Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Iyengar Flavor Index (IFI): A Receptor-Normalized Standard for Cross-Modal Flavor Measurement

Priyanca Ford

scovillesplice@gmail.com

Scoville Splice LLC https://orcid.org/0000-0003-0395-1752

Research Article

Keywords: Iyengar Flavor Index, receptor potency, EC50, reference compound, TAS2R31, TRPV1, TRPA1, taste receptors, ChEMBL, Scoville heat units

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11875881/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

SCOVILLESPLICE.FARM

SCOVILLE SPLICE RESEARCH SERIES

The Iyengar Flavor Index (IFI): A Receptor-Normalized Standard for Cross-Modal Flavor Measurement

A ratio-to-reference score for one compound at one receptor assay: eight axes, a bitter-suppression reference built on TAS2R31, data-derived references for TRPV1 and TRPA1, and a composite score.

P.I. Ford

Scoville Splice LLC · ORCID: 0000-0003-0395-1752

DOI: 10.5281/zenodo.24511320 License: CC BY 4.0 **Patents:** US 64/172,827; US 64/188,197; US 64/192,062; US 64/199,105

Figshare: figshare/35015806 Harvard Dataverse: DVN/JIGL4T OSF: osf.io/cu4pb Hugging Face: capsaicinoid-data

GitHub: paper-04 Website: iyengar-flavor-index

Package record. 104 claims traced to a file or a cited source (64 computed, 20 definition, 8 design assumption, 1 design target, 11 literature value). 16 tables, 5 figures, 42 data files. Data fingerprint `d945ee2fb9f5`. Rerun every result with `bash run_all.sh`.



Pepper greenhouse with a data tablet. Scoville Splice LLC image library.

ABSTRACT

Pungency is rated in Scoville heat units and foods in the glycemic index, and the Iyengar Flavor Index (IFI) adds a score for one compound at one receptor assay. The score is 1,000 × EC50(reference) / EC50(compound), where both EC50 values come from the same assay. The

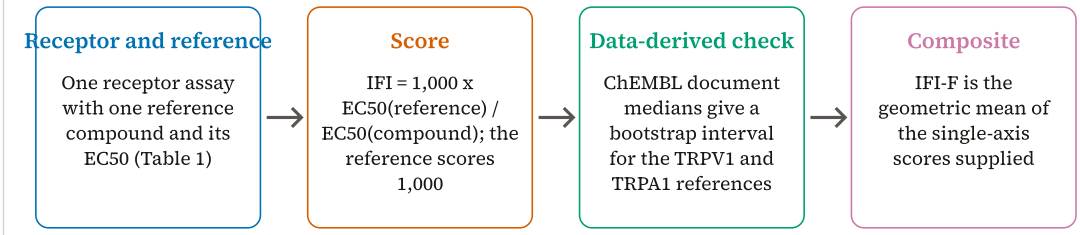
reference compound scores 1,000, a compound twice as potent scores 2,000 and a compound half as potent scores 500. seven single axes each name a receptor and a reference compound: sweet, umami, bitter suppression, kokumi, capsaicinoid pungency, sharp pungency and astringency. A cooling axis (TRPM8) is proposed, and IFI-F is the geometric mean of the single-axis scores supplied. The bitter-suppression axis uses TAS2R31 (gene ID 259290) with the antagonist GIV3727 at a reference of 7.15 µM, the mean of two reported IC50 values. ChEMBL records give data-derived medians of 27.45 nM for capsaicin at TRPV1 (14 documents) and 3,900 nM for allyl isothiocyanate at TRPA1 (9 documents), each with a bootstrap interval. Worked arithmetic, a sensitivity analysis and a proposed measurement series are computed from the files of the package. All results are computed. No experimental data are reported in this paper.

Keywords: Iyengar Flavor Index · receptor potency · EC50 · reference compound · TAS2R31 · TRPV1 · TRPA1 · taste receptors · ChEMBL · Scoville heat units

KEY POINTS

- IFI = 1,000 × EC50(reference) / EC50(compound) scores one compound at one receptor assay; the reference compound scores 1,000.
- The bitter-suppression axis pairs TAS2R31 with GIV3727 at 7.15 µM, the mean of the reported IC50 values 6.4 µM and 7.9 µM.
- ChEMBL records give data-derived medians of 27.45 nM (capsaicin, TRPV1) and 3,900 nM (AITC, TRPA1), with bootstrap intervals.
- IFI-F, the geometric mean of the supplied scores, is computed for worked examples and for the three reference potencies it depends on.
- SHU remains a separate unit for the heat of a product.

Marks under numbers: computed · design target · design assumption · literature value · definition. The evidence table in Appendix A lists the source file of every one.



Graphical abstract. The index in four steps: a receptor and its reference compound, the score formula, the data-derived check of two references and the composite score. Drawn from the steps file. Data: `data/figures/GA.csv`; script: `code/03_data_derived.py`.

1. Introduction

Pungency has a long-used scale. Scoville (1912) rated a pepper extract by the dilution at which heat is no longer sensed, and the result is expressed in Scoville heat units (SHU). The glycemic index (Jenkins et al.,

1981) shows the use of a dimensionless index that rates a food against one fixed reference, glucose. Taste and chemesthesis rest on receptors, each with its own response to a compound: TAS1R2/TAS1R3 for sweet and TAS1R1/TAS1R3 for umami (Li et al., 2002), TAS2R receptors for bitter (Chandrashekar et al., 2006), the calcium-sensing receptor for kokumi (Ohsu et al., 2010), and the ion channels TRPV1, TRPA1 and TRPM8 for capsaicin heat, sharp pungency and cooling (Caterina et al., 1997; Jordt et al., 2004; McKemy et al., 2002).

The Iyengar Flavor Index (IFI) gives each receptor assay its own scale. The EC50 of a reference compound, the concentration that gives half of the maximal response, is divided by the EC50 of the compound under test and multiplied by 1,000. Potent compounds score high, and the reference compound scores 1,000. Every axis is a ratio to its own reference, so scores are dimensionless and each axis keeps the receptor it names.

This paper defines the index, lists its axes with their receptors and reference compounds, and computes the arithmetic that follows from the definition. It builds the reference of the bitter-suppression axis from two reported IC50 values of the antagonist GIV3727 at TAS2R31 (Slack et al., 2010). It derives data-derived reference potencies for capsaicin at TRPV1 and allyl isothiocyanate (AITC) at TRPA1 from ChEMBL (Zdrazil et al., 2024), gives bootstrap intervals, and computes how the composite score IFI-F responds to each reference. A proposed measurement series for every axis is included. The index scores a compound at a receptor assay; it is defined for a compound and not for a cultivar, a construct or a tissue. Companion papers use single axes as references: IFI-U in 10.5281/zenodo.24511324, IFI-S with IFI-C in 10.5281/zenodo.24511316, and the framework in the platform paper 10.5281/zenodo.24511346.

2. Methods

2.1 Definition and scoring

For a compound c on axis a, IFI(c, a) = 1,000 × EC50(reference, a) / EC50(c, a). Both values come from the same receptor assay, in the same unit and for the same type of potency: an agonist EC50 on the agonist axes and an antagonist IC50 on the bitter-suppression axis. IFI-A uses the midpoint of a precipitation titration. A lower EC50 gives a higher score, and the reference compound scores 1,000 on its own axis. The composite IFI-F of a compound is the geometric mean of its positive single-axis scores that are supplied. IFI-F has no reference compound of its own, and the proposed IFI-M is not an input of IFI-F. SHU is a separate unit. All scores are computed with exact fractions (Python `fractions`) and checked by a second implementation in $\log_{10}$ space.

2.2 Axes and receptors

The input file lists nine axes, and eight of them carry a reference compound (Table 1, Table S1). A receptor axis names its human receptor by NCBI Gene symbol and gene ID. The reference compound and its EC50 are stated conventions for the axis, and the basis of each is listed in `inputs/receptor_assays.csv` with the receptor and reference sources. Changing a reference EC50 rescales every score on the axis by one factor, so the ranking of compounds within the axis stays the same.

2.3 Bitter-suppression reference

GIV3727 inhibits TAS2R31 activated by acesulfame K or saccharin, with reported IC50 values of 6.4 µM and 7.9 µM (errors 2.4 and 6.1 µM; three experiments each; Slack et al., 2010). TAS2R31, named TAS2R44 in earlier work, responds to saccharin and acesulfame K with reported EC50 values of 1.1 mM and 2.5 mM (Kuhn et al., 2004). The IFI-B reference EC50 is the arithmetic mean of the two IC50 values, 7.15 µM. The geometric mean, 7.111 µM, and each IC50 alone are used as alternatives in the sensitivity analysis (Table S2).

2.4 Data-derived references

ChEMBL activity records (Zdrazil et al., 2024) of capsaicin at human TRPV1 (target CHEMBL4794) and AITC at human TRPA1 (target CHEMBL6007) were kept when the activity type is EC50 and the relation is '='. The 5 records that ChEMBL flags as potential duplicates are listed apart (Table S5). The median of the records of each document gives one independent value per document. The data-derived reference is the median of the document medians. Its interval runs from the 2.5th to the 97.5th percentile of 200,000 bootstrap medians of the $\log_{10}$ document medians; documents are resampled with replacement with the Python random generator (seed 20261112) and percentiles use linear interpolation. A reference is declared data-derived where at least 5 independent documents exist. The calculation is repeated for the records that name Homo sapiens as the assay organism and for those records plus the flagged duplicates (Table S7). A numpy implementation with 10,000 resamples and a raw-scale implementation with 20,000 resamples are stored in the research record and compared in Table S11.

2.5 Composite score and sensitivity

The score ladder gives 10 compound EC50 values from 0.25 to 4 times the reference EC50 (Table S8). The sensitivity analysis takes a compound whose EC50 at TAS2R31, TRPV1 and TRPA1 equals the stated reference, so its three single-axis scores are 1,000 each. It varies three references one at a time (`one_at_a_time` in `code/common.py`): the IFI-B reference across the two IC50 values, and the IFI-C and IFI-P references across the bootstrap intervals of the ChEMBL medians (all records).

2.6 Proposed measurement series

A measurement series is proposed for each axis (`inputs/assay_design.csv`): HEK293 cells that express one receptor, a promiscuous G-protein subunit for the taste receptors, a calcium fluorescence readout, a series of 10 points with 3-fold steps from 30 times the reference EC50, a four-parameter logistic fit, the reference compound on every plate, and the geometric mean of the scores of 3 independent plates. IFI-A is read as a precipitation titration. These settings are design assumptions with no source. The series spans 19,683-fold in concentration, and Table S10 lists its concentrations for every axis.

2.7 Software

Scripts run on Python 3.9 or newer with the standard library only. The numpy bootstrap in the research record uses numpy 2.5.4. Every table and figure is written from a CSV file in `data/` by the scripts in `code/`.

3. Results

3.1 Axes and reference compounds

The index has seven single axes, each with a receptor assay and a reference compound, the proposed cooling axis IFI-M (TRPM8, menthol at 0.1 µM) and the composite IFI-F (Table 1). The reference EC50 values run from 0.004 µM for brazzein at TAS1R2/TAS1R3 to 2,500 µM for L-glutamate at TAS1R1/TAS1R3 (Figure 1). The other references are 7.15 µM for GIV3727 at TAS2R31, 100 µM for gamma-Glu-Val-Gly at the calcium-sensing receptor, 0.5 µM for capsaicin at TRPV1, 2.1 µM for AITC at TRPA1 and 25 µM for tannic acid in the astringency titration. Brazzein is a heat-stable sweet protein (Ming and Hellekant, 1994).

Table 1. Axes and reference compounds. The nine axes of the index: the receptor (human gene ID) or assay, the reference compound and its reference EC50. IFI-F has no reference of its own; IFI-M is proposed. Data: `data/tables/T01.csv`*; script:* `code/01_reference_table.py`.

Axis	Modality	Receptor	Reference compound	Reference EC50 (µM)
IFI-S	Sweet	TAS1R2 (gene ID 80834) + TAS1R3 (gene ID 83756)	Brazzein	0.004
IFI-U	Umami	TAS1R1 (gene ID 80835) + TAS1R3 (gene ID 83756)	L-Glutamate (monosodium salt)	2500
IFI-B	Bitter suppression	TAS2R31 (gene ID 259290)	GIV3727	7.15
IFI-K	Kokumi	CASR (gene ID 846)	gamma-Glu-Val-Gly	100
IFI-C	Capsaicinoid pungency	TRPV1 (gene ID 7442)	Capsaicin	0.5
IFI-P	Sharp pungency	TRPA1 (gene ID 8989)	Allyl isothiocyanate (AITC)	2.1
IFI-A	Astringency	Salivary proline-rich protein precipitation (no receptor)	Tannic acid	25
IFI-M	Cooling (proposed)	TRPM8 (gene ID 79054)	Menthol	0.1
IFI-F	Composite	all axes supplied	none (no reference of its own)	geometric mean of the supplied scores

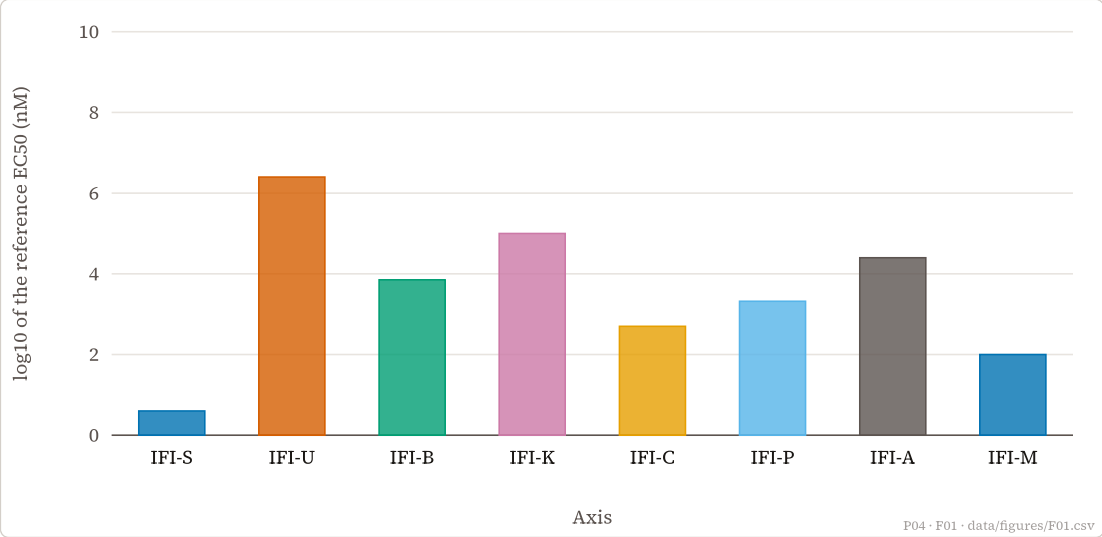


Figure 1. Reference EC50 of each axis. The base-10 logarithm of the reference EC50 of the eight axes that carry a reference compound, in nanomolar. Logarithm of the reference EC50 in Table 1. Data: `data/figures/F01.csv`; script: `code/01_reference_table.py`.

3.2 Worked arithmetic

The arithmetic is the same on every axis (Table 2). A compound twice as potent as the reference has half its EC50 and scores 2,000. A compound half as potent has twice its EC50 and scores 500. On the capsaicin axis, a score of 3,000 corresponds to an EC50 of 0.1667 µM at TRPV1, one third of the reference. Figure 2

and Table S8 show the score for EC50 values from 0.25 to 4 times the reference. A compound at 0.25 times the reference EC50 scores 4,000 on a single axis, and IFI-F with a second axis at 1,000 is 2,000, because the geometric mean moves with the square root of the single-axis change.

Table 2. Worked arithmetic of the score. Each axis: the EC50 of a compound twice as potent as the reference (score 2,000) and half as potent (score 500). Data: `data/tables/T02.csv`*; script:* `code/02_worked_arithmetic.py`.

Axis	Reference EC50 (µM)	EC50 of a compound twice as potent (µM)	IFI	EC50 of a compound half as potent (µM)	IFI
IFI-S	0.004	0.002	2000	0.008	500
IFI-U	2500	1250	2000	5000	500
IFI-B	7.15	3.575	2000	14.3	500
IFI-K	100	50	2000	200	500
IFI-C	0.5	0.25	2000	1	500
IFI-P	2.1	1.05	2000	4.2	500
IFI-A	25	12.5	2000	50	500
IFI-M	0.1	0.05	2000	0.2	500

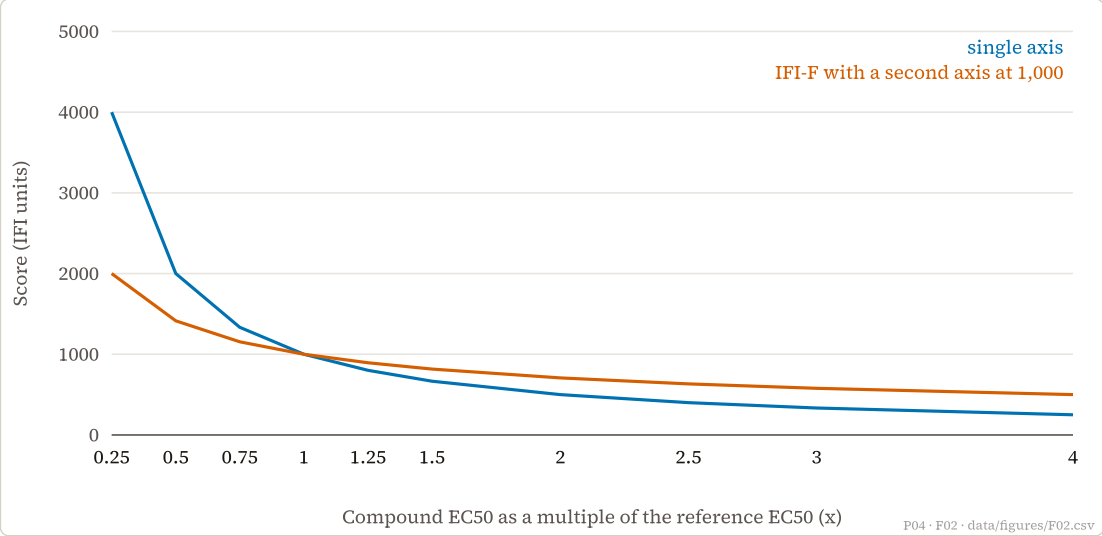


Figure 2. Score of a compound against its EC50. The single-axis score and the IFI-F of a compound whose second axis scores 1,000, for EC50 values from 0.25 to 4 times the reference EC50. Exact formula and geometric mean. Data: `data/figures/F02.csv`; script: `code/02_worked_arithmetic.py`.

3.3 The bitter-suppression reference

The IFI-B reference is 7.15 µM, the arithmetic mean of 6.4 µM (acesulfame K) and 7.9 µM (saccharin) (Table S2). Each value lies 10.5% from the mean, and the geometric mean is 7.111 µM. Table 3 gives the IFI-B scores of five compounds at TAS2R31. A compound with an EC50 of 1 µM scores 7,150, and a compound with an EC50 of 50 µM scores 143.

Table 3. IFI-B scores at TAS2R31. IFI-B of five compounds against the reference of 7.15 µM, the mean of the two reported IC50 values of GIV3727. Data: `data/tables/T03.csv` *; script:* `code/02_worked_arithmetic.py` *.*

Compound	EC50 at TAS2R31 (µM)	IFI-B
Reference compound	7.15	1000
Compound at 3.575 µM	3.575	2000
Compound at 14.3 µM	14.3	500
Compound at 1 µM	1	7150
Compound at 50 µM	50	143

3.4 Data-derived references for TRPV1 and TRPA1

Table 4 gives the data-derived references. For capsaicin at TRPV1, 15 records from 14 documents give a median of 27.45 nM with a 95% interval of 6.00 to 197.48 nM (Figure 3, Figure 4). The reference EC50 of 500 nM is 18.21 times the median, with a ratio interval of 2.53 to 83.33. For AITC at TRPA1, 13 records from 9 documents give a median of 3,900 nM with an interval of 580.00 to 10,471.29 nM; the reference EC50 of 2,100 nM is 0.54 times the median, with a ratio interval of 0.20 to 3.62. Both axes meet the minimum of 5 documents. Records that name Homo sapiens as the assay organism give 13 documents for capsaicin, with a median of 29.00 nM (5.30 to 200.00 nM), and 7 documents for AITC, with a median of 3,900 nM (580.00 to 8,120.00 nM). The input holds 28 records, and three of them carry no assay organism (Table S4). Adding the flagged duplicates gives 16 documents and a median of 27.45 nM for capsaicin, with an interval of 5.30 to 154.92 nM (Table S7). Replacing a reference by its data-derived median rescales every score on the axis by one factor and leaves the ranking within the axis unchanged.

Table 4. Data-derived reference potencies. Median of document-level medians of ChEMBL EC50 records for capsaicin at TRPV1 and AITC at TRPA1, with the 95% bootstrap interval, for all records and for records that name Homo sapiens as the assay organism. Data: `data/tables/T04.csv` *; script:* `code/03_data_derived.py` *.*

Axis	Compound at human receptor (ChEMBL target)	Reference EC50 (nM)	Data-derived median, all records (nM)	95% interval (nM)	Documents	Median, records naming Homo sapiens (nM)	95% interval (nM)	Documents
IFI-C	capsaicin at TRPV1 (CHEMBL4794)	500	27.45	6.00 to 197.48	14	29.00	5.30 to 200.00	13
IFI-P	allyl isothiocyanate at TRPA1 (CHEMBL6007)	2100	3900.00	580.00 to 10471.29	9	3900.00	580.00 to 8120.00	7

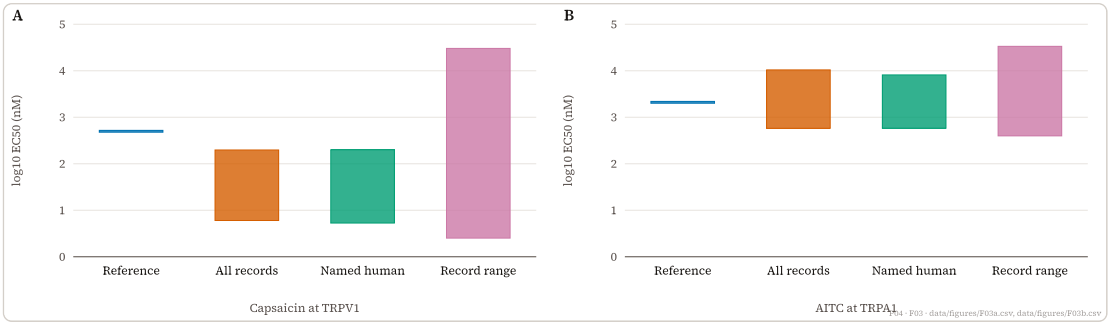


Figure 3. Reference EC50 and the ChEMBL intervals. A, capsaicin at TRPV1. B, AITC at TRPA1. Bars span the 95% bootstrap interval of the median for all records and for records that name Homo sapiens (named human), and the range of all records (record range), in log10 nM; the narrow bar is the reference EC50. Bootstrap of document-level medians. Data: `data/figures/F03a.csv` , `data/figures/F03b.csv` ; script: `code/03_data_derived.py` .

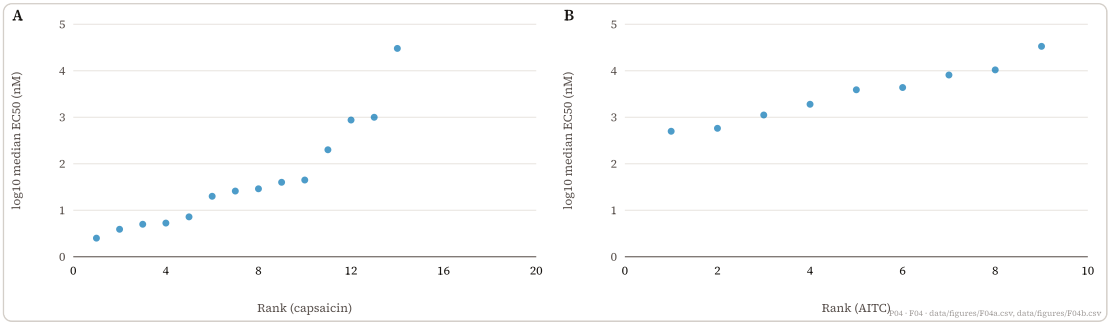


Figure 4. Document-level median EC50 values. A, 14 documents for capsaicin. B, 9 documents for AITC. Each point is the median EC50 of one document, ranked from the lowest. Median of the records of each document. Data: `data/figures/F04a.csv`, `data/figures/F04b.csv`; script: `code/03_data_derived.py`.

3.5 Composite scores

Table 5 gives IFI-F for five sets of single-axis scores. IFI-S of 2,000 and IFI-C of 500 give an IFI-F of 1,000. Raising the IFI-C score to 2,000 gives 2,000. A single supplied score is returned unchanged: IFI-C of 3,000 alone gives 3,000. With four axes supplied (IFI-U 2,000, IFI-K 500, IFI-S 1,500 and IFI-A 800), IFI-F is 1,046.64.

Table 5. Composite scores. IFI-F is the geometric mean of the single-axis scores supplied. Data: `data/tables/T05.csv`; script: `code/02_worked_arithmetic.py`.

Scores supplied	Axes	IFI-F
IFI-S 2,000 and IFI-C 500	2	1000.00
IFI-S 2,000 and IFI-C 2,000	2	2000.00
IFI-B 2,000, IFI-C 500 and IFI-P 1,000	3	1000.00
IFI-C 3,000 alone	1	3000.00
IFI-U 2,000, IFI-K 500, IFI-S 1,500 and IFI-A 800	4	1046.64

3.6 Sensitivity of the composite score

A compound that sits at the stated reference EC50 at TAS2R31, TRPV1 and TRPA1 has an IFI-F of 1,000 (Figure 5, Table S9). When the IFI-B reference moves from 6.4 to 7.9 µM, IFI-F moves from 963.7 to 1,033.8. Across the interval of the capsaicin median, the IFI-C reference moves IFI-F from 228.9 to 733.7, and across the interval of the AITC median, the IFI-P reference moves it from 651.2 to 1,708.4. The two data-derived intervals therefore move the composite score more than the two IC50 values behind the IFI-B reference.

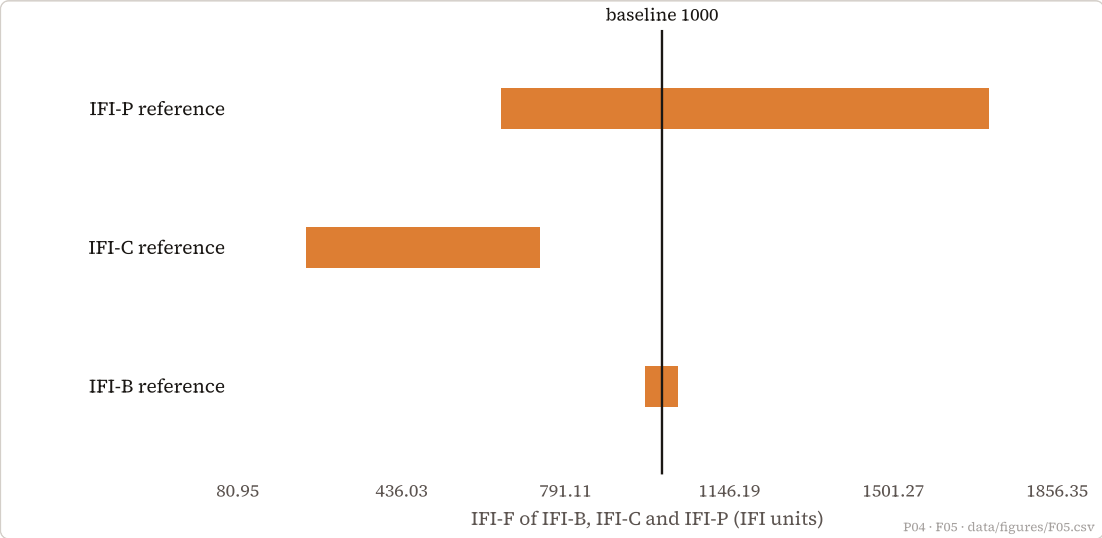


Figure 5. Sensitivity of IFI-F to the reference potencies. IFI-F of a compound whose EC50 at TAS2R31, TRPV1 and TRPA1 equals the stated reference, when one reference moves across its range and the others stay at the stated value: the two IC50 values for IFI-B, and the bootstrap interval of the ChEMBL median (all records) for IFI-C and IFI-P. One-at-a-time analysis of the geometric mean. Data: `data/figures/F05.csv`; script: `code/04_sensitivity_and_supplement.py`.

3.7 Reported potencies and bootstrap agreement

Table S3 lists the reported potencies of the reference compounds. five reported EC50 values of capsaicin at human TRPV1 come from three papers (McIntyre et al., 2001; Gualdani et al., 2024; Maramai et al., 2025) and span 5.3 to 1,900 nM. For menthol at TRPM8, the mouse value is 4.1 µM (Behrendt et al., 2004) and the human values are 81 µM (Journigan et al., 2021) and 118.2 µM (Jahanfar et al., 2022). The three bootstrap implementations of Table S11 give the same medians in all six rows. The numpy implementation with 10,000 resamples gives the same intervals as the primary implementation. The raw-scale implementation gives 6.10 to 200.00 nM for capsaicin (all records) and an upper end of 314.50 nM with the flagged duplicates.

3.8 Measurement series

The proposed series covers 19,683-fold in concentration for every axis (Table S10). For IFI-C it runs from 15 µM down to 0.0007621 µM.

3.9 Scoville heat units

SHU is a separate unit and is not an IFI score. The Phoenix cultivar design (10.5281/zenodo.24511384) carries a heat design target of 2,956,310 SHU. The index assigns no score to a cultivar, a construct or a tissue.

4. Limitations

The reference potencies are stated conventions for each axis. Data-derived values are computed for two axes, from ChEMBL records that combine assay formats and laboratories, and rest on 14 and 9 documents. The IFI-B reference averages two IC50 values obtained with different agonists, each with a large reported error. IFI-M, the astringency axis and the measurement series are proposals. The index scores a compound at a receptor assay and does not state the perceived intensity of a food.

5. Data and code availability

Everything behind this paper is in the data package: the claims table, the data for every table and figure, the scripts that regenerate them, the inputs, the research record and a backup. `bash run_all.sh` reruns every step and compares each regenerated file with the stored file; `bash verify.sh` checks the checksums. The paper and data are licensed under CC BY 4.0 and the code under the MIT licence. The public input datasets are listed in `inputs/input_sources.csv`.

Public records of this paper: Zenodo 10.5281/zenodo.24511320; Figshare: figshare/35015806; Harvard Dataverse: DVN/JIGL4T; OSF: osf.io/cu4pb; Hugging Face: capsaicinoid-data; GitHub: paper-04; Website: iyengar-flavor-index.

Intellectual Property. The engineered constructs, cultivar designs, and methodologies described in this paper are the subject of pending United States patent applications (USPTO Provisional Application Nos. 64/172,827; 64/188,197; 64/192,062; and 64/199,105).

Experimental status. All results reported in this paper are computational. Experimental work on these designs is under way at the authors' laboratory facilities in Montecito, California, and Geneva, Switzerland. Those data are not part of this paper, are shared only with commercial partners under agreement, and are planned for publication next year.

Active experiments for this index include TRPA1 cell assays with allyl isothiocyanate, cinnamaldehyde and related compounds, TRPM8 cell assays with menthol, isopiperitenol and pathway intermediates, a CaSR assay with gamma-glutamyl peptides and glutathione, and the astringency reference assay (salivary protein precipitation).

6. References

1. Behrendt, H. J., Germann, T., Gillen, C., Hatt, H., & Jostock, R. (2004). Characterization of the mouse cold-menthol receptor TRPM8 and vanilloid receptor type-1 VR1 using a fluorometric imaging plate reader (FLIPR) assay. *British Journal of Pharmacology, 141*(4), 737–745. doi:10.1038/sj.bjp.0705652
2. Caterina, M. J., Schumacher, M. A., Tominaga, M., Rosen, T. A., Levine, J. D., & Julius, D. (1997). The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature, 389*(6653), 816–824. doi:10.1038/39807
3. Chandrashekar, J., Hoon, M. A., Ryba, N. J., & Zuker, C. S. (2006). The receptors and cells for mammalian taste. *Nature, 444*(7117), 288–294. doi:10.1038/nature05401
4. Gualdani, R., et al. (2024). TRPV1 corneal neuralgia mutation: enhanced pH response, bradykinin sensitization, and capsazepine rescue. *Proceedings of the National Academy of Sciences, 121*(37). doi:10.1073/pnas.2406186121
5. Jahanfar, S., et al. (2022). Nebivolol as a potent TRPM8 channel blocker: a drug-screening approach. *Membranes, 12*(10), 954. doi:10.3390/membranes12100954
6. Jenkins, D. J., Wolever, T. M., Taylor, R. H., et al. (1981). Glycemic index of foods: a physiological basis for carbohydrate exchange. *American Journal of Clinical Nutrition, 34*(3), 362–366. doi:10.1093/ajcn/34.3.362
7. Jordt, S. E., Bautista, D. M., Chuang, H. H., et al. (2004). Mustard oils and cannabinoids excite sensory nerve fibres through the TRP channel ANKTM1. *Nature, 427*(6971), 260–265. doi:10.1038/nature02282
8. Journigan, V. B., et al. (2021). Structural and in vitro functional characterization of a menthyl TRPM8 agonist and antagonist. *ACS Medicinal Chemistry Letters, 12*(5), 758–767. doi:10.1021/acsmedchemlett.1c00001
9. Kuhn, C., Bufe, B., Winnig, M., Hofmann, T., Frank, O., Behrens, M., ... & Meyerhof, W. (2004). Bitter taste receptors for saccharin and acesulfame K. *Journal of Neuroscience, 24*(45), 10260–10265. doi:10.1523/JNEUROSCI.1225-04.2004
10. Li, X., Staszewski, L., Xu, H., Durick, K., Zoller, M., & Adler, E. (2002). Human receptors for sweet and umami taste. *Proceedings of the National Academy of Sciences, 99*(7), 4692–4696. doi:10.1073/pnas.072090199

11. Maramai, S., et al. (2025). Indole-2-carboxamide as an effective scaffold for the design of new TRPV1 agonists. *Molecules, 30*(3), 721. doi:10.3390/molecules30030721

12. McIntyre, P., McLatchie, L. M., Chambers, A., et al. (2001). Pharmacological differences between the human and rat vanilloid receptor 1 (VR1). *British Journal of Pharmacology, 132*(5), 1084–1094. doi:10.1038/sj.bjp.0703918

13. McKemy, D. D., Neuhausser, W. M., & Julius, D. (2002). Identification of a cold receptor reveals a general role for TRP channels in thermosensation. *Nature, 416*(6876), 52–58. doi:10.1038/nature719

14. Ming, D., & Hellekant, G. (1994). Brazzein, a new high-potency thermostable sweet protein from *Pentadiplandra brazzeana* B. *FEBS Letters, 355*(1), 106–108. doi:10.1016/0014-5793(94)01184-2

15. Ohsu, T., Amino, Y., Nagasaki, H., et al. (2010). Involvement of the calcium-sensing receptor in human taste perception. *Journal of Biological Chemistry, 285*(2), 1016–1022. doi:10.1074/jbc.M109.029165

16. Scoville, W. L. (1912). Note on capsicums. *Journal of the American Pharmaceutical Association, 1*(5), 453–454. doi:10.1002/jps.3080010520

17. Slack, J. P., Brockhoff, A., Batram, C., et al. (2010). Modulation of bitter taste perception by a small molecule hTAS2R antagonist. *Current Biology, 20*(12), 1104–1109. doi:10.1016/j.cub.2010.04.043

18. Zdrazil, B., Felix, E., Hunter, F., et al. (2024). The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Research, 52*(D1), D1180–D1192. doi:10.1093/nar/gkad1004

Supplementary data

The supplementary tables hold the full data behind the paper. Table S1 gives every column of the axis input file. Table S2 derives the IFI-B reference from the two reported IC50 values. Table S3 lists the reported potencies of the reference compounds with their sources. Table S4 lists the 28 ChEMBL records used and Table S5 the 5 records that ChEMBL flags as potential duplicates. Table S6 gives the document-level medians and Table S7 the data-derived references for three record variants. Table S8 is the score ladder, Table S9 the sensitivity of IFI-F, Table S10 the proposed measurement series and Table S11 the agreement of the bootstrap implementations. Every table (S1 to S11) is stored as a CSV file in `data/tables/`, and the scripts that write them are in `code/`. The file map below lists every data, input, code and research-record file.

Table S1. Receptor assay inputs. Every column of the axis input file: receptor, gene ID, reference compound and EC50, potency type, basis, assay design and sources. Data: `data/tables/S01.csv`*; script:* `code/01_reference_table.py`*.*

Axis	Receptor	Gene ID	Reference compound	Reference EC50 (µM)	Potency type	Basis of the reference	Assay design	Receptor source	Reference source
IFI-S	TAS1R2/TAS1R3 heteromer	80834 + 83756	Brazzein	0.004	agonist EC50	stated convention	Cells expressing TAS1R2 and TAS1R3 with a calcium readout	Li et al. 2002 (doi 10.1073/pnas.072090199)	Ming and Hellekant 1994 (doi 10.1016/0014-5793(94)01184-2)
IFI-U	TAS1R1/TAS1R3 heteromer	80835 + 83756	L-Glutamate (monosodium salt)	2500	agonist EC50	stated convention	Cells expressing TAS1R1 and TAS1R3 with a calcium readout	Li et al. 2002 (doi 10.1073/pnas.072090199)	
IFI-B	TAS2R31 (antagonist mode)	259290	GIV3727	7.15	antagonist IC50	mean of two reported IC50 values (6.4 and 7.9 uM)	Cells expressing TAS2R31 with a calcium readout in antagonist mode against saccharin or acesulfame K	Kuhn et al. 2004 (doi 10.1523/JNEUROSCI.1225-04.2004)	Slack et al. 2010 (doi 10.1016/j.cub.2010.04.043)
IFI-K	Calcium-sensing receptor	846	gamma-Glu-Val-Gly	100	agonist EC50	stated convention	Cells expressing CASR with a calcium readout	Ohsu et al. 2010 (doi 10.1074/jbc.M109.029165)	Ohsu et al. 2010 (doi 10.1074/jbc.M109.029165)
IFI-C	TRPV1	7442	Capsaicin	0.5	agonist EC50	stated convention; data-derived median in Table 4	Cells expressing TRPV1 with a calcium readout	Caterina et al. 1997 (doi 10.1038/39807)	Caterina et al. 1997 (doi 10.1038/39807)
IFI-P	TRPA1	8989	Allyl isothiocyanate (AITC)	2.1	agonist EC50	stated convention; data-derived median in Table 4	Cells expressing TRPA1 with a calcium readout	Jordt et al. 2004 (doi 10.1038/nature02282)	Jordt et al. 2004 (doi 10.1038/nature02282)
IFI-A	Salivary proline-rich protein precipitation (no receptor)		Tannic acid	25	precipitation midpoint	stated convention	Protein precipitation titration with a proline-rich protein		
IFI-M	TRPM8	79054	Menthol	0.1	agonist EC50	proposed convention	Cells expressing TRPM8 with a calcium readout	McKemy et al. 2002 (doi 10.1038/nature719); Peier et al. 2002 (doi 10.1016/S0092-8674(02)00652-9)	Behrendt et al. 2004 (doi 10.1038/sj.bjp.0705652)
IFI-F	none		none		geometric mean of supplied single-axis scores	composite			

Table S2. Derivation of the IFI-B reference. The two reported IC50 values of GIV3727 at TAS2R31 (Slack et al., 2010), their means and the spread around the mean. Data: `data/tables/S02.csv`*; script:* `code/01_reference_table.py`*.*

Quantity	Value (µM, or % where stated)	Kind	Reported error (µM)
IC50 against acesulfame K	6.4	reported (n = 3)	2.4
IC50 against saccharin	7.9	reported (n = 3)	6.1
Arithmetic mean of the two IC50 values (the IFI-B reference)	7.15	computed	
Geometric mean of the two IC50 values	7.111	computed	
Half of the difference as a percentage of the mean	10.5	computed	
Reference EC50 in the input file	7.15	input	
Input file equals the arithmetic mean	yes	computed	

Table S3. Reported potencies at the reference receptors. Reported EC50 and IC50 values of capsaicin, menthol, GIV3727, saccharin and acesulfame K, with assay format and source. Data: `data/tables/S03.csv`*; script:* `code/01_reference_table.py`*.*

Compound	Receptor	Potency type	Agonist used	Value as printed	Error as printed	Unit	Value (nM)	Assay format	Source
capsaicin	TRPV1 / VR1 (human)	agonist EC50		0.53	0.03	uM	530	aequorin luminescence	McIntyre et al. 2001, Br J Pharmacol, doi 10.1038/sj.bjp.0703918
capsaicin	TRPV1 / VR1 (human)	agonist EC50		0.19	0.08	uM	190	whole-cell patch clamp (+30 mV)	McIntyre et al. 2001, Br J Pharmacol, doi 10.1038/sj.bjp.0703918
capsaicin	TRPV1 / VR1 (human)	agonist EC50		1.9	0.3	uM	1900	two-electrode voltage clamp	McIntyre et al. 2001, Br J Pharmacol, doi 10.1038/sj.bjp.0703918
capsaicin	TRPV1 (human, WT)	agonist EC50		267	55	nM	267	whole-cell patch clamp	Gualdani et al. 2024, PNAS, doi 10.1073/pnas.2406186121
capsaicin	TRPV1 (human)	agonist EC50		0.0053	0.0004	uM	5.3	calcium imaging, Fluo-4 AM	Maramai et al. 2025, Molecules, doi 10.3390/molecules30030721
menthol	TRPM8 (human)	agonist EC50		81	17	uM	81000	calcium imaging	Journigan et al. 2021, ACS Med Chem Lett, doi 10.1021/acsmedchemlett.1c00001
menthol	TRPM8 (human)	agonist EC50		118.2	14	uM	118200	whole-cell patch clamp (voltage ramp)	Jahanfar et al. 2022, Membranes, doi 10.3390/membranes12100954
(-)-menthol	TRPM8 (MOUSE)	agonist EC50		4.1	1.3	uM	4100	calcium imaging, FLIPR	Behrendt et al. 2004, Br J Pharmacol, doi 10.1038/sj.bjp.0705652
GIV3727	TAS2R31 (human)	antagonist IC50	acesulfame K	6.4	2.4	uM	6400	cell-based receptor assay	Slack et al. 2010, Curr Biol, doi 10.1016/j.cub.2010.04.043
GIV3727	TAS2R31 (human)	antagonist IC50	saccharin	7.9	6.1	uM	7900	cell-based receptor assay	Slack et al. 2010, Curr Biol, doi 10.1016/j.cub.2010.04.043
saccharin	TAS2R31 (human; named TAS2R44 in the source)	agonist EC50		1.1		mM	1100000	cell-based receptor assay	Kuhn et al. 2004, J Neurosci, doi 10.1523/JNEUROSCI.1225-04.2004
acesulfame K	TAS2R31 (human; named TAS2R44 in the source)	agonist EC50		2.5		mM	2500000	cell-based receptor assay	Kuhn et al. 2004, J Neurosci, doi 10.1523/JNEUROSCI.1225-04.2004

Table S4. ChEMBL EC50 records used. The 28 records of capsaicin at TRPV1 and AITC at TRPA1 (relation '=') with assay, organism, document and year. Data: `data/tables/S04.csv`*; script:* `code/03_data_derived.py`*.*

Activity ID	Compound	Gene	EC50 (nM)	Relation	Assay	Assay organism	Document	Year	ChEMBL target
1204433	capsaicin	TRPV1	19.95	=	CHEMBL847558	not recorded	CHEMBL1135551	2002	CHEMBL4794
1951532	capsaicin	TRPV1	40.0	=	CHEMBL889822	Homo sapiens	CHEMBL1146259	2007	CHEMBL4794
2146826	allyl isothiocyanate	TRPA1	33500.0	=	CHEMBL940160	not recorded	CHEMBL1147055	2007	CHEMBL6007
3292346	capsaicin	TRPV1	2.512	=	CHEMBL1121356	Homo sapiens	CHEMBL1158557	2010	CHEMBL4794
3301063	allyl isothiocyanate	TRPA1	1900.0	=	CHEMBL1118242	not recorded	CHEMBL1155439	2010	CHEMBL6007
3501548	allyl isothiocyanate	TRPA1	398.0	=	CHEMBL1259371	Homo sapiens	CHEMBL1255565	2010	CHEMBL6007
3501563	allyl isothiocyanate	TRPA1	22387.21	=	CHEMBL1259374	Homo sapiens	CHEMBL1255565	2010	CHEMBL6007
3501567	allyl isothiocyanate	TRPA1	12022.64	=	CHEMBL1259375	Homo sapiens	CHEMBL1255565	2010	CHEMBL6007
3501571	allyl isothiocyanate	TRPA1	10471.29	=	CHEMBL1259376	Homo sapiens	CHEMBL1255565	2010	CHEMBL6007
3501575	allyl isothiocyanate	TRPA1	5754.4	=	CHEMBL1259377	Homo sapiens	CHEMBL1255565	2010	CHEMBL6007
6328847	allyl isothiocyanate	TRPA1	1120.0	=	CHEMBL1820158	Homo sapiens	CHEMBL1817669	2011	CHEMBL6007
8032045	capsaicin	TRPV1	3.9	=	CHEMBL1954853	Homo sapiens	CHEMBL1949507	2012	CHEMBL4794
13898130	allyl isothiocyanate	TRPA1	500.0	=	CHEMBL3112183	Homo sapiens	CHEMBL3108739	2014	CHEMBL6007
15161504	capsaicin	TRPV1	5.3	=	CHEMBL3398873	Homo sapiens	CHEMBL3396955	2015	CHEMBL4794
15652387	allyl isothiocyanate	TRPA1	580.0	=	CHEMBL3596064	Homo sapiens	CHEMBL3593170	2015	CHEMBL6007
18578233	capsaicin	TRPV1	870.0	=	CHEMBL4222390	Homo sapiens	CHEMBL4219145	2018	CHEMBL4794
18578301	capsaicin	TRPV1	870.96	=	CHEMBL4222390	Homo sapiens	CHEMBL4219145	2018	CHEMBL4794
18589919	capsaicin	TRPV1	1000.0	=	CHEMBL4224182	Homo sapiens	CHEMBL4219255	2018	CHEMBL4794
19019745	allyl isothiocyanate	TRPA1	4365.16	=	CHEMBL4347275	Homo sapiens	CHEMBL4346663	2019	CHEMBL6007
19148118	capsaicin	TRPV1	25.9	=	CHEMBL4379510	Homo sapiens	CHEMBL4376835	2020	CHEMBL4794
19190324	capsaicin	TRPV1	7.2	=	CHEMBL4388810	Homo sapiens	CHEMBL4387732	2019	CHEMBL4794
19400567	allyl isothiocyanate	TRPA1	3900.0	=	CHEMBL4425281	Homo sapiens	CHEMBL4425086	2016	CHEMBL6007
20669535	allyl isothiocyanate	TRPA1	8120.0	=	CHEMBL4620420	Homo sapiens	CHEMBL4619731	2020	CHEMBL6007
22905366	capsaicin	TRPV1	30200.0	=	CHEMBL4732941	Homo sapiens	CHEMBL4732045	2020	CHEMBL4794
22994496	capsaicin	TRPV1	44.8	=	CHEMBL4775807	Homo sapiens	CHEMBL4765488	2021	CHEMBL4794
24785206	capsaicin	TRPV1	29.0	=	CHEMBL5130548	Homo sapiens	CHEMBL5126626	2022	CHEMBL4794
25536712	capsaicin	TRPV1	5.0	=	CHEMBL5342620	Homo sapiens	CHEMBL5338786	2023	CHEMBL4794
29059684	capsaicin	TRPV1	200.0	=	CHEMBL6072919	Homo sapiens	CHEMBL6071163	2024	CHEMBL4794

Table S5. ChEMBL records flagged as potential duplicates. The 5 records that ChEMBL flags as potential duplicates and that are kept out of the main calculation. Data: `data/tables/S05.csv`*; script:* `code/03_data_derived.py`*.*

Activity ID	Compound	Gene	EC50 (nM)	Relation	Assay	Organism	Document	Year	ChEMBL target
2528363	capsaicin	TRPV1	20.0	=	CHEMBL994490	Homo sapiens	CHEMBL1138475	2009	CHEMBL4794
3074188	capsaicin	TRPV1	40.0	=	CHEMBL1068424	Homo sapiens	CHEMBL1153169	2010	CHEMBL4794
22772984	capsaicin	TRPV1	5.3	=	CHEMBL4704866	Homo sapiens	CHEMBL4699615	2020	CHEMBL4794
29059492	capsaicin	TRPV1	1000.0	=	CHEMBL6072897	Homo sapiens	CHEMBL6071163	2024	CHEMBL4794
19019781	allyl isothiocyanate	TRPA1	4000.0	=	CHEMBL4347275	Homo sapiens	CHEMBL4346663	2019	CHEMBL6007

Table S6. Document-level medians. The median EC50 of the records of each document, with the number of records and the base-10 logarithm. Data: `data/tables/S06.csv`*; script:* `code/03_data_derived.py`*.*

Axis	Compound	Document	Year	Records	Median EC50 (nM)	log10 median EC50 (nM)
IFI-C	capsaicin	CHEMBL1135551	2002	1	19.95	1.2999
IFI-C	capsaicin	CHEMBL1146259	2007	1	40.00	1.6021
IFI-C	capsaicin	CHEMBL1158557	2010	1	2.51	0.4000
IFI-C	capsaicin	CHEMBL1949507	2012	1	3.90	0.5911
IFI-C	capsaicin	CHEMBL3396955	2015	1	5.30	0.7243

Axis	Compound	Document	Year	Records	Median EC50 (nM)	log10 median EC50 (nM)
IFI-C	capsaicin	CHEMBL4219145	2018	2	870.48	2.9398
IFI-C	capsaicin	CHEMBL4219255	2018	1	1000.00	3.0000
IFI-C	capsaicin	CHEMBL4376835	2020	1	25.90	1.4133
IFI-C	capsaicin	CHEMBL4387732	2019	1	7.20	0.8573
IFI-C	capsaicin	CHEMBL4732045	2020	1	30200.00	4.4800
IFI-C	capsaicin	CHEMBL4765488	2021	1	44.80	1.6513
IFI-C	capsaicin	CHEMBL5126626	2022	1	29.00	1.4624
IFI-C	capsaicin	CHEMBL5338786	2023	1	5.00	0.6990
IFI-C	capsaicin	CHEMBL6071163	2024	1	200.00	2.3010
IFI-P	allyl isothiocyanate	CHEMBL1147055	2007	1	33500.00	4.5250
IFI-P	allyl isothiocyanate	CHEMBL1155439	2010	1	1900.00	3.2788
IFI-P	allyl isothiocyanate	CHEMBL1255565	2010	5	10471.29	4.0200
IFI-P	allyl isothiocyanate	CHEMBL1817669	2011	1	1120.00	3.0492
IFI-P	allyl isothiocyanate	CHEMBL3108739	2014	1	500.00	2.6990
IFI-P	allyl isothiocyanate	CHEMBL3593170	2015	1	580.00	2.7634
IFI-P	allyl isothiocyanate	CHEMBL4346663	2019	1	4365.16	3.6400
IFI-P	allyl isothiocyanate	CHEMBL4425086	2016	1	3900.00	3.5911
IFI-P	allyl isothiocyanate	CHEMBL4619731	2020	1	8120.00	3.9096

Table S7. Data-derived references for three record variants. Median, interval, record range and ratio to the reference for all records, for records naming Homo sapiens, and for those records with the flagged duplicates. Data: `data/tables/S07.csv`*; script:* `code/03_data_derived.py`*.*

Variant	Axis	Compound	Records	Documents	Median (nM)	Interval low (nM)	Interval high (nM)	Lowest record (nM)	Highest record (nM)	Reference EC50 (nM)	Reference over median	Ratio, low	Ratio, high	At le doc
all records	IFI-C	capsaicin	15	14	27.45	6.00	197.48	2.51	30200.00	500	18.21	2.53	83.33	yes
all records	IFI-P	allyl isothiocyanate	13	9	3900.00	580.00	10471.29	398.00	33500.00	2100	0.54	0.20	3.62	yes
records naming Homo sapiens	IFI-C	capsaicin	14	13	29.00	5.30	200.00	2.51	30200.00	500	17.24	2.50	94.34	yes
records naming Homo sapiens	IFI-P	allyl isothiocyanate	11	7	3900.00	580.00	8120.00	398.00	22387.21	2100	0.54	0.26	3.62	yes
records naming Homo sapiens plus flagged duplicates	IFI-C	capsaicin	18	16	27.45	5.30	154.92	2.51	30200.00	500	18.21	3.23	94.34	yes
records naming Homo sapiens plus flagged duplicates	IFI-P	allyl isothiocyanate	12	7	3900.00	580.00	8120.00	398.00	22387.21	2100	0.54	0.26	3.62	yes

Table S8. Score ladder. The single-axis score and the IFI-F with a second axis at 1,000 for EC50 values from 0.25 to 4 times the reference EC50. Data: `data/tables/S08.csv`*; script:* `code/02_worked_arithmetic.py`*.*

EC50 as a multiple of the reference EC50	Potency as a multiple of the reference	IFI of the single axis	IFI-F with a second axis scoring 1,000
0.25	4	4000	2000.00
0.5	2	2000	1414.21
0.75	1.333333	1333.333	1154.70
1	1	1000	1000.00
1.25	0.8	800	894.43
1.5	0.6666667	666.6667	816.50
2	0.5	500	707.11
2.5	0.4	400	632.46
3	0.3333333	333.3333	577.35
4	0.25	250	500.00

Table S9. Sensitivity of IFI-F. IFI-F of a compound whose EC50 at TAS2R31, TRPV1 and TRPA1 equals the stated reference, when one reference moves across its range (IFI-B: the two IC50 values; IFI-C and IFI-P: the bootstrap interval of the ChEMBL median, all records). Data: `data/tables/S09.csv`*; script:* `code/04_sensitivity_and_supplement.py`*.*

Reference varied	Low reference (µM)	High reference (µM)	IFI-F at the low value	IFI-F at the high value	IFI-F at the stated references
IFI-P reference	0.58	10.471	651.2	1708.4	1000.0
IFI-C reference	0.006	0.19748	228.9	733.7	1000.0
IFI-B reference	6.4	7.9	963.7	1033.8	1000.0

Table S10. Proposed measurement series. Concentrations of the ten-point, three-fold series for every axis, as multiples of the reference EC50 and in micromolar. Data: `data/tables/S10.csv`*; script:* `code/04_sensitivity_and_supplement.py`*.*

Point	Multiple of the reference EC50	IFI-S (µM)	IFI-U (µM)	IFI-B (µM)	IFI-K (µM)	IFI-C (µM)	IFI-P (µM)	IFI-A (µM)	IFI-M (µM)
1	30	0.12	75000	214.5	3000	15	63	750	3
2	10	0.04	25000	71.5	1000	5	21	250	1
3	3.333	0.01333	8333	23.83	333.3	1.667	7	83.33	0.3333
4	1.111	0.004444	2778	7.944	111.1	0.5556	2.333	27.78	0.1111
5	0.3704	0.001481	925.9	2.648	37.04	0.1852	0.7778	9.259	0.03704
6	0.1235	0.0004938	308.6	0.8827	12.35	0.06173	0.2593	3.086	0.01235
7	0.04115	0.0001646	102.9	0.2942	4.115	0.02058	0.08642	1.029	0.004115
8	0.01372	0.00005487	34.29	0.09808	1.372	0.006859	0.02881	0.3429	0.001372
9	0.004572	0.00001829	11.43	0.03269	0.4572	0.002286	0.009602	0.1143	0.0004572
10	0.001524	0.000006097	3.81	0.0109	0.1524	0.0007621	0.003201	0.0381	0.0001524

Table S11. Agreement of bootstrap implementations. Medians and intervals from the primary bootstrap, the numpy bootstrap and the raw-scale alternative for three record variants. Data: `data/tables/S11.csv`*; script:* `code/04_sensitivity_and_supplement.py`.

Variant	Compound	Documents	Median, 200,000 resamples (nM)	Interval (nM)	Median, numpy, 10,000 resamples (nM)	Interval (nM)	Median, raw-scale alternative (nM)	Interval (nM)	Medians equal	Intervals equal (primary and numpy)
all records	capsaicin	14	27.45	6.00 to 197.48	27.45	6.00 to 197.48	27.45	6.10 to 200.00	yes	yes
all records	allyl isothiocyanate	9	3900.00	580.00 to 10471.29	3900.00	580.00 to 10471.29	3900.00	580.00 to 10471.29	yes	yes
records naming Homo sapiens	capsaicin	13	29.00	5.30 to 200.00	29.00	5.30 to 200.00	29.00	5.30 to 200.00	yes	yes
records naming Homo sapiens	allyl isothiocyanate	7	3900.00	580.00 to 8120.00	3900.00	580.00 to 8120.00	3900.00	580.00 to 8120.00	yes	yes
records naming Homo sapiens plus flagged duplicates	capsaicin	16	27.45	5.30 to 154.92	27.45	5.30 to 154.92	27.45	5.30 to 314.50	yes	yes
records naming Homo sapiens plus flagged duplicates	allyl isothiocyanate	7	3900.00	580.00 to 8120.00	3900.00	580.00 to 8120.00	3900.00	580.00 to 8120.00	yes	yes

Supplementary data: file map

Path	Contents
`data/tables/T01.csv`	Table 1: Axes and reference compounds
`data/tables/T02.csv`	Table 2: Worked arithmetic of the score
`data/tables/T03.csv`	Table 3: IFI-B scores at TAS2R31
`data/tables/T04.csv`	Table 4: Data-derived reference potencies
`data/tables/T05.csv`	Table 5: Composite scores
`data/tables/S01.csv`	Table S1: Receptor assay inputs
`data/tables/S02.csv`	Table S2: Derivation of the IFI-B reference
`data/tables/S03.csv`	Table S3: Reported potencies at the reference receptors
`data/tables/S04.csv`	Table S4: ChEMBL EC50 records used
`data/tables/S05.csv`	Table S5: ChEMBL records flagged as potential duplicates
`data/tables/S06.csv`	Table S6: Document-level medians
`data/tables/S07.csv`	Table S7: Data-derived references for three record variants
`data/tables/S08.csv`	Table S8: Score ladder
`data/tables/S09.csv`	Table S9: Sensitivity of IFI-F
`data/tables/S10.csv`	Table S10: Proposed measurement series
`data/tables/S11.csv`	Table S11: Agreement of bootstrap implementations
`data/figures/GA.csv`	Data behind the graphical abstract
`data/figures/F01.csv`	Data behind Figure 1
`data/figures/F02.csv`	Data behind Figure 2
`data/figures/F03a.csv`	Data behind Figure 3
`data/figures/F03b.csv`	Data behind Figure 3
`data/figures/F04a.csv`	Data behind Figure 4
`data/figures/F04b.csv`	Data behind Figure 4
`data/figures/F05.csv`	Data behind Figure 5

Path	Contents
`data/results/ifi_reference_table.csv`	The nine axes with receptor, human gene ID, reference compound, reference EC50 in micromolar and nanomolar, potency type and basis
`data/results/ifib_reference_derivation.csv`	The two reported IC50 values of GIV3727 at TAS2R31, their arithmetic and geometric means and the spread around the mean
`data/results/reported_potency_ranges.csv`	Lowest and highest reported EC50 of capsaicin at human TRPV1 and of menthol at TRPM8 with the number of values and papers
`data/results/worked_arithmetic.csv`	Scores and EC50 values of a compound twice and half as potent as the reference on every axis
`data/results/ifib_example_scores.csv`	IFI-B scores of five example compounds at TAS2R31
`data/results/ifi_c_three_times.csv`	EC50 at TRPV1 that gives IFI-C = 3000
`data/results/score_ladder.csv`	Single-axis score and IFI-F with a second axis at 1,000 for EC50 values from 0.25 to 4 times the reference
`data/results/composite_examples.csv`	IFI-F of five example sets of single-axis scores
`data/results/data_derived_reference.csv`	Median of document medians, bootstrap interval, record range and ratio to the reference for capsaicin at TRPV1 and AITC at TRPA1, for three record variants
`data/results/document_medians.csv`	Median EC50 of the records of each ChEMBL document (all records variant)
`data/results/ifi_f_sensitivity.csv`	IFI-F of a compound at the stated references when one reference moves across its range
`data/results/assay_series.csv`	Concentrations of the proposed ten-point series for each axis, as multiples of the reference EC50 and in micromolar
`data/results/assay_series_span.csv`	Number of points, fold step, top and lowest multiple, and fold span of the series
`data/results/bootstrap_agreement.csv`	Medians and intervals from three bootstrap implementations for three record variants
`data/results/shu_design_target.csv`	Heat design target of the Phoenix cultivar design in SHU and its status as a non-IFI value
`data/figures/*`	Data behind the figures (one CSV per figure or panel)
`data/wet_lab_active.csv`	Experiments active in the authors' laboratories for this paper: register identifier (WL-01 to WL-84), experiment name and stage code (S0 paperwork, S1 calibration and cheap decisive tests, S2 transient and in vitro mechanism tests, S3 stable lines, S4 tests on engineered material, S5 optional and long-lead)
`inputs/receptor_assays.csv`	The nine IFI axes: receptor, human gene ID, reference compound, reference EC50, potency type, basis, assay design and citations
`inputs/literature_potencies.csv`	Reported potencies of the reference compounds at their receptors (capsaicin, menthol, GIV3727, saccharin, acesulfame K) with source and DOI
`inputs/chembl_potency_records.csv`	28 ChEMBL EC50 records (relation '=') for capsaicin at TRPV1 and AITC at TRPA1 with assay, organism and document identifiers
`inputs/chembl_potency_records_flagged_duplicates.csv`	5 ChEMBL EC50 records that ChEMBL flags as potential duplicates (4 capsaicin, 1 AITC)
`inputs/settings.csv`	Score scale, bootstrap settings and the minimum number of documents for a data-derived reference
`inputs/assay_design.csv`	Parameters of the proposed measurement series (design assumptions)
`inputs/shu_design_target.csv`	Heat design target in SHU of the Phoenix cultivar design
`inputs/input_sources.csv`	Public input datasets (ChEMBL and NCBI Gene)
`code/01_reference_table.py`	Axes, reference compounds, IFI-B reference from two reported IC50 values, reported potencies
`code/02_worked_arithmetic.py`	Exact-fraction worked arithmetic, score ladder, IFI-B example scores, composite examples
`code/03_data_derived.py`	Document-level medians and bootstrap intervals of ChEMBL EC50 records (three record variants)
`code/04_sensitivity_and_supplement.py`	Sensitivity of IFI-F to the reference potencies, measurement series, bootstrap agreement table, SHU example
`research_record/scripts/ifi_bootstrap_numpy.py`	Bootstrap of document-level medians with numpy (10,000 resamples, seed 20261112)
`research_record/scripts/ifi_bootstrap_numpy_output.txt`	Printed output of the numpy bootstrap
`research_record/scripts/ifi_exact_arithmetic.py`	Worked arithmetic with exact fractions
`research_record/scripts/ifi_exact_arithmetic_output.txt`	Printed output of the exact arithmetic

Path	Contents
`research_record/outputs/bootstrap_numpy.csv`	Medians and intervals from the numpy bootstrap for the three record variants
`research_record/second_script_rerun/ifi_logspace_check.py`	Second implementation of the arithmetic in log10 space, compared with the stored result files
`research_record/second_script_rerun/ifi_logspace_check_output.txt`	Output of the log10 check
`research_record/second_script_rerun/ifi_bootstrap_alt.py`	Second implementation of the bootstrap on the raw scale (20,000 resamples, nearest-rank percentiles)
`research_record/second_script_rerun/ifi_bootstrap_alt_output.txt`	Printed output of the raw-scale bootstrap
`research_record/second_script_rerun/bootstrap_alt.csv`	Medians and intervals from the raw-scale bootstrap for the three record variants

Cite this paper

Ford, P.I. (2026). The Iyengar Flavor Index (IFI): A Receptor-Normalized Standard for Cross-Modal Flavor Measurement. Scoville Splice Research Series. https://doi.org/10.5281/zenodo.24511320

```
@misc{Ford2026_P04,
  author = {Ford, P.I.},
  title = {The Iyengar Flavor Index (IFI): A Receptor-Normalized Standard for Cross-Modal Flavor Measurement},
  year = {2026},
  publisher = {Scoville Splice LLC},
  series = {Scoville Splice Research Series},
  doi = {10.5281/zenodo.24511320},
  url = {https://doi.org/10.5281/zenodo.24511320}
}
```

Also in `CITATION.cff`, `citation.bib` and `citation.ris` in the data package.

Related papers in this series

ID	Title	DOI
P03	Sweet Heat: In Silico Design of Fruit-Specific Brazzein Expression in Capsicum annuum	10.5281/zenodo.24511316
P05	Umami-Enriched Capsicum: In Silico Design of Glutamate Accumulation by Glutamate Decarboxylase Knockout and Glutamate Dehydrogenase Expression	10.5281/zenodo.24511324
P11	The Capsaicin Design Platform: Architecture, Sixteen Design Tools and the Flavor Index Framework of a Subscription Service	10.5281/zenodo.24511346

Appendix A. Evidence table

Claim ID	Claim	Value	Kind	Source file	Location
P04-C001	The score of the reference compound, and the multiplier of the formula, is 1,000	1,000	definition	`inputs/settings.csv`	row ifi_scale
P04-C002	A compound twice as potent as the reference scores 2,000	2,000	computed	`data/results/worked_arithmetic.csv`	column ifi_twice_as_potent, every row
P04-C003	A compound half as potent as the reference scores 500	500	computed	`data/results/worked_arithmetic.csv`	column ifi_half_as_potent, every row
P04-C004	Seven single axes each carry a receptor assay and a reference compound	seven	computed	`data/results/ifi_reference_table.csv`	rows IFI-S, IFI-U, IFI-B, IFI-K, IFI-C, IFI-P, IFI-A
P04-C005	TAS2R31 is human gene ID 259290	259290	definition	`data/results/ifi_reference_table.csv`	row IFI-B, column ncbi_gene_id
P04-C006	The IFI-B reference EC50 is 7.15 µM, the mean of two reported IC50 values	7.15 µM	computed	`data/results/ifib_reference_derivation.csv`	row Arithmetic mean of the two IC50 values
P04-C007	The data-derived median EC50 of capsaicin at TRPV1 is 27.45 nM	27.45 nM	computed	`data/results/data_derived_reference.csv`	row all records, IFI-C, column median_of_document_medians_nM
P04-C008	The capsaicin median comes from 14 documents	14	computed	`data/results/data_derived_reference.csv`	row all records, IFI-C, column documents
P04-C009	The data-derived median EC50 of AITC at TRPA1 is 3,900 nM	3,900 nM	computed	`data/results/data_derived_reference.csv`	row all records, IFI-P, column median_of_document_medians_nM
P04-C010	The AITC median comes from 9 documents	9	computed	`data/results/data_derived_reference.csv`	row all records, IFI-P, column documents
P04-C011	GIV3727 inhibits TAS2R31 activated by acesulfame K with an IC50 of 6.4 µM	6.4 µM	literature value	`literature`	Slack et al. 2010, doi 10.1016/j.cub.2010.04.043
P04-C012	GIV3727 inhibits TAS2R31 activated by saccharin with an IC50 of 7.9 µM	7.9 µM	literature value	`literature`	Slack et al. 2010, doi 10.1016/j.cub.2010.04.043
P04-C013	The input file lists nine axes	nine	computed	`inputs/receptor_assays.csv`	9 data rows
P04-C014	Eight axes carry a reference compound	eight	computed	`data/results/ifi_reference_table.csv`	rows with a reference EC50
P04-C015	The reported errors of the two IC50 values are 2.4 and 6.1 µM	2.4 and 6.1 µM	literature value	`literature`	Slack et al. 2010, doi 10.1016/j.cub.2010.04.043
P04-C016	Each IC50 comes from three independent experiments	three	literature value	`literature`	Slack et al. 2010, doi 10.1016/j.cub.2010.04.043
P04-C017	Saccharin activates TAS2R31 (named TAS2R44 in the source) with a reported EC50 of 1.1 mM	1.1 mM	literature value	`literature`	Kuhn et al. 2004, doi 10.1523/JNEUROSCI.1225-04.2004
P04-C018	Acesulfame K activates TAS2R31 (named TAS2R44 in the source) with a reported EC50 of 2.5 mM	2.5 mM	literature value	`literature`	Kuhn et al. 2004, doi 10.1523/JNEUROSCI.1225-04.2004
P04-C019	The geometric mean of the two IC50 values is 7.111 µM	7.111 µM	computed	`data/results/ifib_reference_derivation.csv`	row Geometric mean of the two IC50 values
P04-C020	ChEMBL target CHEMBL4794 is human TRPV1	CHEMBL4794	definition	`inputs/chembl_potency_records.csv`	column target_chembl_id
P04-C021	ChEMBL target CHEMBL6007 is human TRPA1	CHEMBL6007	definition	`inputs/chembl_potency_records.csv`	column target_chembl_id
P04-C022	Five records are flagged by ChEMBL as potential duplicates	5	computed	`inputs/chembl_potency_records_flagged_duplicates.csv`	5 data rows
P04-C023	The lower interval end is the 2.5th percentile	2.5th	definition	`inputs/settings.csv`	row interval_low_percentile

Claim ID	Claim	Value	Kind	Source file	Location
P04-C024	The upper interval end is the 97.5th percentile	97.5th	definition	`inputs/settings.csv`	row interval_high_percentile
P04-C025	The bootstrap uses 200,000 resamples	200,000	design assumption	`inputs/settings.csv`	row bootstrap_resamples
P04-C026	The bootstrap random seed is 20261112	20261112	design assumption	`inputs/settings.csv`	row bootstrap_seed
P04-C027	A data-derived reference is declared where at least 5 independent documents exist	5	definition	`inputs/settings.csv`	row minimum_documents
P04-C028	The numpy implementation uses 10,000 resamples	10,000	design assumption	`research_record/scripts/ifi_bootstrap_numpy.py`	size=(10000, ...) in the sampling call
P04-C029	The raw-scale implementation uses 20,000 resamples	20,000	design assumption	`research_record/second_script_rerun/ifi_bootstrap_alt.py`	range(20000)
P04-C030	The score ladder has 10 EC50 values	10	computed	`data/results/score_ladder.csv`	10 data rows
P04-C031	The score ladder starts at an EC50 of 0.25 times the reference	0.25	definition	`data/results/score_ladder.csv`	first row
P04-C032	The score ladder ends at an EC50 of 4 times the reference	4	definition	`data/results/score_ladder.csv`	last row
P04-C033	The sensitivity analysis varies three references	three	computed	`data/results/ifi_f_sensitivity.csv`	3 data rows
P04-C034	The proposed series has 10 points	10	design assumption	`inputs/assay_design.csv`	row dilution_points
P04-C035	The step between points is 3-fold	3	design assumption	`inputs/assay_design.csv`	row dilution_factor
P04-C036	The top concentration is 30 times the reference EC50	30	design assumption	`inputs/assay_design.csv`	row top_concentration_multiple
P04-C037	Scores from 3 independent plates are combined	3	design assumption	`inputs/assay_design.csv`	row independent_plates
P04-C038	The series spans 19,683-fold in concentration	19,683	computed	`data/results/assay_series_span.csv`	column span_fold
P04-C039	Scripts run on Python 3.9 or newer	3.9	definition	`code/LIBRARIES.txt`	line 1
P04-C040	The numpy research script uses numpy 2.5.4	numpy 2.5.4	definition	`code/LIBRARIES.txt`	line 2
P04-C041	The proposed IFI-M reference EC50 of menthol is 0.1 µM	0.1 µM	definition	`data/results/ifi_reference_table.csv`	row IFI-M
P04-C042	The IFI-S reference EC50 of brazzein is 0.004 µM	0.004 µM	definition	`data/results/ifi_reference_table.csv`	row IFI-S
P04-C043	The IFI-U reference EC50 of L-glutamate is 2,500 µM	2,500 µM	definition	`data/results/ifi_reference_table.csv`	row IFI-U
P04-C044	The IFI-K reference EC50 of gamma-Glu-Val-Gly is 100 µM	100 µM	definition	`data/results/ifi_reference_table.csv`	row IFI-K
P04-C045	The IFI-C reference EC50 of capsaicin is 0.5 µM	0.5 µM	definition	`data/results/ifi_reference_table.csv`	row IFI-C
P04-C046	The IFI-P reference EC50 of AITC is 2.1 µM	2.1 µM	definition	`data/results/ifi_reference_table.csv`	row IFI-P
P04-C047	The IFI-A reference tannic acid midpoint is 25 µM	25 µM	definition	`data/results/ifi_reference_table.csv`	row IFI-A
P04-C048	An IFI-C score of 3,000 corresponds to an EC50 of 0.1667 µM at TRPV1	3,000	computed	`data/results/ifi_c_three_times.csv`	row 1, column ifi_c

Claim ID	Claim	Value	Kind	Source file	Location
P04-C049	EC50 behind IFI-C = 3,000	0.1667 µM	computed	`data/results/ifi_c_three_times.csv`	row 1, column ec50_uM
P04-C050	A compound at 0.25 times the reference EC50 scores 4,000 on a single axis	4,000	computed	`data/results/score_ladder.csv`	row 0.25, column ifi_single_axis
P04-C051	IFI-F of that compound with a second axis at 1,000 is 2,000	2,000	computed	`data/results/score_ladder.csv`	row 0.25, column ifi_f_with_second_axis_at_1000
P04-C052	The IFI-B reference EC50, number only	7.15	computed	`data/results/ifib_reference_derivation.csv`	row Arithmetic mean of the two IC50 values
P04-C053	IC50 against acesulfame K, number only	6.4	literature value	`literature`	Slack et al. 2010, doi 10.1016/j.cub.2010.04.043
P04-C054	IC50 against saccharin, number only	7.9	literature value	`literature`	Slack et al. 2010, doi 10.1016/j.cub.2010.04.043
P04-C055	The two IC50 values lie 10.5% either side of their mean	10.5%	computed	`data/results/ifib_reference_derivation.csv`	row Half of the difference as a percentage of the mean
P04-C056	EC50 of 1 µM	1 µM	computed	`data/results/ifib_example_scores.csv`	row Compound at 1 µM
P04-C057	A compound with an EC50 of 1 µM at TAS2R31 scores 7,150	7,150	computed	`data/results/ifib_example_scores.csv`	row Compound at 1 µM
P04-C058	EC50 of 50 µM	50 µM	computed	`data/results/ifib_example_scores.csv`	row Compound at 50 µM
P04-C059	A compound with an EC50 of 50 µM at TAS2R31 scores 143	143	computed	`data/results/ifib_example_scores.csv`	row Compound at 50 µM
P04-C060	The capsaicin median is built from 15 records	15	computed	`data/results/data_derived_reference.csv`	row all records, IFI-C, column records
P04-C061	The 95% bootstrap interval of the capsaicin median is 6.00 to 197.48 nM	6.00 to 197.48 nM	computed	`data/results/data_derived_reference.csv`	row all records, IFI-C, columns interval_low_nM, interval_high_nM
P04-C062	The IFI-C reference EC50 is 500 nM	500 nM	definition	`data/results/data_derived_reference.csv`	row all records, IFI-C, column convention_nM
P04-C063	The capsaicin reference is 18.21 times the data-derived median	18.21	computed	`data/results/data_derived_reference.csv`	row all records, IFI-C, column convention_over_median
P04-C064	The interval of that ratio is 2.53 to 83.33	2.53 to 83.33	computed	`data/results/data_derived_reference.csv`	row all records, IFI-C, columns ratio_low, ratio_high
P04-C065	The AITC median is built from 13 records	13	computed	`data/results/data_derived_reference.csv`	row all records, IFI-P, column records
P04-C066	The 95% bootstrap interval of the AITC median is 580.00 to 10,471.29 nM	580.00 to 10,471.29 nM	computed	`data/results/data_derived_reference.csv`	row all records, IFI-P, columns interval_low_nM, interval_high_nM
P04-C067	The IFI-P reference EC50 is 2,100 nM	2,100 nM	definition	`data/results/data_derived_reference.csv`	row all records, IFI-P, column convention_nM
P04-C068	The AITC reference is 0.54 times the data-derived median	0.54	computed	`data/results/data_derived_reference.csv`	row all records, IFI-P, column convention_over_median
P04-C069	The interval of that ratio is 0.20 to 3.62	0.20 to 3.62	computed	`data/results/data_derived_reference.csv`	row all records, IFI-P, columns ratio_low, ratio_high
P04-C070	Records naming Homo sapiens give 13 documents for capsaicin	13	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens, IFI-C, column documents

Claim ID	Claim	Value	Kind	Source file	Location
P04-C071	The median for those records is 29.00 nM	29.00 nM	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens, IFI-C, column median_of_document_medians_nM
P04-C072	The interval for those records is 5.30 to 200.00 nM	5.30 to 200.00 nM	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens, IFI-C
P04-C073	Records naming Homo sapiens give 7 documents for AITC	7	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens, IFI-P, column documents
P04-C074	The AITC median for those records is 3,900 nM	3,900 nM	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens, IFI-P
P04-C075	The AITC interval for those records is 580.00 to 8,120.00 nM	580.00 to 8,120.00 nM	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens, IFI-P
P04-C076	The input holds 28 ChEMBL records	28	computed	`inputs/chembl_potency_records.csv`	28 data rows
P04-C077	Three of the 28 records carry no assay organism	three	computed	`data/tables/S04.csv`	rows with 'not recorded' in column Assay organism
P04-C078	With the flagged duplicates capsaicin has 16 documents and a median of 27.45 nM	16 documents and a median of 27.45 nM	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens plus flagged duplicates, IFI-C
P04-C079	The interval is then 5.30 to 154.92 nM	5.30 to 154.92 nM	computed	`data/results/data_derived_reference.csv`	row records naming Homo sapiens plus flagged duplicates, IFI-C
P04-C080	IFI-S of 2,000 and IFI-C of 500 give an IFI-F of 1,000	2,000	computed	`data/results/composite_examples.csv`	row 1
P04-C081	IFI-C score in the first composite example	500	computed	`data/results/composite_examples.csv`	row 1
P04-C082	IFI-F of the first example is 1,000	1,000	computed	`data/results/composite_examples.csv`	row 1, column ifi_f
P04-C083	IFI-S 2,000 and IFI-C 2,000 give an IFI-F of 2,000	2,000	computed	`data/results/composite_examples.csv`	row 2
P04-C084	IFI-C 3,000 alone gives an IFI-F of 3,000	3,000	computed	`data/results/composite_examples.csv`	row 4
P04-C085	The last example supplies four axes	four	computed	`data/results/composite_examples.csv`	row 5, column axes_supplied
P04-C086	With IFI-U 2,000, IFI-K 500, IFI-S 1,500 and IFI-A 800 the IFI-F is 1,046.64	1,046.64	computed	`data/results/composite_examples.csv`	row 5, column ifi_f
P04-C087	IFI-F of a compound at the stated references is 1,000	1,000	computed	`data/results/ifi_f_sensitivity.csv`	column baseline_ifi_f
P04-C088	IFI-F is 963.7 when the IFI-B reference is 6.4 µM	963.7	computed	`data/results/ifi_f_sensitivity.csv`	row IFI-B reference, column ifi_f_at_low
P04-C089	IFI-F is 1,033.8 when the IFI-B reference is 7.9 µM	1,033.8	computed	`data/results/ifi_f_sensitivity.csv`	row IFI-B reference, column ifi_f_at_high
P04-C090	IFI-F is 228.9 at the low end of the IFI-C reference interval	228.9	computed	`data/results/ifi_f_sensitivity.csv`	row IFI-C reference, column ifi_f_at_low
P04-C091	IFI-F is 733.7 at the high end of the IFI-C reference interval	733.7	computed	`data/results/ifi_f_sensitivity.csv`	row IFI-C reference, column ifi_f_at_high
P04-C092	IFI-F is 651.2 at the low end of the IFI-P reference interval	651.2	computed	`data/results/ifi_f_sensitivity.csv`	row IFI-P reference, column ifi_f_at_low
P04-C093	IFI-F is 1,708.4 at the high end of the IFI-P reference interval	1,708.4	computed	`data/results/ifi_f_sensitivity.csv`	row IFI-P reference, column ifi_f_at_high
P04-C094	Five reported EC50 values of capsaicin at human TRPV1 come from three papers	five	computed	`data/results/reported_potency_ranges.csv`	row capsaicin, columns values, papers

Claim ID	Claim	Value	Kind	Source file	Location
P04-C095	The reported capsaicin values span 5.3 to 1,900 nM	5.3 to 1,900 nM	computed	`data/results/reported_potency_ranges.csv`	row capsaicin, columns lowest_nM, highest_nM
P04-C096	The mouse TRPM8 EC50 of (-)-menthol is reported as 4.1 µM	4.1 µM	literature value	`literature`	Behrendt et al. 2004, doi 10.1038/sj.bjp.0705652
P04-C097	One human TRPM8 EC50 of menthol is reported as 81 µM	81 µM	literature value	`literature`	Journigan et al. 2021, doi 10.1021/acsmedchemlett.1c00001
P04-C098	Another human TRPM8 EC50 of menthol is reported as 118.2 µM	118.2 µM	literature value	`literature`	Jahanfar et al. 2022, doi 10.3390/membranes12100954
P04-C099	The raw-scale implementation gives 6.10 to 200.00 nM for capsaicin (all records)	6.10 to 200.00 nM	computed	`data/results/bootstrap_agreement.csv`	row all records, capsaicin, column interval_alt_nM
P04-C100	The raw-scale implementation gives an upper end of 314.50 nM for capsaicin with the flagged duplicates	314.50 nM	computed	`data/results/bootstrap_agreement.csv`	row records naming Homo sapiens plus flagged duplicates, capsaicin
P04-C101	For IFI-C the top concentration of the series is 15 µM	15 µM	computed	`data/results/assay_series.csv`	row point 1, column IFI-C_uM
P04-C102	For IFI-C the lowest concentration of the series is 0.0007621 µM	0.0007621 µM	computed	`data/results/assay_series.csv`	row point 10, column IFI-C_uM
P04-C103	The Phoenix cultivar design carries a heat design target of 2,956,310 SHU	2,956,310 SHU	design target	`data/results/shu_design_target.csv`	row SS-0096, column shu_design_target
P04-C104	Supplementary tables S1 to S11 hold the full data	S1 to S11	computed	`data/tables/tables_index.csv`	rows S01 to S11

Appendix B. Glossary, abbreviations and units

Term	Meaning	Type
IFI	Iyengar Flavor Index: 1000 x EC50(reference) / EC50(compound) for one compound at one receptor assay	abbreviation
IFI-S	IFI axis for sweet taste (TAS1R2/TAS1R3; reference brazzein)	abbreviation
IFI-U	IFI axis for umami taste (TAS1R1/TAS1R3; reference L-glutamate)	abbreviation
IFI-B	IFI axis for bitter suppression (TAS2R31 in antagonist mode; reference GIV3727)	abbreviation
IFI-K	IFI axis for kokumi (calcium-sensing receptor; reference gamma-Glu-Val-Gly)	abbreviation
IFI-C	IFI axis for capsaicinoid pungency (TRPV1; reference capsaicin)	abbreviation
IFI-P	IFI axis for sharp pungency (TRPA1; reference allyl isothiocyanate)	abbreviation
IFI-A	IFI axis for astringency (salivary proline-rich protein precipitation; reference tannic acid)	abbreviation
IFI-M	Proposed IFI axis for cooling (TRPM8; reference menthol)	abbreviation
IFI-F	Composite score: the geometric mean of the single-axis scores supplied	abbreviation
EC50	Concentration of an agonist that gives half of the maximal response	abbreviation
IC50	Concentration of an antagonist that lowers the response to an agonist by half	abbreviation
SHU	Scoville heat unit: the unit of pungency of a product	unit
TAS1R1	Taste receptor type 1 member 1 (umami receptor subunit)	abbreviation

Term	Meaning	Type
TAS1R2	Taste receptor type 1 member 2 (sweet receptor subunit)	abbreviation
TAS1R3	Taste receptor type 1 member 3 (shared sweet and umami receptor subunit)	abbreviation
TAS2R31	Taste receptor type 2 member 31 (bitter receptor; formerly named TAS2R44)	abbreviation
CASR	Calcium-sensing receptor	abbreviation
TRPV1	Transient receptor potential cation channel subfamily V member 1 (capsaicin receptor)	abbreviation
TRPA1	Transient receptor potential cation channel subfamily A member 1 (allyl isothiocyanate receptor)	abbreviation
TRPM8	Transient receptor potential cation channel subfamily M member 8 (cold and menthol receptor)	abbreviation
AITC	Allyl isothiocyanate: the pungent compound of mustard	abbreviation
GIV3727	Small-molecule antagonist of TAS2R31 and several other bitter taste receptors	term
gamma-Glu-Val-Gly	A gamma-glutamyl tripeptide that activates the calcium-sensing receptor	term
PRP	Salivary proline-rich protein	abbreviation
ChEMBL	Public database of bioactivity records from the literature	term
HEK293	Human embryonic kidney cell line used to express receptors	abbreviation
Geometric mean	The n-th root of the product of n positive numbers	term
Bootstrap	Resampling of the observed values with replacement to estimate an interval	term
Document-level median	The median EC50 of all records that come from one document	term
Gene ID	Identifier of a gene record in NCBI Gene	term
nM	Nanomolar: 10^-9 mol per litre	unit
µM	Micromolar: 10^-6 mol per litre	unit
mM	Millimolar: 10^-3 mol per litre	unit
IFI units	The dimensionless unit of an IFI score	unit

Appendix C. Computational environment

Item	Value
Python	3.15.2
Platform	Darwin x86_64
Libraries	standard library only

Item	Value
Steps and run time	1 code/01_reference_table.py: 0.1 s; 2 code/02_worked_arithmetic.py: 0.1 s; 3 code/03_data_derived.py: 2.8 s; 4 code/04_sensitivity_and_supplement.py: 0.1 s
Input files	9 files, SHA-256 in `code/environment.json`
Data fingerprint	`d945ee2fb9f5385ae7173c59aa11965d9a7ea472939f62d641ff46de3cca41e5`

LEGAL NOTICES

Research use only. All results in this paper are computed designs, predictions and targets. They are not food, a dietary supplement, a drug or a consumer product, and they are not medical, dietary, agronomic, regulatory or financial advice. Concentrated capsaicinoids can cause severe pain and injury. Do not expose any person or animal to material derived from these designs.

No warranty; limitation of liability. This paper, its data and its code are provided "as is" without warranty of any kind. To the fullest extent permitted by law, Scoville Splice LLC and the authors are not liable for any claim, loss or damage arising from their use. Users are responsible for complying with all biosafety, biosecurity, plant-health, food-safety, import and export, and other laws that apply to them, and for obtaining every approval needed before any laboratory, greenhouse or field work.

Licence. The paper and the data are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0); the code is licensed under the MIT licence. These licences cover copyright in the files only and grant no right under any patent, patent application, trade secret or trademark. See `legal/DISCLAIMER.txt`, `legal/TERMS_OF_USE.txt` and `legal/NOTICE.txt` in the data package.

DOI 10.5281/zenodo.24511320 · US provisional applications 64/172,827, 64/188,197, 64/192,062 and 64/199,105 · © 2026 Scoville Splice LLC and P.I. Ford · Published under CC BY 4.0 · scovillesplice.farm · ORCID: 0000-0003-0395-1752