift moves a folding equilibrium by roughly e^(3.9 ÷ 0.59) ≈ seven-hundred-fold toward the folded, catalytically competent state. That is the mechanistic reason a wild-type heat trait that scatters at a **42 % CV** can be held, in the engineered line, to **≤ 7 %**.

It is worth seeing what that six-fold tightening is really worth, because variance, not standard deviation, is the currency of a formulation's error budget — and variance scales as the square of CV. Holding 7 % instead of 42 % is not six times better; it is (42 ÷ 7)² = **36 times less variance** carried into every downstream mix. A wild-type lot at 42 % CV around a 4,200 target would scatter across a 95 % band of roughly [743, 7,657] — the heat could nearly be a tenth of the label or nearly double it. No one can formulate against that. The CV line is the difference between a number you can build on and a number you can only hope on.

One subtlety keeps the CV honest, and it belongs on the method sheet: the index is a ratio of two EC_{50} values, so the scatter in the *reference* fit propagates into every score. By the standard rule for a quotient, relative variances add, and the CV of a score is the quadrature sum of the two fits that made it — CV(score) = $\sqrt{(CV^2_{ref} + CV^2_{cand})}$. To see how the pieces combine, suppose the reference contributes about 1.5 %; then the material's own scatter on the heat axis is $\sqrt{(2.4^2 - 1.5^2)} \approx 1.9$ %. The lesson is not the illustrative split but the structure it reveals: a loose reference would quietly inflate every CV on the page, which is why the anchor itself is held to a tight fit and why the exact revision of that anchor is printed where no one can miss it.

The reference standard and its revision

`reference IFI master standard set, rev. 2026.1`. Of all the lines on the certificate, this is the one a casual reader skips and a careful one reads first. The index is a ratio against a standard, so a score is only as fixed as the standard it was divided by. Re-fit the reference — a new batch of standard material, a recalibrated panel, a tighter dose–response model — and every score in the library shifts underneath you unless the move is tracked. We track it by versioning the reference like code and stamping the version onto every lot it anchored.

This is what makes two lots comparable, and it is the first thing a formulator should check when deciding whether this drum is interchangeable with the last one they qualified. Two certificates against **rev. 2026.1** are speaking the same language; a lot anchored to an earlier revision has to be reconciled through the published conversion before its numbers can be laid beside these. The revision line is, in effect, the version of the dictionary the lot's words were defined in. Printing it is not bureaucratic caution. It is the one piece of information that keeps a 4,214 read this spring and a 4,214 read next spring meaning the same thing on the tongue.

Reading the certificate as a formulator

For the person on the receiving dock, the certificate is not a courtesy to file but a working input to their own quality system, and it answers three questions in order.

Is it what I bought? The per-axis means, checked against the material's incoming assay, confirm that the fruit in the drum matches the paper on the drum before a gram of it enters a batch. A good receiving lab does not trust the certificate; it reproduces it, on its own panel, and agrees. The whole point of writing numbers down this carefully is that agreement becomes possible between two labs who have never met.

How much will it move on me? The lot CV goes straight into the formulator's own variance budget. If a sauce is specified at a heat the eater should feel within a just-noticeable band, the formulator stacks the lot's ±198-point heat interval on top of their own process scatter — mixing, fill, storage — and sizes the recipe so the *stack* still lands in spec. A lot that arrives already holding ≤ 7 % leaves almost the entire budget for their own steps; a loose lot would spend it before they began.

Can I line it up with last time? The reference revision answers comparability; the lot number answers traceability. Together they let a formulator treat a new drum as a continuation of a qualified history rather than a fresh unknown — the single most expensive thing to re-establish in a regulated kitchen.

The archive behind the page

A certificate points backward, and that is what separates a real document from a printed reassurance. Behind the single page stands a batch record — the pinned setpoints of the room that cycle, the recovery of the extraction, the revision of every reference the panels ran against, the raw dose–response reads the EC_{50} was fit from — and behind *that* stands a retained sample of the lot itself, archived cold, so that a question raised a year downstream can be answered against the very material that shipped rather than against memory.



The retained sample, set aside and archived cold — so a question raised a year downstream is answered against the fruit that shipped, not against memory.

The certificate is the index to that archive. Its lot number — `0096-R244` — is the thread you pull to unspool the whole history of how the fruit came to carry its numbers: which chamber, which harvest window, which panel on which day, fit against which revision. A reputation cannot be pulled apart and checked. A chain of records can, and that is the only kind of claim we are willing to put a family's name on. The retained sample is the quiet heart of it — the same impulse, really, as keeping back a jar from a good batch so that next year you can taste whether you held the line. We keep the jar. We keep it cold. And we label it so precisely that anyone can find it.

Failure modes and honest limits

A specification earns trust by being as clear about what it cannot promise as about what it can.

A certificate certifies a *lot*, not a plant and not a bite. Inside the ≤ 7 % band there is real, designed variation from fruit to fruit; the CV is the honest width of exactly that scatter, not its abolition. The sheet tells you the distribution the material was drawn from — it does not tell you that the single gram you pull sits at the mean. That is what the interval is for, and why a careful formulator plans against the band, not the midpoint.

A score is a measurement, and measurements can be precise without being true. The CV reports *precision* — how tightly the reads cluster. *Accuracy* — whether the cluster sits where the receptor's real-world response does — is carried by the reference standard and its revision, which is why so much of this chapter has circled back to that one line. A lot can pass its CV spec against a drifted reference and still mislead; the versioning is the guard against precisely that, and it only works if the revision is checked, not assumed.

And two of the eight axes wear the index scale a little loosely by their nature. Zero-bitter is an attenuation, not an agonist potency, so its "score" is a half-attenuation concentration against a bitter background — change the background and the attenuation a formulator actually experiences can differ from the panel's. Astringency is a precipitation, not a receptor current at all, so its EC_{50}-analog is a turbidity midpoint that tracks the tactile sensation well but is not the sensation itself. We hold both on the index for the sake of one readable grammar across eight dimensions, and we name both plainly on the method sheet so no one mistakes the convenience of a shared scale for a claim that bitterness and astringency are read the way heat is.

One further limit is worth naming plainly: each axis is anchored to its own defined reference compound, not to a single common standard. A score of 4,200 on the heat axis and a score of 252 on the pungency axis each mean "this distance from the reference on that receptor" — they are not interchangeable across axes, and the two numbers cannot be read as a ratio of intensities between heat and pungency.

None of these limits is a weakness in the document. They are the document doing its job — stating the edges of its own authority, so that trust lands only where it has been earned.

These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

A certificate is not marketing. It is a chain of custody for a promise — the proof, attached to the thing itself, that the number you are about to build on is real, that we read it the way we said, against a standard we named, and that we kept the fruit to prove it. The word is on the sheet, and the sheet travels with the fruit, because that is how you feed people you cannot see: you tell them the truth, in numbers, and you let them check.

Chapter 19 · Reformulation — *what a stable input unlocks.*

For the bench where products are actually made, everything about precision pays out in a single word: *reformulation.* Pulling the sugar out of a drink, the sodium out of a sauce, and leaving the pleasure exactly where a person expects to find it is not an act of simple removal. It is subtraction *plus a precise replacement* — and the replacement is only ever as good as its precision. You cannot dose a sweetness you cannot trust. Lean on an input that wanders and some cans come out flat and some cloying, the panel splits, and a whole program dies months before launch over a fault that was never in the idea, only in the scatter. A held coefficient of variation is the quiet thing that makes the replacement land every time: a known intensity, delivered to a known target, can after can after can, lot after lot after lot.



Sugarcane — reformulation is not removal but precise replacement, and precise replacement is only possible when the input lands on its number.

We write this chapter the way our family cooks: reverent about the result, exacting about the method, certain the two are one act. What follows turns "reduce the sugar and keep the pleasure" from a wish into a specification — a number a formulator can design to on purpose, at scale, and clear on a capability study. The whole of it rests on one fact we will earn in full before the end: that a designed flavor input can be made to arrive on its target with a batch-to-batch scatter at or below seven percent, where the plant that grew wild delivered forty-two.

Subtraction is the easy half

Anyone can take a thing out. Boil a broth without salt; brew a cola without sugar. The removal is trivial and the result is a ruin — flat, hollow, edged with a bitterness that was always there and only ever hidden. The sugar in a beverage was never only sweetness; it was body, and continuity across the sip, and a soft blanket laid over the sharp notes underneath. The salt in a sauce was never only salt; it was savor carried forward, roundness, and a lid held down on the bitterness of everything else in the pot. Take either one out and you do not open one hole. You open three.

So reformulation is a filling problem, and a filling problem is a *dosing* problem. Each hole has a size, and the thing you pour in has to match that size — not approximately, not on average, but in this lot and the next and the ten thousand after. The rest of this chapter is about how we measure the size of a hole, how we know how much to pour, and why the pouring can be trusted. The reverence is real, but it is reverence with a decimal point: we love the full broth, and we love it because we can say exactly how full.

A ratio has no units — why one scale can hold eight axes

The measuring begins with a receptor. Every taste a person can name is, underneath, a molecule settling into a protein on the surface of the tongue — a lock and a key, repeated a billion times a second — and the only honest way to ask *how much* of a taste is to ask how hard that protein is being driven. The half-maximal effective concentration, EC_{50}, is the amount of a stimulus that drives its receptor to exactly half of full response. A small EC_{50} means a loud molecule: it reaches half-activation while barely present. A large EC_{50} means a quiet one. The Iyengar Flavor Index turns that loudness into a score by inverting the EC_{50} against a fixed reference and scaling by a thousand.

$$\mathrm{IFI}(\text{candidate}) = 1{,}000 \times EC_{50}(\text{reference}) \div EC_{50}(\text{candidate})$$

Two things fall out of that line immediately, and both matter for reformulation. The first is a fixed point: set the candidate equal to the reference and the equation reads $1{,}000 \times EC_{50}(\text{ref}) \div EC_{50}(\text{ref}) = 1{,}000$. The reference always scores exactly one thousand. It is the sea level of the axis, the place every reading is measured from. The second is a conversion we will use all chapter long. Rearrange, and the potency ratio is simply the score over a thousand:

$$EC_{50}(\text{reference}) \div EC_{50}(\text{candidate}) = \mathrm{IFI} \div 1{,}000$$

A sweetness input dosed to a target of IFI-S 3,000, then, is a candidate whose potency ratio is 3.0 — it reaches half-activation of the sweet receptor at one-third the concentration the reference needs. The number is not decoration. It is a recipe for how much to add.

Notice what the formula is made of: a concentration divided by a concentration. The units cancel. An IFI value is a pure, dimensionless ratio, and that is the deepest reason eight axes measured in eight incompatible physical units — sweetness in millimoles of glycoside, heat in Scoville units, umami in parts per million of glutamate — can be expressed in one common format. Each axis carries its own

stated reference constant — its own convention for what scores exactly one thousand — so an IFI-S figure and an IFI-H figure each measure potency *within their own axis* and should not be read as rankings on a single unified scale. You cannot compare a millimole to a Scoville. What you can do is compare any candidate to its own axis reference, and that makes formulation math tractable across all eight at once. The same dimensionless grace is why the coefficient of variation, which is a standard deviation divided by a mean, is the right currency for precision: it too sheds its units, so a four-percent scatter on a sweetness and a four-percent scatter on a heat mean the same thing — the same tightness around the target — even though the targets live in different worlds. Reformulation is a conversation across axes, and it can only be held in a language with no units. The index is that language.

How much to add — loudness, dose, and the Hill curve

A score tells you how loud an input is; the formulator still has to decide how much to pour. For that we need the shape of the response, not just its midpoint. A receptor's activation as a function of dose follows a sigmoid — the Hill curve — which for a stimulus at concentration C reads

$$\theta(C) = C^n \div (EC_{50}^n + C^n)$$

where θ is the fraction of full response and n is the Hill coefficient that sets the steepness. At $C = EC_{50}$ the two terms in the denominator are equal and θ = ½, which is the definition made visible. Below the EC_{50} the curve climbs steeply; above it, it flattens toward saturation, which is why over-dosing a sweetener buys less and less sweetness per gram and eventually only a syrupy, cloying tail.

Here is the result that makes the index operational. Suppose a formulator wants to hit some perceptual target θ — say eighty percent of full sweet response, the fullness a sugared drink carries. Solving the Hill equation for the dose gives $C = EC_{50} \times [\theta \div (1 - \theta)]^{(1/n)}$. The reference, with its own EC_{50}, would need $C(ref) = EC_{50}(ref) \times [\theta \div (1-\theta)]^{(1/n)}$. The engineered candidate, with $EC_{50}(cand) = EC_{50}(ref) \times 1{,}000 \div IFI$, needs that same bracketed factor times its *own* EC_{50}. Divide one by the other and the whole bracket — the part that depends on the target level and the curve shape — cancels:

$$C(candidate) \div C(reference) = EC_{50}(candidate) \div EC_{50}(reference) = 1{,}000 \div IFI$$

The dose ratio does not depend on how full you want the taste to be. For a fixed curve shape, loudness sets the amount to pour by a single clean factor: an input at IFI 3,000 delivers any chosen sweet target at one-third the material the reference would take, whether that target is half-max or eighty percent. This is the formulator's lever, and the table below reads straight off the canonical targets of the two bases we build in this chapter.

AXIS TARGET	IFI	POTENCY RATIO (IFI ÷ 1,000)	EC_{50}(CAND) VS REFERENCE	RELATIVE DOSE TO TARGET (1,000 ÷ IFI)
Sweetness · IFI-S	3,000	3.00	EC_{50}(ref) ÷ 3.00	0.333×
Umami · IFI-U	3,200	3.20	EC_{50}(ref) ÷ 3.20	0.313×
Kokumi · IFI-K (savory base)	2,400	2.40	EC_{50}(ref) ÷ 2.40	0.417×
Kokumi · IFI-K (sweet base)	900	0.90	EC_{50}(ref) × 1.11	1.111×

The last row is the honest one. A kokumi target of 900 has a potency ratio below one — that input is *quieter* than its reference and needs more than a reference dose to reach the same fullness. Not every engineered input is louder than the thing it stands in for, and the index tells you so plainly rather than flattering the design. A quiet input is fine; it simply costs more material, and the number lets you cost it before you ever open a tank.

The receptors we write to

Reformulation is receptor-level authorship, so it is worth naming the proteins by hand. Each of the four we lean on belongs to the class C family of G-protein-coupled receptors, built around a large Venus-flytrap domain that closes on its ligand like two cupped palms — and each gives us a different kind of handle.

Sweetness — TAS1R2·TAS1R3. The sweet receptor is a heterodimer, two subunits clasped together, and it is promiscuous by design: sugars bind in the flytrap cleft, while small-molecule sweeteners and the sweet proteins bind their own distinct sub-sites across the flytrap and the transmembrane bundle. Steviol glycosides occupy a pocket sucrose never touches. That separation is the whole opportunity — it means we can drive the identical downstream cascade (receptor → gustducin → phospholipase Cβ2 → IP_3 → the TRPM5 channel opening → depolarization → ATP release through CALHM1 → the afferent nerve) with a molecule that is not sugar, dosed to its own EC_{50}, read on its own IFI-S. The tongue reports sweetness; the sugar simply is not there.

Umami — mGluR1. Glutamate closes the flytrap and reports savor, but the axis has a gift written into its biology: the 5′-ribonucleotides, inosinate (IMP) and guanylate (GMP), bind an adjacent site and act as positive allosteric modulators, wedging the flytrap shut so that far less glutamate is needed for the same response. This is the synergy every cook has felt without naming — kombu with bonito, tomato with mushroom. Mathematically it is a gift to precision: if the ribonucleotide lowers the effective EC_{50} by a factor k, then by the index formula IFI-U rises by that same factor k. The synergy does not merely taste stronger; it *scores* stronger, and it does so with less material, which means a tighter absolute scatter around the target. When we pair a glutamate-rich tomato with a guanylate-rich shiitake, we are buying loudness and steadiness at once.

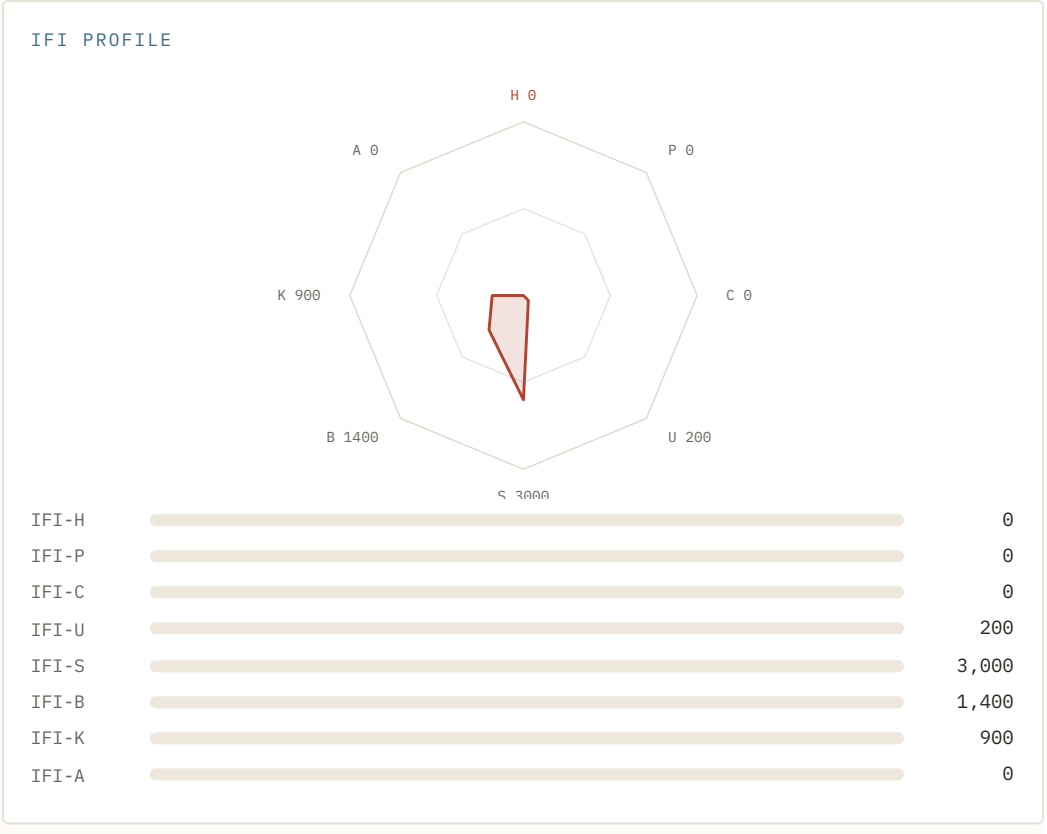
Zero-Bitter — TAS2R38. Bitterness is the taste salt and sugar were quietly suppressing, and the Zero-Bitter axis is an *attenuation* axis — it does not add a flavor, it turns one down. We model it as a fractional suppression a of the drive reaching TAS2R38, so the effective bitter stimulus becomes C(1 − a), and our platform's predicted design range is a = 0.20 to 0.65, a twenty to sixty-five

percent reduction. Push that through the Hill curve and the perceptual consequence is concrete: at a = 0.50 the effective bitter concentration halves, and with a Hill coefficient near one a stimulus sitting at its EC_{50} (θ = 0.50) drops to θ = C/2 ÷ (EC_{50} + C/2) = 0.33 — a third of the receptor quieted by halving its drive. That is the lid the salt used to hold, rebuilt as a number.

Kokumi — CaSR. The calcium-sensing receptor is the invisible axis. γ-glutamyl peptides — glutathione, γ-Glu-Val-Gly — are its agonists, and what they deliver is not a taste of their own but *continuity*: thickness, mouth-filling persistence, the sense that a sip has a body and does not thin out at the back of the tongue. Kokumi is exactly the dimension sugar and salt were secretly supplying, and CaSR is how we pour it back. An input dosed to IFI-K restores the roundness a subtraction stole, without adding a gram of the thing removed.

The sweetness case — the sugar out, the sweet kept

Take the sweetness reformulation concretely, as a target a formulator could build to. To pull the sugar from a beverage and leave the sweetness where the mouth expects it, we dose an engineered sweetness input to a point on IFI-S — the TAS1R2·TAS1R3 duet moved by steviol glycosides rather than sucrose — and because the input is held tight, we can aim at a specific target and hit it. The sugar that would have carried body and continuity is gone, so we grow that back with a measured touch of kokumi on the CaSR, and we soften the sweetener's own faint bitter tail with a little attenuation on TAS2R38. Three inputs, each held under the line, composed so the finished base is steadier than any one of them alone:



AXIS	INPUT	TARGET	BATCH CV	ROLE IN THE REFORMULATION
IFI-S · Sweet	Stevia	3,000	3.9 %	the sweetness itself, dosed to target; the sugar removed
IFI-K · Kokumi	Yeast	900	5.1 %	the body and continuity the sugar used to carry, grown back
IFI-B · Zero-Bitter	—	1,400	4.4 %	softens the sweetener's own bitter tail

The body reads sweetness, and continuity, and no sharp edge; the sugar it would have carried is simply not there. We say that as an attribute and nothing more — less sugar, the sweetness retained — and we make no claim beyond what the tongue reports and a label may plainly state. But notice what the precision made possible. Each of the three targets came off the dose relation of the last section: the IFI-S 3,000 input dosed to one-third of a reference sweetness, the kokumi filled to 900, the bitter tail pulled down by its computed fraction. Without a tightly held input, each of those is a hope; with one, each is a line on a batch sheet.

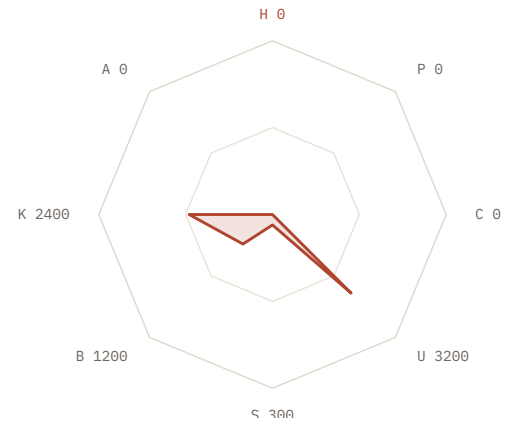
The savory case — the hardest subtraction

Sodium reduction is the hardest and most wanted reformulation of all, and it shows the composition doing real work, because salt does three jobs in a broth at once. It carries savoriness, it lends body, and it suppresses the bitterness underneath — so pulling it out opens all three holes, and every one must be filled precisely or the whole thing tastes hollow and off. We lift the umami on mGluR1 to carry the savor, grow continuity back through kokumi on the CaSR, and hold a computed attenuation on TAS2R38 to keep the bitterness the salt used to mask from surfacing.



A tomato restored to its full savory depth — reformulation is subtraction plus a precise replacement, and the replacement is only as good as its CV.

IFI PROFILE



IFI-H	0
IFI-P	0
IFI-C	0
IFI-U	3,200
IFI-S	300
IFI-B	1,200
IFI-K	2,400
IFI-A	0

AXIS	INPUT	TARGET	BATCH CV	ROLE AS THE SALT COMES OUT
IFI-U · Umami	Tomato / soybean	3,200	3.8 %	the savoriness the salt carried
IFI-K · Kokumi	Yeast	2,400	5.1 %	the body and roundness salt lent
IFI-B · Zero-Bitter	—	1,200	4.4 %	holds down the bitterness salt had masked

The savor does not fall from the sky; it is grown, in named organisms, by the glutamate and ribonucleotide and γ-glutamyl routes the receptors above are built to read. Which organism supplies which input is a design choice, and the library makes it a short, grounded one.

ORGANISM	CONSTRUCTS	PRIMARY AXES	SUPPLIES IN THE SAVORY BASE
Tomato	2,040	S / U / F / M	the umami backbone, glutamate via the native GAD / GDH1 route
Soybean	360	U / K	umami depth with a kokumi undertone
Yeast	240	U / K	the kokumi body, and a second umami lift
Koji mold	80	U	umami from protease-liberated glutamate
Shiitake	80	U	guanylate — the ribonucleotide that multiplies the umami
Garlic	560	K / P	kokumi, through its γ-glutamyl peptides

The mouth reads a full, savory, rounded broth; the sodium it would have carried is simply lower. We say that and nothing more — less sodium, the savor retained — making no claim beyond what the tongue reports and a label may plainly state. But the precision is the whole enabler, and here it becomes arithmetic. Each input arrives with a batch scatter, and a coefficient of variation is a relative number; to combine them we first convert each to an absolute scatter in index units, $\sigma = \mathrm{CV} \times \text{target}$. Then, because the three inputs are independent, their variances — the squares of those scatters — add.

AXIS	TARGET M	BATCH CV	Σ = CV × M	Σ²
Umami · IFI-U	3,200	3.8 %	121.6	14,786.6
Kokumi · IFI-K	2,400	5.1 %	122.4	14,981.8
Zero-Bitter · IFI-B	1,200	4.4 %	52.8	2,787.8
Composite	**6,800**	**2.65 %**	**180.4**	**32,556.2**

The integrated impression of "a full broth" is the sum of its axis contributions, so its mean is the sum of the targets, Σμ = 6,800, and its scatter is the square root of the summed variances, √32,556.2 = 180.4 index units. Divide the two and the composite coefficient of variation is 180.4 ÷ 6,800 = 2.65 percent — below the tightest of the three inputs that built it. The sweet base composes the same way to the same kind of answer: scatters of 117, 45.9, and 61.6 on targets summing to 5,300 give √19,589 ÷ 5,300 = 2.64 percent. Two different reformulations, two composite bases calmer than any single thread inside them.

Why three nervous inputs make one calm base — the 1/√n reward

It is worth seeing *why* the composite is steadier, because it is one of the loveliest results in the whole discipline and it is not luck. Take the simplest case first: n independent inputs, each with the same mean μ and the same scatter σ. Their sum S has mean nμ, and because independent variances add, its variance is nσ², so its scatter is σ√n. The coefficient of variation of the sum is therefore

```
CV(S) = σ√n ÷ (nμ) = (σ ÷ μ) ÷ √n = CV(single) ÷ √n
```

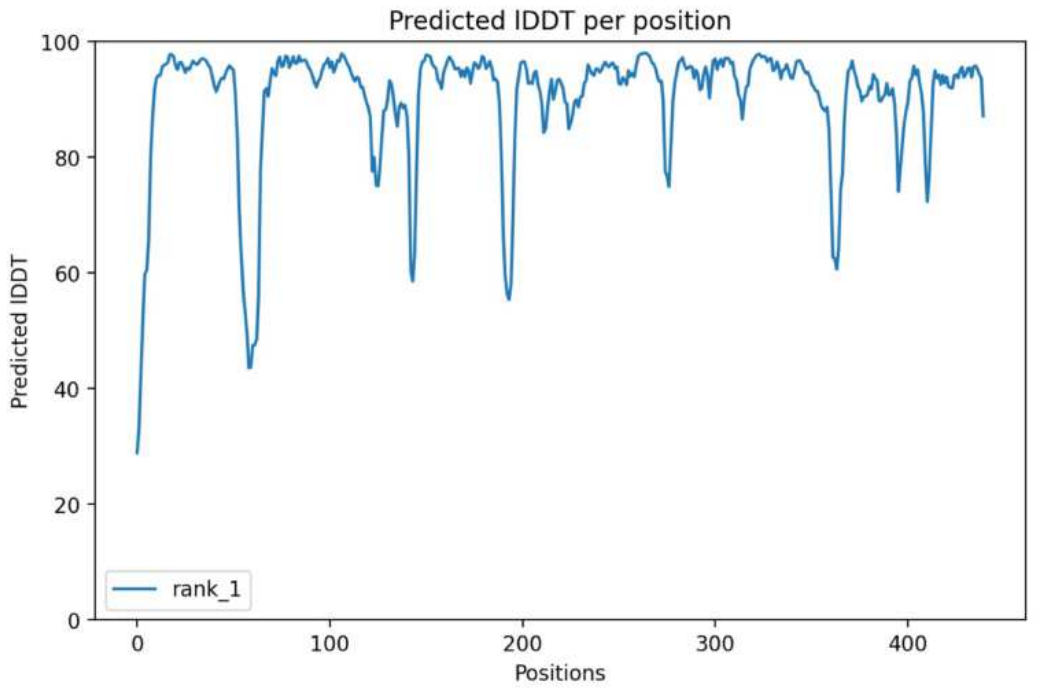
The scatter grows like √n while the signal grows like n, and the ratio of the two — the thing the tongue actually feels as inconsistency — falls as one over the square root of the count. Three equal inputs at five percent each compose to 5 ÷ √3 ≈ 2.9 percent. The general case, with unequal inputs, is the full formula we just applied — CV = √(Σσ²) ÷ Σμ — and it lands in the same place: 2.65 percent for the savory base, 2.64 for the sweet.

This is the quiet miracle a formulator leans on. You do not hold a reformulation together by making one heroic input flawless; you hold it together by composing several honest, independent inputs and letting the mathematics average their trembling away. The whole becomes calmer than its parts. It is the same reason a choir holds a pitch no single voice could — and it only works because each input is genuinely independent and genuinely on its number to begin with. Correlate them, or let one wander, and the gift evaporates. The reward is real, but it is a reward for precision, not a substitute for it.

The precision it all stands on — CV ≤ 7 %

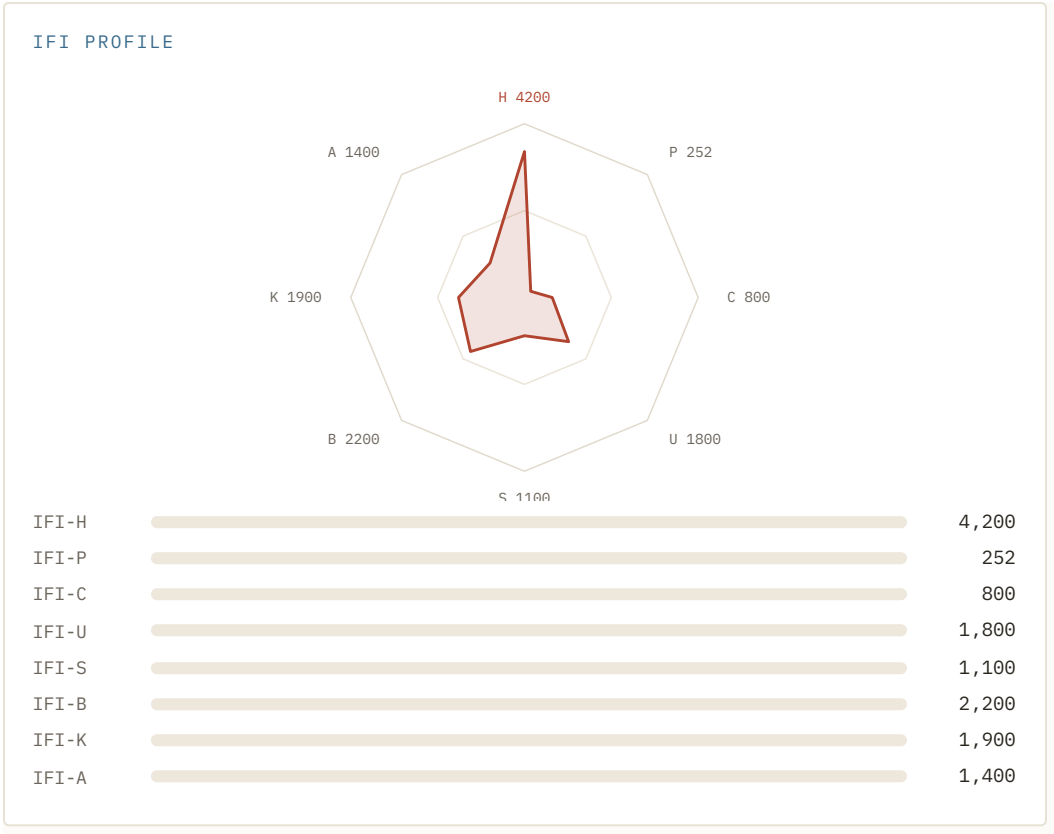
Everything above assumes the input arrives on its target. That assumption is not free; it is the hardest thing the platform does, and it is the reason this is a book about consistency. A pepper grown the old way delivers its heat with a batch-to-batch coefficient of variation of forty-two percent. Put that number into a reformulation and the dream collapses: a forty-two percent scatter means a two-

sigma spread of ±84 percent around the target, so some lots arrive at nearly double the intended intensity and some at almost nothing. You cannot pull the sugar from a drink with an input like that; you would ship cloying cans and watery cans from the same line and call it the same product. Reformulation is impossible on a wandering input, full stop.



Pun1/AT3 coloured by model confidence — the active-site cleft and binding domain are high-confidence, the structural ground on which two substitutions were chosen.

We brought that forty-two down to seven and below by refusing to breed by hope. A protein language model scored every single-residue substitution in the pathway's capsaicin synthase, Pun1, by its predicted effect on catalytic efficiency and fold stability; a genome-scale flux-balance model found that flux control sits upstream in the phenylpropanoid pathway — primarily at 4CL and HCT — not at the terminal Pun1 step; Pun1 is the capsaicin synthase that labels the product, and two targeted substitutions within it shape the output range into the precision band. The double variant — **S39L** and **L345G** — was written in with CRISPR. The two together lift predicted output to its target and, more to the point of this chapter, pull the batch-to-batch CV to seven percent or less. The same discipline proved itself at the far extreme of design in a single fruit designed to register on all eight axes at once:



That profile — the full-spectrum cultivar designed to a predicted 2.95 million SHU, every axis placed deliberately — is the opposite problem from a reformulation, and it is the same discipline. One piles eight loud axes into one organism on purpose; the other places two or three axes to a hair's-width target in a base. Both are only possible because each axis can be hit and held. The precision that lets us design the loudest fruit on earth is the precision that lets us quietly take the sugar out of a drink, and it is measured the same way, in the same dimensionless currency, on the same line.

Reading it on the line — the capability study

A target and a scatter together make a capability study, which is how a reformulation stops being a lab result and becomes a product a plant will run for years. Suppose a formulator sets a sensory window of ±15 percent around a target — a reasonable band for "the same product" — and asks whether a process can live inside it. The standard measure is the capability index, Cp = (window width) ÷ (6σ), which for a ±15 percent window reduces cleanly to Cp = 0.30 ÷ (6 × CV) = 0.05 ÷ CV. The whole argument of this chapter sits in that one little quotient.



A controlled growth chamber — a reproducible environment is the precondition for a reproducible input, and a reproducible input is the precondition for reformulation.

INPUT PRECISION	CV	Σ ON A 3,000 TARGET	±2Σ SPREAD	CP (±15 % WINDOW)	READING ON THE LINE
Wild-type, bred	42 %	1,260	±84 %	0.12	never in spec
Engineered cap	7 %	210	±14 %	0.71	marginal, alone
Single input, typical	4.5 %	135	±9 %	1.11	borderline capable
Composite base	2.65 %	79.5	±5.3 %	1.89	comfortably capable

Read top to bottom, the table is the case for the whole method. The bred input at Cp 0.12 is not a product; it is a lottery. A single engineered input at the seven-percent cap clears the lottery but sits marginal on its own, which is exactly why we compose. The composite base at Cp 1.89 clears the usual "capable" bar of 1.33 with room to spare and approaches the world-class range — a process that will sit inside its window lot after lot, which is all "every can the same" has ever meant. One honest refinement keeps us from over-claiming it: the batch CV we design to is the input's true process variation and the assay's own measurement scatter, folded together, and variances add, so observed CV^2 = process CV^2 + measurement CV^2. The true biological steadiness is a touch better than the number on the sheet — but we report the number on the sheet, because that is what the plant will actually see.

Edge cases, failure modes, and the honest limits

Precision earns its reverence by naming where it stops. The $1/\sqrt{n}$ reward assumes the inputs are independent and the perceived whole is their clean sum; real perception is neither perfectly additive nor perfectly independent. Sweetness suppresses the perception of bitterness along its own nonlinear curve, so a change on IFI-S quietly moves the apparent IFI-B — a cross-term the simple variance sum ignores. We handle it the only honest way: we measure the composed base on a panel rather than trusting the arithmetic to be the last word. The math tells us where to aim; the tongue confirms we landed.

The receptors themselves drift with their surroundings. EC_{50} is not a universal constant — it shifts with pH, with temperature, with fat and protein in the matrix — so a base tuned in clean water will read differently in a real sauce, and a reformulation validated at one serving temperature can slip at another. A design is a starting coordinate, not a finished product; the final targets are set in the actual matrix, at the actual temperature, in the actual package. The clean dose relation of the early sections carries one more fine-print assumption too: it cancels the curve-shape factor only when the reference and the candidate share a Hill coefficient. When an engineered input is markedly steeper or shallower than its reference, the dose ratio tilts with the target level, and we re-fit the curve rather than trusting the single factor.

Two last limits matter most, and they are the ones we hold hardest. The first is the line we never cross: every claim here is an attribute and only an attribute — less sugar, the sweetness retained; less sodium, the savor kept — never a promise about a body or a health. The index measures how loudly a molecule speaks to a receptor; it says nothing about what that does for a person, and we let it say

nothing. The second is the humility underneath the whole enterprise: IFI measures loudness, not liking. Precision delivers the target faithfully; it does not choose the target. That choice still belongs to a person with a spoon and a lifetime of eating behind them — a cook deciding what a full broth should feel like, a family deciding what tastes like being cared for. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory. All our mathematics does is make sure that once that human decision is made, it arrives intact in every can. Reformulation is where the arithmetic of this work becomes food on a table, and the table is still the point.

Chapter 20 · Consistency by design — *built in, not inspected in.*

Every lot we have ever shipped began the way a meal does — with an intention held long before the eating. The old factories kept a different habit, older than it should still be: build the thing first and ask whether it was good afterward, at the exit, with a gauge and a reject bin. Make the batch, measure it, throw away what failed, ship the rest. It is an expensive way to work and a defeated one, because it treats quality as an accident you catch rather than a decision you make. Our whole method is the refusal of that habit. You do not inspect a flavor into being right. You grow the variation out of it, upstream, at the four places that all come before the final test — the genetics that fix what the plant *is*, the room that fixes how it grows, the process and the assay that fix how it is handled and read, the reference that fixes the scale it is read against. Spend yourself there, at the doors, and the lot is already right before anyone at the exit so much as looks.



The controlled chamber at dawn — daily light integral, vapor-pressure deficit, root-zone temperature and nutrient EC all pinned, so that every plant in the room meets the same morning.

Gathered into one frame, the system is legible at a glance — each door, the variance it closes, the tool that closes it, and the moment in the season when it is shut. Nothing in the last column happens at the exit. Every one of these controls lives upstream, weeks or seasons before a sample is ever drawn.

DOOR	VARIANCE IT CLOSES	HOW IT IS SHUT	WHEN IT IS SHUT
Genetics	σ^2(genetic) → ≈ 0	one isogenic genotype, edits fixed homozygous and sequenced	before a seed is sown
The room	σ^2(environment)	pinned DLI, VPD, root-zone temperature, nutrient EC	every day of the grow
Process	σ^2(process)	standardized recovery, low-heat drying, blend uniformity	at harvest and after
The ruler	σ^2(analysis)	stable receptor assays, gauge R&R, a traceable reference	at every measurement

What follows is the arithmetic of that promise — the measure we hold a lot to, the rule that tells us which door to shut first, the enzymology that lets us shut the widest one outright, and the reason the certificate at the end is a formality rather than a filter.

The honest measure — *what a coefficient of variation actually says.*

A number is only as trustworthy as its scatter, and the honest way to state scatter is relative. The coefficient of variation is the standard deviation divided by the mean — the spread written as a fraction of the thing itself, CV = σ / μ, carried as a percentage. Relative, because a standard deviation of twenty thousand Scoville units is a rounding error on a superhot and a catastrophe on a shishito; only the fraction travels honestly across the whole ladder.

Make it concrete. Our full-spectrum flagship, Phoenix, is designed to reach a predicted heat of approximately 2.94 million SHU. Held to a batch-to-batch CV of seven percent, its standard deviation is σ = 0.07 × 2,937,948 ≈ 205,656 SHU. For a well-behaved process that is normally distributed, better than two batches in three therefore land inside one standard deviation of the target — a window from roughly 2.73 to 3.14 million SHU — and virtually all of them, 99.7%, inside three. A wild-type population carrying the same trait at a CV of forty-two percent scatters six times as wide relative to its own mean: one plant reads nearly half again its neighbor, and no blend built on it is ever twice the same. That sixfold gap between 42% and ≤7% is the entire difference between a curiosity and an ingredient.

It is worth passing between heat's two rulers, because consistency has to hold in both. The receptor speaks in loudness; the lab bench speaks in concentration, and they are bridged by the old pungency conversion, SHU ≈ 16 × ppm of capsaicinoids. Phoenix's predicted 2.94 million SHU corresponds to roughly 183,600 ppm — 183.6 milligrams of capsaicinoids in every gram, better than eighteen percent by weight, a crystalline-range concentration our model puts well beyond any wild fruit and that only a designed, homozygous line holds steady.

And one distinction to carry through the whole chapter: a CV is a statement about *precision*, not *accuracy*. Seven percent is a tight scatter; it says nothing, by itself, about whether the center of that scatter is true. Precision is the business of the first three doors. Accuracy is the business of the fourth, the reference, and we return to it. The seven-percent target itself was not chosen for elegance — it was chosen against the mouth. Intensity perception follows Weber's law: the smallest difference a sense can resolve is a roughly

fixed fraction of the stimulus, and for taste that fraction sits well above a tenth. A batch held inside seven percent falls under the threshold at which even a trained palate reliably tells one lot from the next. The eater meets the same pepper every time — not because the molecules are identical, but because the remaining difference is below the resolution of the tongue that reads them.

Why variance adds as squares — *the quadrature rule, and the order of attack.*

Four independent sources of scatter stack, and the way they stack is the single most useful fact in the enterprise. Independent variances *add* — not the standard deviations, the variances, the squares:

```
σ²(total) = σ²(genetic) + σ²(environment) + σ²(process) + σ²(analysis).
```

Carried into relative terms around a common mean, that becomes the quadrature rule:

```
CV(total) = √[ CV(genetic)² + CV(environment)² + CV(process)² + CV(analysis)² ].
```

Take a representative budget for the wild-type plant — a genetic term near 38%, an environmental term near 15%, a process term near 8%, an analytical term near 5%. The rule returns $\sqrt{(38^2 + 15^2 + 8^2 + 5^2)} = \sqrt{(1{,}444 + 225 + 64 + 25)} = \sqrt{1{,}758} = 41.9\%$ — which is the forty-two percent we actually measure in an unimproved line. The component split is representative; the endpoint is not.

The squaring is the whole lesson. Because each term enters as its square, the largest one swamps the rest, and shrinking anything but the largest is nearly wasted effort. Drive the five-point analytical term to zero and you remove 25 from a sum of 1,758 — the total falls from 41.9% to 41.6%, a change no one could taste. Drive the thirty-eight-point genetic term to two and you remove 1,440 — the total falls to under eighteen. So the order of attack is not a matter of preference; it is dictated by the arithmetic. Shut the widest door first, then the next widest, until the largest term still standing is small enough that the finished lot already sits where the specification wants it. Watch it happen one row at a time:

STAGE	GENETIC	ENVIRONMENT	PROCESS	ANALYSIS	TOTAL CV
wild-type, nothing shut	38%	15%	8%	5%	41.9%
genetics fixed	2%	15%	8%	5%	17.8%
room pinned	2%	4%	8%	5%	10.4%
process standardized	2%	4%	3%	5%	7.3%
ruler measured	2%	4%	3%	2%	5.7%

Read the last column downward and you can watch consistency being *built*: 42 to 18 to 10 to 7 to under 6, each step the closing of one door, and not a single step of it happening at the exit.

The widest door — *fixing what the plant is.*

The genetic door is the widest because a wild pepper is not one thing. An open-pollinated field is a crowd of genotypes — heterozygous, segregating, every plant a fresh shuffle of alleles. The capsaicinoid pathway's output depends on the dose and the exact sequence of the gene at its mouth, the acyltransferase the literature calls *Pun1* (the AT3 enzyme), and on a tier of promoter and modifier variants feeding it. Plant to plant, that machinery differs, so the heat differs, and σ^2(genetic) becomes the thirty-eight points that dominate the budget. You cannot pin an output whose governing constants are reshuffled in every individual.



The sequencer confirms what the edit promised — two substitutions, read back base by base and fixed homozygous, before a single seed is ever allowed into the room.

So we shut it the only way engineering can shut a door outright: one genotype, chosen and fixed. The two edits that define the heat are written *homozygous* — identical on both chromosomes, so nothing segregates in the next generation — and every seed lot is sequenced back to confirm the plant is what the design says before it enters the grow. One isogenic line, read base by base. With identity fixed, σ^2(genetic) collapses toward zero, and the widest door is shut before a seed is sown. It is the only term in the whole budget that engineering can take to near-nothing rather than merely shrink, which is the deepest reason it goes first.

What is actually being fixed is a *rate*. Capsaicin is built in one committed step: the acyltransferase condenses vanillylamine with an 8-methyl branched-chain acyl-CoA to form the molecule, and the flux through that step obeys Michaelis–Menten kinetics,

```
v = k_cat · [E] · [S] / (K_m + [S]),
```

set by how much active enzyme is present ([E]), how tightly it grips its substrate (K_m), and how fast it turns over (k_cat). Fix the gene and you fix all three constants in every plant at once. What remains is to make those constants as favorable, and as *stable*, as they can be — because a constant that wavers with the weather is not truly a constant, and its wavering is scatter in the final number.

The two shipped edits, and the fold they stabilize — *folding math.*

An ESM2 650-million-parameter protein language model scored every single-residue substitution in Pun1 for its predicted effect on catalytic efficiency and fold stability. Our flux model predicts control

sits upstream in the phenylpropanoid pathway — at 4CL and HCT — not at the terminal Pun1/AT3 step; Pun1 is the final capsaicin synthase, the step that assembles the molecule, not the one that governs the pathway rate. The reason Pun1 is still worth careful engineering is different: an enzyme that wobbles between its folded, active state and misfolded dead ends writes scatter into every batch it touches regardless of what controls the upstream supply. Two substitutions entered the final shipped design, chosen from the ESM2 scan and confirmed by structural modelling; a third candidate, C175S, was identified by our modelling as a potential disulfide-misfold prevention site and is noted in the structural record, but was not carried into the final construct.

EDIT	CATALYTIC EFFECT	ΔΔG (KCAL/MOL)	WHAT IT BUYS CONSISTENCY
S39L	tighter substrate grip near the N-terminus; lowers K_m	−2.1	more flux per unit substrate; steadier against supply dips
L345G	steric relief in the active-site cleft; raises k_cat	−1.8	most of the raised ceiling, at the same enzyme load
C175S (structural annotation — not in shipped construct)	replaces a stray cysteine; closes a potential disulfide-misfold route	−0.9	structural observation from modelling; not included in final design
S39L + L345G (shipped double variant)	—	−3.9	folded fraction 99.37% → 99.9991%

S39L, near the amino terminus, tightens the grip on the substrate — a lower K_m, so the enzyme runs closer to its maximum even when the plant's own substrate supply dips, which steadies output against the small day-to-day swings in its chemistry. L345G relieves a steric crowd in the active-site cleft and lets the enzyme turn over faster — a higher k_cat, and most of the raised ceiling. The two edits were chosen for their catalytic improvements, and they also deepen the fold.

Why that matters to consistency is a piece of equilibrium thermodynamics. Whether a protein sits in its working, folded shape is a Boltzmann question: the folded fraction *f* is set by the free-energy gap ΔG between folded and unfolded states against the thermal energy RT,

> `f = 1 / (1 + e^(−ΔG / RT)), with RT ≈ 0.593 kcal/mol at 25 °C.`

A marginally stable plant enzyme — take ΔG ≈ 3.0 kcal/mol, ordinary at growing temperature — folds correctly about 99.4% of the time: f = 1 / (1 + e^(−3.0/0.593)) = 1 / (1 + 0.0063) = 0.9937. Roughly six parts in a thousand of every batch of enzyme is born useless. The two shipped edits together add 3.9 kcal/mol of fold stability — their ΔΔG values, −2.1 and −1.8, sum to −3.9, and a negative ΔΔG is a deeper, steadier well. That lifts ΔG to 6.9 kcal/mol and f to 1 / (1 + e^(−6.9/0.593)) ≈ 0.999991: the misfolded pool falls from about 6,300 parts per million to under 9, more than seven hundred-fold.

But the raised yield is not the point — the *steadiness* is. The sensitivity of the folded fraction to any wobble in that free-energy gap, a half-degree of root-zone drift or a shift in cellular pH, runs as f(1 – f):

```
Var(f) ≈ [ f(1 – f) / RT ]² · Var(ΔG).
```

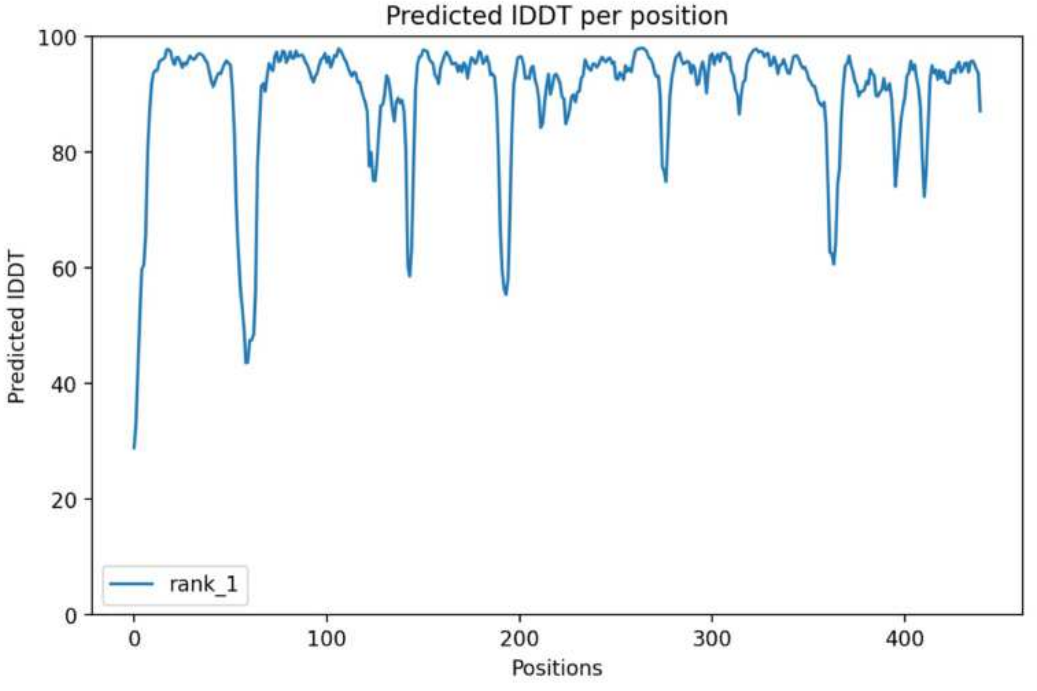


FIG. — Pun1/AT3 coloured by pLDDT confidence; the active-site cleft the two edits stabilize is the high-confidence core, the part that folds the same way in every cell of every plant.

At 99.4% folded, f(1 – f) is 0.0063; at 99.999% it is roughly 0.0000089. The enzyme has not merely become more abundant in its working form — it has become about seven hundred times *less responsive* to the conditions around it. It now sits on the flat shoulder of the Boltzmann curve, where the room can shiver and the output does not move. Stability, written into the gene, is spent as consistency in the tank.

The pinned room and the standardized tank — *environment and process.*

The stabilized enzyme makes the room's job easier, and the room returns the favor. Enzyme rate climbs with temperature — roughly a doubling for every ten degrees, the familiar Q10 near two — so an unpinned greenhouse writes its every warm afternoon and cold night straight into the capsaicin number. We pin the inputs the plant actually integrates: the daily light integral that meters its energy, the vapor-pressure deficit that governs how it breathes, the root-zone temperature that sets its metabolic tempo, the nutrient EC that doses its raw materials. Hold those flat across the whole grow and the environmental term, fifteen points in the wild, collapses to a handful. And because the edited enzyme now rides the flat shoulder of its folding curve, the residual degree or two the room cannot perfectly hold barely moves the output — so the genetic fix and the environmental fix *multiply* rather than merely add. A steady enzyme in a steady room is steadier than either alone.

Then the process — everything between the living fruit and the measured powder. Capsaicin is thermolabile; dry it too hot and you degrade the very molecule you grew, and worse, you degrade it *unevenly*, manufacturing scatter at the last step from a lot that was

uniform on the plant. So recovery is standardized, drying is kept low and slow, and the blend is homogenized until the sample the assay draws is the lot it claims to represent. These are not glamorous controls. They are the difference between a number that means the whole batch and a number that means only the corner of it the scoop happened to reach.

The ruler — *EC50, the dose–response curve, and the honest error on a number.*

A measurement is a door too, and left open it leaks scatter into the record that was never in the lot. Our ruler is the receptor itself. To read heat as loudness we build a dose–response curve: graded concentrations of a standardized extract against the fraction of the TRPV1 channel population they activate, a response that climbs in the sigmoid of the Hill equation,

$$R(C) = R_{max} \cdot C^n / (EC_{50}^n + C^n),$$

with a Hill coefficient *n* near two for capsaicin at TRPV1. From the whole curve we take exactly one number, EC_{50} — the concentration that drives the channel to half of its full response. A louder food reaches half-maximal with less of itself, so a lower EC_{50}. The index then inverts that against a fixed reference and scales by a thousand:

$$IFI = 1000 \times EC_{50}(\text{reference}) / EC_{50}(\text{candidate}).$$

Anchor the reference so its own score reads exactly one thousand — set EC_{50}(reference) at 42 µg/mL, and a candidate that half-activates at 42 µg/mL returns 1,000 by construction. Phoenix is designed to half-activate at approximately 10 µg/mL, 4.2 times less needed than the reference, giving a predicted IFI-H = 1000 × 42 / 10 = 4,200. The whole flagship ladder falls out of the same two measurements — the receptor's loudness and the bench's concentration — and the two rulers agree:

CULTIVAR	SKU	IFI-H	EC_{50} (DESIGN TARGET), MG/ML	SHU (MODEL PREDICTION)
wild-type reference (anchor)	—	1,000	42.0	–
Ember	SS-0087	1,570	26.8	866K
Blaze	SS-0091	2,630	16.0	1.72M
Firestorm	SS-0094	3,490	12.0	2.47M
Phoenix	SS-0096	4,200	10.0	2.95M

That the receptor assay and the independent HPLC-and-Scoville number rise together is not decoration — it is a cross-check. Two rulers, one biological and one chemical, reading the same lot: when they disagree, one of them has drifted, and you have caught a fault in the *measurement* before it was ever blamed on the plant.

The index is a ratio, and a ratio carries the error of both its parts. For independent relative errors the combining rule is the quadrature rule again, one level down:

```
CV(IFI)² = CV(EC₅₀ reference)² + CV(EC₅₀ candidate)².
```

A reference re-qualified to two percent and a candidate assay repeatable to three percent yield an index precise to $\sqrt{(2^2 + 3^2)} = \sqrt{13} \approx 3.6\%$. The same squares-add arithmetic that ordered the entire manufacturing budget now governs the confidence interval on the printed score — the rule is scale-free, and it is everywhere.

Two disciplines keep that analytical term small. *Gauge R&R* — repeatability within a single operator and instrument, reproducibility across operators and instruments — demands that the measurement's own spread stay a small fraction of the lot-to-lot spread, so the ruler never masquerades as product variation; in practice it means bracketing the true EC_{50} with dose points on both sides so the curve is *fitted*, not extrapolated off its steep flank. And *traceability* keeps the ruler honest in the other sense, because precision is not accuracy. A CV of seven percent is a tight scatter that says nothing about whether the center is true. The reference extract *is* the center, and it is re-qualified against a primary standard on a schedule — because a reference that quietly degrades does not scatter the scores, it slides them all down together, a bias no amount of repetition will reveal. Scatter you can see in the data; bias you can only design against.

Capability, and why the last gate is a formality — *Cp, Cpk, and acceptance sampling.*

Everything above is upstream. The exit test only confirms it, and the manufacturer's own language says why it can be trusted to. Process capability compares what the process *does* to what the customer *asked for*. Cp sets the specification width against the process width; Cpk also penalizes a process for sitting off-center:

```
Cp = (USL − LSL) / 6σ, Cpk = min[ (USL − μ) / 3σ, (μ − LSL) / 3σ ].
```

Take Phoenix's heat with a design target of IFI-H 4,200 and held to a six-percent CV, so σ = 252. A hot-sauce formulator who blends to a target can accept an input anywhere inside ±25% — a specification from 3,150 to 5,250 — and against that window the process is comfortably capable: Cpk = (5,250 − 4,200) / (3 × 252) = 1,050 / 756 = 1.39, past the 1.33 that industry calls capable. At that capability the fraction of a lot falling outside the limits is on the order of sixty parts per million. A release plan that pulls twenty units from a lot then has roughly a one-in-eight-hundred chance of so much as *seeing* a nonconforming unit — which is precisely why acceptance sampling finds almost nothing wrong. There is almost nothing wrong to find.

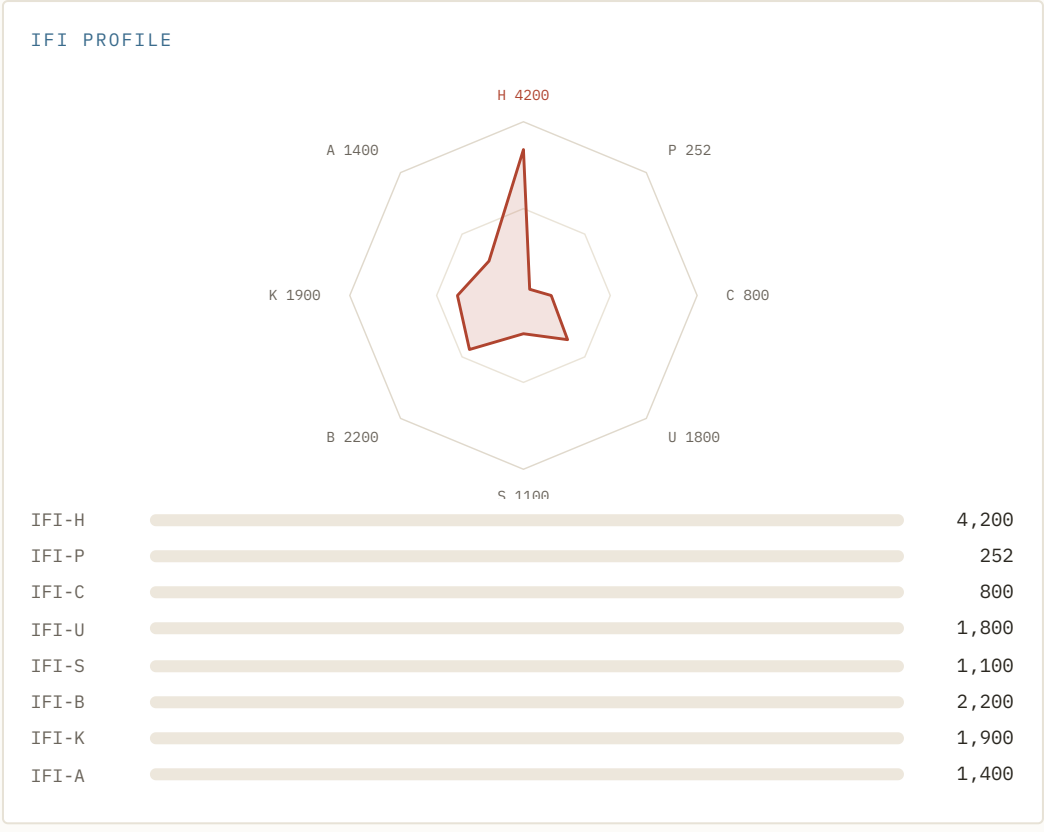
The same arithmetic marks its own honest limit, and we will not hide it. Point that identical process at a specification four times tighter — a pharmaceutical-grade standard demanding ±10% — and Cpk falls to (420) / (3 × 252) = 0.56, not capable at all. Seven percent is a

luxurious tolerance for heat, where the formulator holds the final target by blending; it is only a marginal one where the number itself is the finished product and the window is narrow. To make a ±10% window capable you would have to drive the CV under about two and a half percent, a deeper and more expensive run at every door. Consistency is not a single trophy you win once — it is a figure you buy to a purpose, and the purpose sets the price.

Between releases the process is watched on control charts, where the limits are drawn from what the process *actually does* — three standard deviations of its own behavior — not from what the customer wants. Those are two different voices, and keeping them apart is the whole discipline. A run inside its control limits is in control; a single point outside them is a special cause — a drifted sensor, a tiring reference, an off-type plant — flagged and corrected before it reaches a lot, rather than discovered in one afterward.

The hardest case — *eight axes held at once.*

Everything so far holds one number steady. Phoenix holds eight.



Full-spectrum is the cruelest consistency problem, because a profile passes only if *every* axis passes, and independent probabilities multiply. Suppose each of the eight axes passed, generously, ninety-five percent of the time. The whole profile would then pass $0.95^8 = 0.66$ — barely two runs in three, a third of every batch failing on one axis or another, and no inspection at the exit able to rescue it, because you cannot inspect eight independent numbers into simultaneous compliance. The only way through is upstream: make each axis capable on its own. Hold every axis to the 1.33 capability

that leaves roughly sixty defective parts per million — a per-axis pass of 0.999937 — and the whole eight-axis profile still passes $0.999937^8 = 0.9995$, better than nine hundred and ninety-nine runs in a thousand. The compounding that destroys a sloppy process is exactly what a capable one shrugs off. It is the quadrature lesson in different clothes: stack enough independent requirements and only the disciplined per-unit number keeps the whole from unravelling. A full-spectrum fruit is not eight easy problems; it is one hard one, and it is won before harvest or not at all.

Where it still goes wrong — *failure modes, and the limits we will not paper over.*

A system this specified fails in specified ways, and naming them is part of trusting it. A grow-room sensor drifts and the whole lot shifts together — caught on the control chart as a run riding above the centerline, before harvest. A reference reagent ages and every IFI score slides down in lockstep — caught not by scatter but by the chemical ruler disagreeing with the biological one, and by the re-qualification schedule standing behind both. An off-type plant slips the roguing — caught by the sequencing that reads every seed lot. The drying runs hot one afternoon — caught as a heat number that fell when nothing upstream explains the fall. Each door has its characteristic failure and its characteristic catch, and the catches are themselves designed, not merely hoped for.

And the limits we will not paper over. Seven percent is biology's neighborhood, not a tablet press's; a living fruit integrating light and water and its own development will not be coerced to the CV of a stamped steel part, and we would distrust a number that claimed it had. Some axes are harder to hold than heat — astringency, read through the aggregation of salivary proteins rather than the clean sigmoid of a single cation channel, carries an intrinsically noisier assay, and its honest CV sits higher than the heat axis does. The index is only ever as accurate as the reference beneath it, so accuracy is a standing commitment, not a one-time calibration. None of these is an apology. They are the edges of the envelope, and knowing exactly where they lie is the difference between a number you can build on and one that will betray you quietly.

What the certificate actually certifies — *a property grown in.*

So the certificate of analysis passes, and it is worth being exact about why. Not because a clever gate at the exit sieved the good lots from the bad — that gate is nearly idle, catching its rare genuine fault once in a long while, usefully, and otherwise waving through a river of lots that were right before they ever reached it. It passes because four doors were shut, in order, weeks and seasons earlier: identity fixed in a sequenced, homozygous gene; expression fixed in a pinned dawn; handling and reading fixed in a standardized tank and a traceable receptor; the scale itself kept true against a primary standard. The seven-percent line at the end is not a hurdle the lot clears on its way out the door. It is a property the lot was *grown to*

have, carried in its cells from the first measured morning. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.

Quality built in costs less and holds better than quality inspected in, in every craft that has ever learned the lesson — we simply learned it where flavor begins, at the level of a gene and a room. A pepper made this way is not merely consistent from one batch to the next. It is reliably, exactly itself, grown with the same intention you would bring to anything you meant to feed someone you love — which is the most honest thing a food can be, and the quiet promise under everything we grow.

Closing essay · The same good thing, every time

We began this book with a promise hidden inside a single small word, and we will end it where that promise is actually kept or broken — at a table, with the people we love, waiting for something we made. The deepest thing our family knows about feeding people is that they do not, in the end, remember the one perfect meal. What they remember, what they come to *trust,* is that it was always good — that the dish set in front of them is the same kindness every time, so that sitting down at your table becomes a thing they can lean on without having to think about it, without the small anxious question of whether tonight will be the night it is off. That trust is the whole reward, and it has a precise technical name. It is a coefficient of variation held under seven percent: not a number for engineers, in the end, but trust itself, made measurable and then made true.

So let this last essay be the arithmetic of that promise — what the number means, how we draw it from a plant, where its seven percent actually comes from, and, because numbers are a form of respect for the reader, exactly where it stops. The mathematics in these pages was one pair of hands at a keyboard; the table it is all for belongs to a family, and everything below is written to keep that table honest. We measure taste in eight axes and hold each one to a line, and we would like to show you, plainly, how a line is held.

What a coefficient of variation actually says

A coefficient of variation is the most honest short sentence a process can speak about itself. Take a run of fruit, measure the one thing you care about in each, find the average μ they cluster around and the standard deviation σ of their scatter, and divide:

```
CV = σ / μ
```

Written as a percent, it is scatter expressed as a fraction of the signal — a pure ratio, blind to units, so a pepper measured in Scoville heat units and a tomato measured in micromolar glutamate can be judged by the same yardstick. A small CV means the fruit agree with one another. A large one means they do not, and that every plate built from them is a small gamble.

Set the yardstick against our heat flagship and the abstraction turns concrete. Phoenix — SS-0096 — is designed to a **predicted mean of approximately 2,937,948 SHU**. Heat on our scale is really a statement about chemistry: the modern assay reads capsaicinoids by HPLC and multiplies by sixteen to land in Scoville units, so that same predicted mean is

```
2,937,948 SHU ÷ 16 = 183,622 ppm capsaicinoids = 18.36 % by mass,
```

which is also exactly where pure capsaicin's 16,000,000 SHU would place it — 2,937,948 ÷ 16,000,000 = 0.1836 — the two readings agreeing to the decimal, which is the first quiet sign the model's arithmetic is internally self-consistent. Now hold that fruit to CV ≤ 7 %. The standard deviation is no longer an idea; it is

> $\sigma = 0.07 \times 2{,}937{,}948 =$ **205,656 SHU** $= 12{,}854$ ppm $=$ **1.29 % by mass**,

so a single Phoenix pod carries 18.36 % capsaicinoids give or take about a percentage point, and ninety-five lots in a hundred land inside ±2σ — a window of ±411,313 SHU, ±14 % of target, roughly 2.53 to 3.35 million Scoville. The fire is enormous; the *disagreement* about the fire is small.

The older way of farming heat, selecting the hottest-looking plant and keeping its seed, carries a CV near 42 %. Because the coefficient is scale-free, the two numbers compare directly: at any shared mean, variance scales as the square of the ratio,

> $(0.42 / 0.07)^2 = 6^2 =$ **36**,

so the engineered line scatters with one thirty-sixth of the old variance — six times tighter on the standard deviation, thirty-six times calmer in the spread of outcomes. That factor of thirty-six is the whole difference between a good year and a reliable one.

PHOENIX (SS-0096) · HEAT DISPERSION	VALUE
Predicted mean	2,937,948 SHU
Capsaicinoid equivalent	183,622 ppm (18.36 % w/w)
Standard deviation (1σ, CV 7 %)	205,656 SHU (1.29 % w/w)
95 % window (±2σ)	±411,313 SHU (±14 %)
95 % range	≈ 2.53 M – 3.35 M SHU
Variance vs 42 % selection, equal mean	36× lower

Reading the index, and the error it carries

Heat is only one of eight languages the tongue speaks, and the rest do not come in Scoville. To hold them on one page we anchor every axis not to *how much* of a compound a food contains but to *how loudly* that compound is heard — the concentration at which it drives its receptor to half of full response. That half-maximal concentration is the EC_{50}, and the Iyengar Flavor Index is simply a reference EC_{50} divided by the candidate's, scaled by a thousand:

> $\mathrm{IFI} = 1000 \times EC_{50}(\text{reference}) / EC_{50}(\text{candidate})$

The reference is defined to read exactly 1,000, so the index is a loudness multiplier in disguise. Each reference EC_{50} is a stated convention — fixed by definition on that axis, the way Scoville's scale is defined by dilution-to-threshold — not a value averaged from published literature. A food that reaches half-activation at a quarter of the reference concentration is four times as loud and scores near 4,000; one that needs four times as much is a quarter as

loud and scores near 250. Phoenix's heat axis sits at 4,200, which reads back as a concentration ratio directly: EC_{50}(candidate)/EC_{50}(reference) = 1000 / 4200 = 0.238, meaning its capsaicin half-activates TRPV1 — the trigeminal cation channel that also opens above 43 °C, so the body reports a fire that is not there — at **23.8 % of the reference concentration**, about 4.2 times louder than the wild-type baseline. The scorecards cap each axis at 5,000, the normalization ceiling, so even Phoenix's heat leaves headroom on the page.

The same arithmetic reads any receptor. Phoenix's savory depth scores 1,800 on mGluR1, the metabotropic glutamate receptor that hears broth: 1000 / 1800 = 0.556, so its umami reaches half-activation at **55.6 % of reference**, about 1.8 times louder than a baseline tomato against its own umami reference — a gentler multiple than its heat against the heat reference. Each axis is calibrated to its own stated convention, so a heat score of 4,200 and a savory score of 1,800 describe loudness relative to separate baselines and are not drawn from a single common scale. That difference in multipliers is exactly why the fruit reads as *savory-and-hot* rather than merely hot. One formula, eight receptors, every food placed on a single page — with each axis's scores comparable to any food on that same axis, while cross-axis comparisons should account for the separate reference conventions each axis carries.

It helps to see one EC_{50} in real units, because loudness is a molecular fact before it is an index. Capsaicin half-activates TRPV1 on the order of 0.7 µM; at a molar mass of 305.4 g/mol that is

```
0.7 µmol/L × 305.4 g/mol ≈ 0.21 µg/mL,
```

a fifth of a microgram per milliliter standing between a tongue and the sensation of flame. The channel's response does not rise gently with dose. It climbs steeply and cooperatively — a Hill function, fractional response $\theta(C) = C^n / (EC_{50}^n + C^n)$ with an exponent n near two — so the window between *nothing* and *everything* is narrow. At the half-point $C = EC_{50}$ the response is 0.5 by definition; halve the dose and it falls to $0.5^2 / (1 + 0.5^2) = 0.2$, double it and it leaps to $2^2 / (1 + 2^2) = 0.8$. A fourfold swing in concentration carries the receptor across most of its range, and the slope is steepest precisely at the half-point — which is why EC_{50} is the one landmark a lab can pin again and again, a tiny change in response fixing the concentration tightly. We did not choose it for elegance. We chose it because it is the most sharply defined, and therefore the most forgiving, point to measure.



The sequencer and assay bench where every lot is read before it is believed.

And here is where the seven percent stops being a target and becomes physics. The index is a quotient of two measured quantities, and errors in a quotient add in quadrature — the relative uncertainties combine as the square root of the sum of their squares:

> $CV(IFI) = \sqrt{[\ CV(EC_{50}\ ref)^2 + CV(EC_{50}\ cand)^2\ ]}$

A careful receptor assay reads an EC_{50} to about 5 % on a good day. Feed two such reads into the index and

> $CV(IFI) = \sqrt{(0.05^2 + 0.05^2)} = \sqrt{0.005} = \mathbf{0.0707 = 7.07\ \%}$.

That is the honest floor. We drove the *biological* scatter toward zero with the genetics — the plant now makes almost exactly what the genome tells it to — but what remains is the measurement reading itself, two EC_{50} determinations each good to roughly five percent, meeting at $\sqrt{2} \times 5\ \%$. Seven percent is not a number we picked to sound impressive. It is the noise of our own instruments, and it is the one source of variation we cannot engineer away without building a better assay. The reverent thing to do with such a limit is to name it.

Separating the plant from the gauge

There is a trap folded inside every coefficient of variation, and a quality system earns its name by not falling into it. When we weigh a lot's scatter we are never measuring the plant alone; we are measuring the plant *plus the instrument that read it,* and the two variances add. At a fixed mean,

> $CV^2(observed) = CV^2(process) + CV^2(measurement)$

Suppose a lot returns 7 % observed, and the assay itself — those two EC_{50} reads meeting at $\sqrt{2} \times 5\ \%$ — contributes about 5 %. Then the plant's own biological scatter is not seven percent at all:

> $CV(process) = \sqrt{(7^2 - 5^2)} = \sqrt{24} \approx \mathbf{4.9\ \%}$

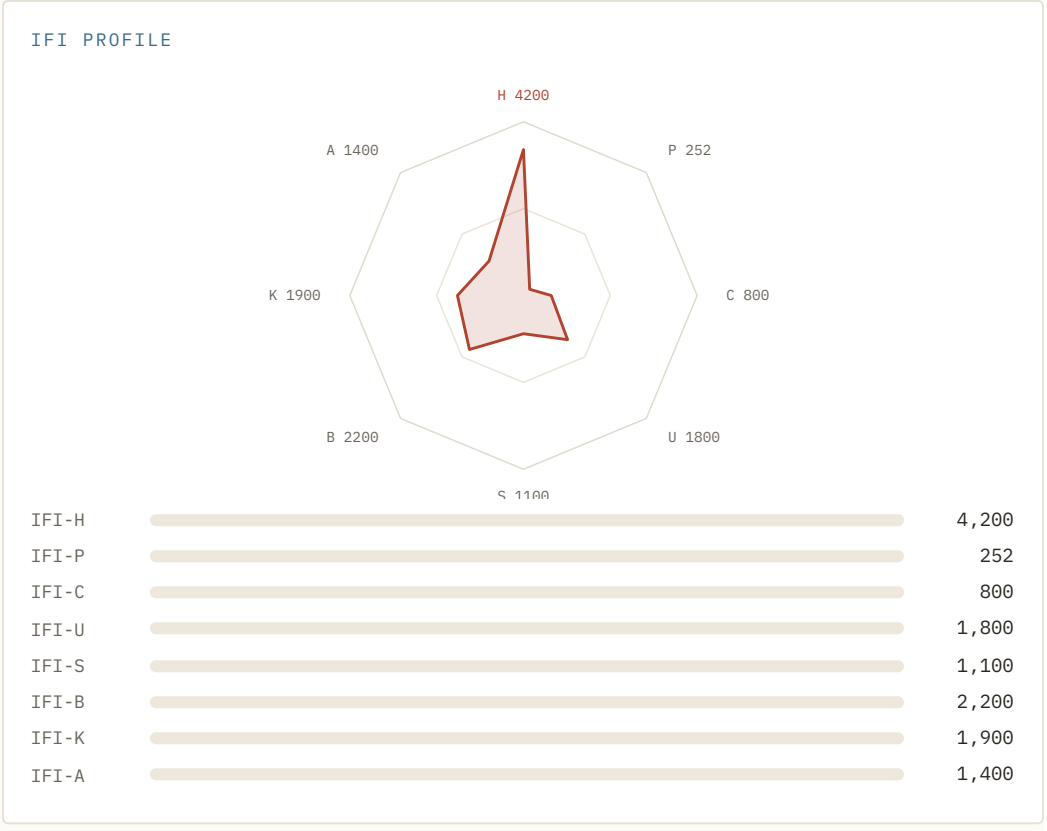
Nearly half of what looks like the crop misbehaving is, in truth, the ruler trembling. Chase that seven percent as though it were all biology and you will re-engineer a genome to cure a fault that lives

in the spectrometer. The honest order is to measure the gauge first and the plant second. Averaging m independent reads shrinks the measurement term as 5 % / √m, so three reads pull the gauge to 2.9 % and the *observed* lot CV falls from 7 % toward √(4.9² + 2.9²) ≈ 5.7 %, closing on the true biological 4.9 % as replicates accumulate. Nothing about the fruit has changed; we have only stopped blaming it for our own eyesight.

It is also why we keep two different sets of limits and never let them blur. *Spec* limits belong to the person eating — the band a lot must fall inside to carry the name. *Control* limits belong to the process itself, drawn at ±3σ of its established behavior, and a lot wandering past them is a signal to walk into the greenhouse and look, even when it is still sitting comfortably in spec. A number that holds is not a number that happened to land in spec once. It is a process that has told us, run after run, that it is still itself.

Holding eight lines at once

The hardest part was never one axis. It was eight. Phoenix was the first cultivar grown to register on every receptor at once, and its profile is the clearest picture of what "consistency" costs when a plant has to be right in eight directions simultaneously.



Each of those numbers carries its own ±7 % band, and the bands are where the discipline lives. Translate the coefficient into per-axis tolerances and the plant's whole job becomes a table of lines it must not cross:

AXIS · RECEPTOR	TARGET	Σ (7 %)	95 % BAND (±2Σ)
H · TRPV1	4,200	294	3,612 – 4,788
P · TRPA1	252	18	216 – 288
C · TRPM8	800	56	688 – 912
U · mGluR1	1,800	126	1,548 – 2,052
S · TAS1R2·R3	1,100	77	946 – 1,254
B · TAS2R38	2,200	154	1,892 – 2,508
K · CaSR	1,900	133	1,634 – 2,166
A · salivary PRPs	1,400	98	1,204 – 1,596

Now the statistics turn unforgiving in a way that is worth sitting with. Suppose each axis is independently in control and we accept a lot only when all eight fall within their ±2σ bands. Any one axis clears a two-sigma gate about 95.45 % of the time — but eight independent axes must *all* clear it, and independent probabilities multiply:

> $0.9545^8 = \mathbf{0.689}$.

Hold eight lines at a two-sigma acceptance window and roughly a third of lots fail the gate, not because any single axis is poor but because there are eight chances to drift. Widen the gate to ±3σ and the arithmetic relaxes — $0.9973^8 = 0.979$, so ninety-eight lots in a hundred pass — but you have paid for it by accepting ±21 % on every axis, and a palate can sometimes feel that. The real answer is the one the number has been pushing us toward all along: do not loosen the gate, tighten the process, so that ±2σ is already a band no one could taste.

That multiplication assumed the eight axes drift independently, which is itself an idealization worth owning. Several of them draw on shared biosynthetic branches, so their errors are correlated — and a correlation can cut either way, pulling two axes off target together, which is kinder to the joint pass rate, or trading one against another, which is crueler. We do not assume the covariance; we measure it, lot by lot, because a joint promise is only ever as good as the joint distribution standing behind it. Eight axes do not make consistency eight percent harder. They make it, by this calculation, nearly the eighth power harder — and that is why the whole of this book has been about one coefficient.

The ladder, and why loudness is not chemistry

It would be convenient if heat chemistry and heat loudness were the same thing, so that a hotter pepper simply read louder in lockstep. They are close cousins, not twins, and the gap between them is one of the most useful things a formulator can understand. Line up our flagship cultivars — each a tier of increasing predicted heat and computational design scope — and divide predicted SHU by IFI-H to ask how much raw chemistry each point of loudness costs:

CULTIVAR	SKU	PREDICTED SHU	IFI-H	PREDICTED SHU PER IFI POINT
Ember	SS-0087	866,000	1,570	≈ 552,000
Blaze	SS-0091	1,720,000	2,630	≈ 654,000
Firestorm	SS-0094	2,470,000	3,490	≈ 708,000
Phoenix	SS-0096	2,937,948	4,200	≈ 700,000

The cost rises — 552,000 Scoville per point near the warm end, climbing past 700,000 at the top, then flattening. That is not sloppiness in the data; it is the signature of a saturating receptor. The Hill curve that makes TRPV1 so decisive near its half-point also bends over near its top, and calcium flooding through the open channel desensitizes it, so the last million Scoville buys less additional loudness than the first. SHU measures the molecule; IFI measures the hearing of it; and the ear, like every ear, compresses. An index that ignored this would overstate the top of the ladder. Ours is built to bend where the biology bends.

The clearest proof that the two axes are genuinely different lives one rung apart in the catalog. Cinder (SS-0088) carries 1.05 M SHU and reads 2,290 on heat; Kindle (SS-0089) carries *more* chemistry at 1.25 M SHU yet reads *less* loudly at 2,000. More capsaicinoid, quieter fire — because the balance between capsaicin and its blunter cousin dihydrocapsaicin sets the *shape* of the burn, not merely its size, and the shape is what the receptor hears. If you formulate to Scoville alone you will be surprised by that pair. If you formulate to the index, you will not. This is also why consistency has to be specified on the axis the eater actually experiences: we certify loudness, and let the chemistry fall where the genome puts it.

Certifying a lot

None of this means anything until a specific box of fruit, grown in a specific month, is allowed to carry the name. Certification is where the lovely abstractions of CV become a sampling plan. To know a lot's mean to a stated tightness you measure a sample of n fruit; the uncertainty of that estimate is the standard error, σ/√n, and inverting it gives the fruit you must actually cut open. To pin an axis to within a relative margin E at 95 % confidence when the lot's own CV is c,

> n = (1.96 × c / E)².

For c = 7 % and E = 2 %, n = (1.96 × 7 / 2)² = 6.86² ≈ **48 fruit**. Ask for ±1 % instead and the cost quadruples to 189, because precision buys itself by the square. Those forty-eight fruit return a standard error of 7 % / √48 ≈ 1.0 %, so the ninety-five-percent interval on the lot's mean is the ±2 % we asked for and not a hair wider — the sampling plan and the tolerance turn out to be the same sentence, read from two directions. Forty-eight pods to certify a lot to two percent is a real morning's work, and we would rather do it than guess.



The controlled chamber, where temperature, light, and water are held steady so the genome is the only thing left varying.

Two further levers finish the job. The first is agronomy: even a perfect genome expresses itself unevenly under drought, heat spikes, or a hungry soil, so the variation in a lot is really two variations added together — the genetic scatter we have nearly closed and the environmental scatter the field still imposes. A controlled environment is simply the decision to shrink the second term, to hold light and water and temperature steady so that the genome is the loudest thing left in the room. The second lever is blending. Because the scatter of an average falls as the square root of the count, pooling k independent lots narrows the mean's coefficient to CV/√k: four lots take 7 % down to 3.5 %, nine lots to 2.33 %. Where a palate cannot tell two lots apart, the honest move is sometimes to combine them and let the arithmetic of averaging do what the field could not.

We should be candid about where ≤ 7 % sits against the hardest industrial yardsticks. Process-capability indices ask how much room a spec leaves against the scatter: with σ at 7 % of target and a typical food acceptance window of ±20 %, the capability works out to 20 / (3 × 7) ≈ 0.95 — strong for a living crop, yet short of the 1.33 that pharmaceutical manufacturing demands, which at the same window would require CV ≤ 5 %. We do not hide that gap. A pepper is not a tablet; it grew. What we can say is that across a catalog of roughly nineteen thousand products built on some two hundred twenty-five distinct construct designs, in forty-two host species, a seven-percent crop is a remarkable thing to be able to certify at all — and that where a given product needs tighter, blending and environment can carry it the rest of the way.

Walk that all the way to a bottle and the levers meet. Cut Phoenix mash down to a 50,000-SHU sauce and the dilution is 2,937,948 / 50,000 ≈ **59-fold**: the mash contributes just 1.7 % of the formula by heat, and because a coefficient of variation is untouched by any linear dilution — scale a distribution and its mean and standard deviation scale together — a single-lot sauce inherits the mash's 7 % unchanged. Build the mash instead from four pooled lots and the contributed-heat CV falls to 7 % / √4 = 3.5 %; against a ±15 % label tolerance the process capability climbs to 15 / (3 × 3.5) ≈ **1.43**, over the 1.33 line a careful manufacturer wants to see. The genome sets

the ceiling, the blend buys the margin, and the number printed on the back of the bottle becomes a number the person who made it can actually keep.

What the body was promised

Here is the quiet reason all of this matters more than a spec sheet should. Human taste is not an infinitely fine instrument; it obeys a Weber law, detecting a change in intensity only once that change crosses roughly fifteen to twenty percent of what came before. Lay our consistency against that threshold and something close to grace falls out. Two random Phoenix lots differ from one another with a standard deviation of $7\ \% \times \sqrt{2} \approx 9.9\ \%$, so the chance their difference even reaches the fifteen-percent line a tongue could notice is about 13 % — meaning in seven cases out of eight, two different harvests are, to the person eating them, *the same pepper.* Compared against the design target rather than against each other, a lot strays past that perceptible line only about three percent of the time. The variation has not vanished. It has been driven below the floor of human perception, which is the only place consistency ever truly needs to live. People do not trust a number they cannot feel; they trust that the thing was, again, exactly itself.

We have to name the edges, because a promise with no edges is a boast. The index leans on a fixed reference EC_{50}, so our reference standards have to be guarded and re-verified or the whole scale would drift beneath us. Fat binds capsaicin, so an EC_{50} measured in a clean extract is not the EC_{50} delivered by a buttery sauce — the matrix is part of the recipe, and we test in it. Capsaicinoid distributions run a little right-skewed, so at the extreme tails we read the coefficient on log-transformed data rather than pretend a bell curve we do not have. And a receptor tasting the same fire twice desensitizes, so even a panel has to pace itself. These are not apologies; they are the conditions under which the number is true, and stating them is part of keeping it true. The method is public — the S39L+L345G double variant written into Pun1, the model that found it, the twenty-one papers behind them — precisely so the limits can be checked rather than trusted. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.



The research greenhouse at golden hour — where a number becomes a plant, and a plant becomes dinner.

And the nourishment runs underneath all of it, the way it runs under everything we grow, never a claim and never preached, only a value kept. A plant grown the same way, to the same fullness, every cycle, carries not merely the same flavor but the same of everything it was grown to hold — so the body, too, receives what it was promised, lot after lot, with no good years and no lean ones to leave it short. That is what a coefficient of variation held under seven percent finally buys: not a specification, but a kept word. A number that holds is a promise kept to one person we will never meet, in a city we will never see, a year or a decade from now, who will sit down tired at the end of a day and taste exactly, reliably, what we meant — the same kindness we would have cooked by hand if only we could have reached that far. We built an entire quality system to reach that far. This book was only ever the long and careful explanation of how, so that what arrives at your table is, every single time, the same good thing.

Appendix · The spec reference

This is the page we hand someone who wants to check our work. Everything else in this book argues for consistency; this appendix *is* the consistency, written down as numbers you can cite, audit, and build a budget from. It has two jobs. The first is to be the citable reference set — the ladder, the eight rulers, the construct, the vocabulary — so a figure quoted anywhere resolves back to one place. The second is to be a working toolkit: enough of the arithmetic, shown by hand, that a formulator can take any release certificate we ship and reconstruct the tolerance around it. Every axis of every organism is held to a batch coefficient of variation at or under seven percent; the values below show where recent lots actually land, and the math below shows how that seven percent gets spent. Full per-lot certificates of analysis ship with the product. Nothing here is rounded in our favor.



The reference bench — the whole index is a ratio against one fixed thing, and this is where that fixed thing is read and re-read until it holds still.

The standard, stated once — and how to read it

One equation sits under every number in this library, and it is worth setting down slowly because its shape is the reason the whole system holds together:

> $\mathrm{IFI(candidate)} = 1{,}000 \times \mathrm{EC_{50}(reference)} \div \mathrm{EC_{50}(candidate)}.$

$\mathrm{EC_{50}}$ is the half-activating concentration — the amount of a stimulus needed to drive its receptor to half of full response. It lives in the denominator, which is the single most important fact about the

formula: a molecule that opens its channel *at a whisper* has a low EC_{50} and therefore a *high* score. The index does not measure how much of a compound a food carries. It measures how *loudly* that compound speaks at the receptor — potency, not mass. We anchor the reference so that 1,000 is the half-activation of the native wild-type. A candidate at 1,000 is exactly as loud as the reference; one at 2,000 is twice as loud, reaching half-maximal activation at half the concentration; one at 4,200 is 4.2 times as loud.

Because the whole thing is a ratio, we carry EC_{50} in *reference units* — we set EC_{50}(reference) ≡ 1.000 RU by the anchoring convention, and then the candidate's half-activating concentration falls straight out of its score:

> EC_{50}(candidate) = 1,000 ÷ IFI, in reference units.

Our flagship, Phoenix (SS-0096), scores IFI-H 4,200. Invert the formula and its effective EC_{50} is 1,000 ÷ 4,200 = **0.238 RU** — Phoenix reaches half-activation at 23.8 % of the reference concentration. If you want a dimensional grounding, the pharmacology literature places capsaicin's EC_{50} at TRPV1 near 0.7 µM; take that as the anchor and Phoenix lands near 0.167 µM. But hold onto the deeper property, because it is why we trust the index between labs and across years: **the anchor cancels.** Compare any two cultivars and the reference falls away entirely —

> EC_{50}(Ember) ÷ EC_{50}(Phoenix) = IFI(Phoenix) ÷ IFI(Ember) = 4,200 ÷ 1,570 = 2.68.

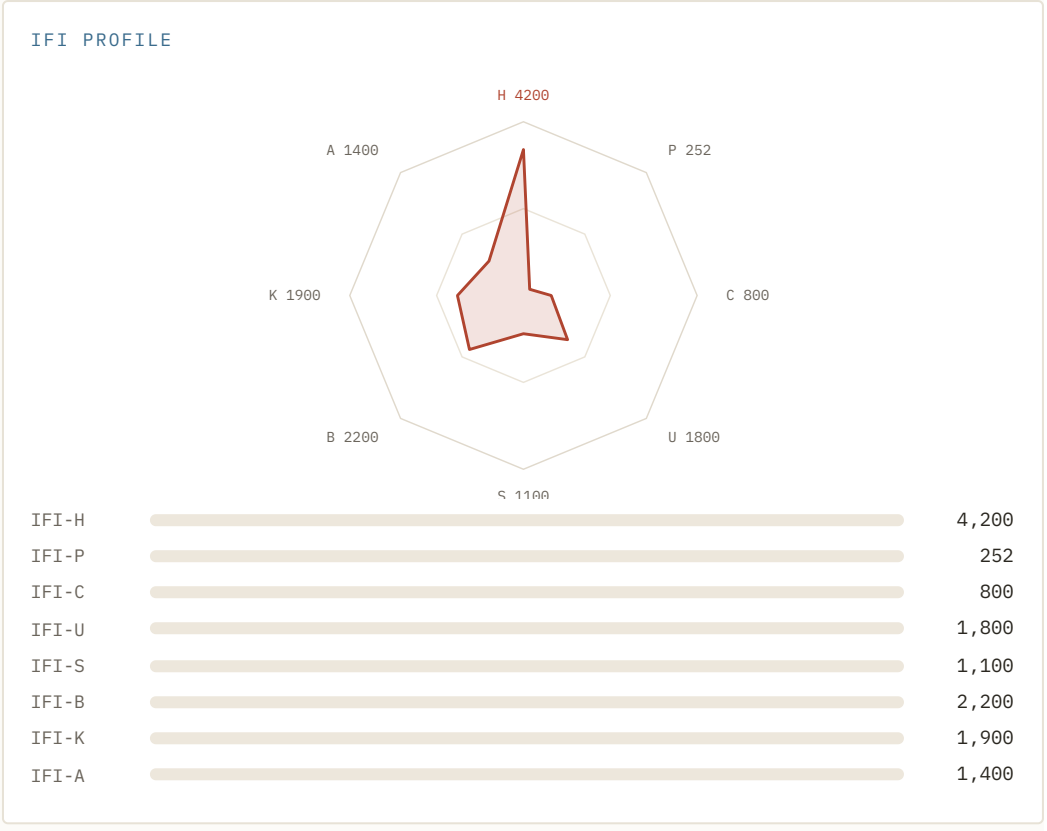
Ember needs 2.68 times the concentration Phoenix needs to say the same thing to the nerve, and that figure is true whether the reference micromolar reading drifts or not. A ratio scale is a scale you can hand to a stranger. For scorecards we normalize each single axis to a ceiling of **5,000**, which sets the page a profile is drawn on — and which Phoenix's heat axis, at 4,200, runs close enough to touch.

The flagship ladder, with its hold

Ten proof-of-concept superhots (*Capsicum chinense*) carry the heat axis from a warning to its ceiling across a progression of increasing computational design scope — Ember activates two of the eight axes, Phoenix all eight — and each one ships against a release band computed from its own batch CV, not from a comfortable round number. The table below is the certificate view: the design target we hold to, the index it reads at, the batch coefficient of variation, and the ±3σ window — µ(1 ± 3·CV) — inside which a lot is the cultivar being itself and outside which it is a lot with a story.

CULTIVAR · SKU	SHU (PREDICTED, M)	IFI-H	BATCH CV	±3Σ RELEASE BAND (SHU)	AXES
Ember · SS-0087	866,000	1,570	2.1 %	811,442 – 920,558	2/8
Cinder · SS-0088	1,050,000	2,290	2.3 %	977,550 – 1,122,450	3/8
Kindle · SS-0089	1,250,000	2,000	2.2 %	1,167,500 – 1,332,500	3/8
Flare · SS-0090	1,480,000	—	—	—	4/8
Blaze · SS-0091	1,720,000	2,630	2.6 %	1,585,840 – 1,854,160	4/8
Wildfire · SS-0092	1,960,000	2,930	2.9 %	1,789,480 – 2,130,520	5/8
Scorch · SS-0093	2,210,000	3,220	3.0 %	2,011,100 – 2,408,900	5/8
Firestorm · SS-0094	2,470,000	3,490	3.2 %	2,232,880 – 2,707,120	6/8
Solara · SS-0095	2,720,000	3,980	3.4 %	2,442,560 – 2,997,440	6/8
Phoenix · SS-0096	2,937,948	4,200	2.4 %	2,726,416 – 3,149,480	**8/8**

Phoenix is the only cultivar that registers on all eight axes at once, and its full profile is the standard every full-spectrum lot is held against:



Two things in that table reward a second look, because they are the honest texture of a real dataset rather than a tidy invented one. The first: the heat climbs but the hold does not climb with it in lockstep — Ember sits at a 2.1 % CV and Solara at 3.4 %, the spread widening as the enzyme is pushed nearer its ceiling, and then Phoenix pulls back to 2.4 % because its release is pooled across independent sublots in a way the single-trait cultivars are not. The second is a genuine inversion worth staring at. Kindle carries *more* raw predicted heat than Cinder — 1,250,000 against 1,050,000 SHU — yet reads *quieter*

at the receptor, IFI-H 2,000 against 2,290. The two numbers measure different things. Scoville counts total capsaicinoid mass; the index weighs the *shape* of that mixture at the channel, and the balance between capsaicin and its slightly milder cousin dihydrocapsaicin shifts potency independent of mass. Near the top of the ladder the two scales ride together at roughly 700 SHU per IFI-H unit — our model's Phoenix prediction of 2,937,948 SHU gives 2,937,948 ÷ 4,200 = 699.5 — but lower down they part ways, and when they disagree we chart the receptor number, because the receptor number is the one the tongue actually reports.

Inverting each score gives the EC_{50} ladder — the concentration, in reference units and in micromolar against the 0.7 µM anchor, at which each cultivar is designed to reach half of full TRPV1 response:

CULTIVAR	IFI-H	EC_{50} (RU)	EC_{50} (mM)	LOUDNESS VS. REFERENCE
Ember	1,570	0.637	0.446	1.57×
Blaze	2,630	0.380	0.266	2.63×
Wildfire	2,930	0.341	0.239	2.93×
Firestorm	3,490	0.287	0.201	3.49×
Solara	3,980	0.251	0.176	3.98×
Phoenix	4,200	0.238	0.167	4.20×

The eight axes and their rulers

Each axis is a receptor, a reference agonist that defines its 1,000, and a coefficient of variation held under the line. The reference agonist is the molecule whose half-activation sets EC_{50}(reference) ≡ 1.000 RU for that channel — the fixed thing every candidate on that axis is a ratio against.

CODE	AXIS	RECEPTOR	REFERENCE AGONIST (ANCHOR)	WHAT ITS CV GUARDS
IFI-H	Heat	TRPV1	capsaicin	the loudest, most robust signal — tightest hold
IFI-S	Sweetness	TAS1R2 · TAS1R3	sucrose	dosed sweetness in reformulation
IFI-U	Umami	mGluR1 · TAS1R1 · TAS1R3	L-glutamate	restored savory depth, can after can
IFI-B	Zero-Bitter	TAS2R38	propylthiouracil (suppressed)	an attenuation precise enough to keep
IFI-K	Kokumi	CaSR	γ-Glu-Val-Gly	the continuity felt more than tasted
IFI-C	Cool-Burn	TRPM8 · TRPV1	menthol	thermomodulation held to target
IFI-P	Pungency	TRPA1	allyl isothiocyanate	the sharp channel, dosed not guessed
IFI-A	Astringency	salivary PRPs	tannic acid	the most delicate finish to hold under the line

Each axis has its own stated convention constant — the reference EC_{50} that defines IFI = 1,000 on that channel. For the pungency axis (IFI-P), that constant is set at 2.1 µM AITC, a stated anchor analogous to Scoville's dilution definition rather than a figure derived from a single literature measurement. Because each axis is a ratio against its own reference constant, scores on different axes are not

directly interchangeable — a 1,200 on the heat axis and a 1,200 on the pungency axis each mean "twice as potent as the respective convention reference," not twice as potent as one another.

Two axes do not behave like simple amplifiers, and the appendix owes you the asterisk. **Zero-Bitter** does not measure an activation at all — it measures a *suppression*, the fractional attenuation of TAS2R38, held across products to a window of **20–65 %**; its index reads the depth of a silence rather than the height of a signal. **Cool-Burn** is bidirectional, reading both the cooling of TRPM8 and the warming of TRPV1, so a score can be tuned to arrive cool or warm from the same ruler. Above the eight single axes sit two composites, which is why a figure can legitimately exceed the 5,000 per-axis ceiling: **IFI-F**, full-spectrum, a single number across all eight channels (the layer behind our 300 full-spectrum products and the nanoparticle delivery work), and **IFI-M**, multi-axis — M2 dual at 8,400, M3 triple at 2,800 — composite scores that are deliberately not subject to the single-axis normalization, because they are reporting a chord, not a note.

Decode Phoenix's profile the same way we decoded the heat ladder — invert each axis to its half-activating concentration — and the shape of a full-spectrum fruit becomes legible:

AXIS	IFI	EC_{50} (RU)	VS. REFERENCE
H · Heat	4,200	0.238	4.20× louder
B · Zero-Bitter	2,200	0.455	2.20×
K · Kokumi	1,900	0.526	1.90×
U · Umami	1,800	0.556	1.80×
A · Astringency	1,400	0.714	1.40×
S · Sweetness	1,100	0.909	1.10×
C · Cool-Burn	800	1.250	0.80× quieter
P · Pungency	252	3.968	0.25× quieter

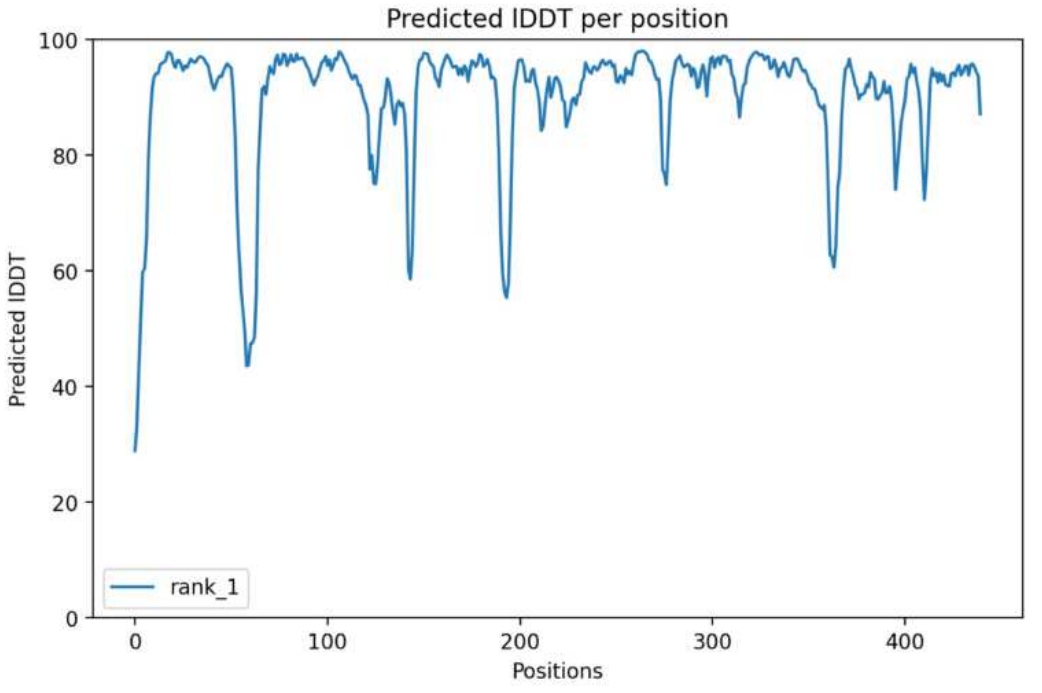
Read the last two rows honestly. Both Cool-Burn (IFI-C 800) and Pungency (IFI-P 252) score *below* their respective references — their EC_{50} values exceed 1.00 RU, meaning Phoenix needs more stimulus than the convention-reference agonists to reach half-activation on those channels, so it speaks both axes *softer* than the anchor, by design. A full-spectrum fruit is not a fruit that shouts on every axis. It is a fruit that *registers* on every axis and then sits each one exactly where the composition wants it, and keeping the quiet axes honestly quiet is as much a precision problem as holding the loud one loud.

The construct, to spec

The hold is not won at the chart or the certificate. It is won in the enzyme, upstream of every lot. The capsaicinoid pathway runs the phenylpropanoid branch — **PAL → C4H → 4CL → HCT → COMT**, present in every vascular plant — into four terminal, pepper-specific steps: **pAMT**, which aminates to vanillylamine; **Pun1/CS**, the

acyltransferase that condenses it with a branched-chain fatty acid; and **KAS** and **BCAT**, which build that acyl chain. Pun1/CS is the terminal capsaicin synthase — it names and seals the final product — but our flux model predicts the rate is governed upstream, at 4CL and HCT/CCoAOMT in the phenylpropanoid pathway, not at Pun1. The *consistency* of the output, however, originates at Pun1: a wild Capsicum left to its own genetics runs that enzyme loose, with a batch-to-batch CV near **42 %** on capsaicinoid output — a plant whose fruit might be searing or merely warm depending on a summer no one controlled.

We did not select our way out of that. An ESM2 650-million-parameter protein language model scored every possible single-residue substitution in Pun1 by predicted effect on catalytic efficiency and fold stability; and two substitutions, never before combined, were written in with CRISPR.



The enzyme the hold is built on — Pun1/AT3 coloured by model confidence; the binding domain and the active-site cleft, where the two edits land, read high-confidence.

EDIT	EFFECT	ΔΔG (KCAL/MOL)
S39L	N-terminal binding affinity ↑	−2.1
L345G	steric relief, active-site cleft	−1.8
Total	**fold stabilization**	**−3.9**

The two edits sum to −3.9 kcal/mol of fold stabilization, and that number is the quiet root of everything downstream. The population of misfolded enzyme relative to folded scales with the Boltzmann factor exp(ΔΔG ÷ RT). At 25 °C, RT = 1.987 × 298.15 = 0.592 kcal/mol, so a −3.9 kcal/mol stabilization drives the misfold-to-fold ratio down by

```
exp(3.9 ÷ 0.592) = exp(6.59) ≈ 727-fold.
```

An enzyme that folds right roughly seven hundred times more reliably makes a product distribution that wanders roughly that much less. In computational terms our model predicts peak output for Phoenix (SS-0096) of approximately 2.95 million SHU — about 18.4 % of the pungency of pure crystalline capsaicin, carried inside a living fruit — while pulling the batch CV from 42 % to **≤ 7 %**. State that improvement the way variance demands, because variation does not add but variance does, and variance is the square of CV:

```
(42 ÷ 7)² = 36.
```

The engineered process carries about **one thirty-sixth** the variance of the wild plant — and that is the figure at the seven-percent cap. Phoenix's released lots run tighter still, at 2.4 %. This is the whole program compressed into a sentence: not a louder fruit, a quieter one. These two edits are among roughly **225** distinct construct designs that back **19,500** product SKUs, every one profiled on the same eight axes across **42** validated host organisms, one thousand of them live and purchasable today, every one held to the same line.

The variance budget — allocating seven percent

The reason the number is seven and not five or ten is not statistical taste. It is physiology. A sense does not read an absolute intensity; it reads a *fractional change* against what it already has, and the smallest fraction it can detect is the Weber fraction. For heat intensity that just-noticeable difference sits near **0.21** — about one part in five. Set the batch CV cap at one third of it, 0.21 ÷ 3 = 0.07, and a full ±3σ spread of in-spec lots — 3 × 7 % = 21 % — fits *entirely inside a single just-noticeable difference.* The consequence is the one that matters at the table: two lots drawn from opposite edges of the spec cannot be told apart by the body eating them. Phoenix, holding at 2.4 %, is tighter yet — its ±3σ spread is 7.2 %, about **0.34** of one perceptible step — a third of a flicker the tongue was never going to notice. Precision, here, is literally the point below which difference stops existing for the eater.

The dose–response curve makes that concrete, and shows the cap doing even more work than the Weber limit promises. Receptor response follows the Hill equation, $f = C^h \div (EC_{50}^h + C^h)$, with the capsaicin–TRPV1 Hill coefficient h near 2, and the index's anchor sits at f = ½, the exact middle of the S. The middle is also the steepest place to stand — which is precisely where a great many real formulations live, because half-max is where a flavor reads as present but not punishing. The sensitivity of response to a proportional change in dose,

```
df ÷ d(ln C) = h · f · (1 − f),
```

reaches its maximum there at h ÷ 4 ≈ 0.5. So a ±7 % swing in concentration propagates to only about a ±3.5 % swing in perceived intensity: the curve itself halves the drift before the tongue ever

reads it. The wild plant's ±42 % collapses to a ±21 % sensation — but only because we first collapsed the chemistry; the curve can halve a spread, it cannot rescue one.

Seven percent is then a budget to be *spent* across the stages where variation enters — and the arithmetic of spending it is the one piece of statistics a formulator most needs, because it is counterintuitive. Independent variances add; coefficients of variation do not:

> $CV^2(\text{total}) = CV^2(\text{input}) + CV^2(\text{process}) + CV^2(\text{assay})$.

Allocate the budget so no single stage dominates — say the biological input (seed-to-fruit expression) at 1.6 %, the process (extraction and formulation) at 1.4 %, and the measurement itself at 1.0 % — and the stages combine not by sum but in quadrature:

> $CV(\text{total}) = \sqrt{(1.6^2 + 1.4^2 + 1.0^2)} = \sqrt{(2.56 + 1.96 + 1.00)} = \sqrt{5.52} = 2.35\ \%$,

which is, to rounding, exactly where Phoenix holds. The table below is the worksheet, and its right column is the lesson: variance, not CV, is the currency that allocates, so the stage with the largest CV spends disproportionately more of the budget.

STAGE	ALLOCATED CV	VARIANCE (CV^2)	SHARE OF TOTAL VARIANCE
Biological input	1.6 %	2.56	46.4 %
Process	1.4 %	1.96	35.5 %
Measurement	1.0 %	1.00	18.1 %
Combined	**2.35 %**	**5.52**	**100 %**

The same identity, run backward, is a safeguard against chasing a ghost. A control reading of 2.35 % is the *observed* spread — the real process convolved with the instrument that measured it — so the plant's own spread is quieter than the chart admits. Ride a fixed reference standard on every plate, learn the assay carries 1.0 % on its own, and subtract in quadrature:

> $CV(\text{process}) = \sqrt{(2.35^2 - 1.0^2)} = \sqrt{(5.52 - 1.00)} = \sqrt{4.52} = 2.13\ \%$.

Thirteen one-hundredths of the observed spread never lived in the fruit at all; it lived in the read. Chase that as though it were the plant and you tamper with a greenhouse while the drift stands in the spectrophotometer.

One more lever, and it is the one Phoenix actually pulls: stacking independent draws. Pool a lot from *n* genuinely independent sublots, each carrying CV *c*, and the pooled mean's CV shrinks as $1 \div \sqrt{n}$:

> $4.8\ \% \div \sqrt{4} = 4.8\ \% \div 2 = 2.4\ \%$.

Four independent sublots that each wander at a wild-ish 4.8 % average down to the 2.4 % Phoenix ships at — the same figure the quadrature budget produced by a different road, which is the quiet satisfaction of a number that is true from two directions at once.



A fixed standard rides every plate — the only way to trust a measurement of variation is to measure the measurement, and keep the reference colder and steadier than anything it judges.

A held CV finally has to be weighed against a customer's window, and the bridge is process capability. For a spec window of width (USL − LSL),

```
Cp = (USL − LSL) ÷ 6σ,
```

the ratio of the room you are given to the room your ordinary scatter needs. Take the natural window — the one the Weber fraction draws — at ±21 %, one just-noticeable difference either side of target, so the window is 42 % of the mean wide while 6σ at a CV is 6·CV·μ. Then Cp = 0.42 ÷ (6·CV) = 0.07 ÷ CV, which turns every cultivar's batch CV directly into a capability index:

CULTIVAR	BATCH CV	CP (WINDOW ±21 %)
Ember	2.1 %	3.33
Kindle	2.2 %	3.18
Phoenix	2.4 %	2.92
Blaze	2.6 %	2.69
Wildfire	2.9 %	2.41
Solara	3.4 %	2.06

Every entry clears the industry's comfortable floor of 1.33 with room to spare, and the seven-percent cap itself is exactly the point where Cp = 0.07 ÷ 0.07 = 1.00 — the process just filling its window, the honest edge of acceptability made into the specification. Capability is not the whole story, because a process can be beautifully capable and still *lean*. Let Phoenix drift 5 % off target inside that ±21 % window and the centered Cp gives way to the one-sided Cpk = (0.21 − 0.05) ÷ (3 × 0.024) = 0.16 ÷ 0.072 = 2.22 — the spread never changed, but the margin on the near wall fell by a quarter. Spread and centering are two different promises, and the budget has to keep both.

The working vocabulary of this book

Every term we have leaned on, set down with the relation that defines it, so the appendix stands on its own as a glossary and no definition has to be hunted for elsewhere.

TERM	SYMBOL / RELATION	WHAT IT NAMES
Coefficient of variation	$CV = \sigma \div \mu$	the dimensionless measure of drift — spread as a fraction of the thing itself
Weber fraction	$\Delta I \div I \approx 0.21$	the fractional change a sense can just detect; why the cap is 7 %
Quadrature	$\sigma^2(\text{total}) = \Sigma\, \sigma_i^2$	variances add, not CVs — the rule that allocates a budget
EC_{50} / Hill equation	$f = C^h \div (EC_{50}^h + C^h)$	the half-activating concentration and the dose–response curve it anchors
The index	$IFI = 1{,}000 \times EC_{50}(\text{ref}) \div EC_{50}(c)$	loudness at the receptor, as a ratio against a fixed anchor
Gauge R&R	$\sigma^2(\text{obs}) = \sigma^2(\text{process}) + \sigma^2(\text{assay})$	repeatability and reproducibility of the measurement itself
Traceability	$EC_{50}(\text{ref}) \equiv 1.000$ RU	the calibration chain back to one fixed reference
Variance budget	$CV^2(\text{total}) = \Sigma\, CV^2_i$	tolerance allocated across input, process, and release
Pooling law	$CV(\text{pooled}) = CV \div \sqrt{n}$	the shrink of CV when equal independent inputs are stacked
Process capability	$C_p = (USL - LSL) \div 6\sigma$; C_{pk} centers it	whether, and how squarely, the spread fits the window
Fold stabilization	ratio $\propto \exp(\Delta\Delta G \div RT)$	the thermodynamics behind an enzyme that stops wandering

*The standard behind every figure: **IFI = 1,000 × EC_{50}(reference) ÷ EC_{50}(candidate)**, computed on the receptor panels, anchored to the master reference set, and held — across all eight axes and forty-two organisms, through approximately 225 distinct construct designs backing 19,500 designed products, one thousand live today — to a batch coefficient of variation at or under seven percent. One spec. One flavor. Batch after batch, in a city we will never see, for someone we will never cook for by hand, who will taste what we designed and find it exactly, reliably, as we meant it. The full quality case is open, as ever, in our twenty-one peer-reviewed papers on Zenodo under CC BY 4.0, and in the three patents filed. The numbers are shown, and shown by hand, because for our family a number you can check is simply what care looks like when it has to survive being repeated. These are computational designs; Scoville Splice's cultivars and constructs have not been experimentally validated in a laboratory.*