Quantitative ^{1}H NMR and Chemometric Characterization of Heirloom Apple Juices and of Commercial and Craft Ciders

Apples and their origins, the history and practice of cider making, composition, natural cider, dryness and further scales, nutrition labeling and multivariate modeling

John C. Edwards, Ph.D.
Process NMR Associates, LLC, 84 Patrick Lane, Ste 115, Poughkeepsie, NY 12603, USA
White Paper Series · PNA-WP-CIDER · Version 3.0 · 13 November 2026
Correspondence: john@process-nmr.com · Web: www.process-nmr.com · Blog: www.nmrblog.com
© Process NMR Associates, LLC 2026
All NMR data were generated by the Process NMR Associates NMR laboratory.

Abstract

Proton NMR spectra of commercial and craft ciders and of apple juices, acquired at Process NMR Associates between 2013 and 2026, were consolidated and evaluated by a single quantitative procedure with maleic acid as internal standard. The material comprises 96 ciders of 36 producers, each included once and 87 of them measured both as received and after freeze-drying, and the juices of 56 apple varieties, 53 of them old varieties from one collection. The background is presented first: the origin and classification of the apple and the origins of the varieties examined, the history of cider from the Carolingian estates to the present revival in North America, the cider-making process with the marks that each of its steps leaves in the spectrum, the natural ciders of the Basque Country and Asturias, the designations of dryness and the rules of nutrition labeling.

In the apple juices fructose, glucose, sucrose, sorbitol, malic, quinic and citramalic acids, L-rhamnitol, asparagine and the hydroxycinnamates were determined, malic acid by a fit of its complete ABX multiplet. Relative to the three sugars the old varieties contain 59 % fructose, 13 % glucose and 28 % sucrose on average, with sucrose shares from 5 to 52 %, and 3.4 to 10.4 g of malic acid per 100 g of sugars, which places 48 of the 53 lots among the sharps; sorbitol varies more than tenfold and is highest in the russets. Principal component analysis, hierarchical clustering, t-SNE and UMAP of the juice spectra show a continuum governed by the sucrose share, by glucose and sorbitol, and by malic acid. PLS regression, examined for six descriptions of the spectrum and seven scalings in nested cross-validation, reproduces the principal constituents with ratios of performance to deviation of 10 to 28.

In the ciders, ethanol (3.0 to 10.4 % v/v in the alcoholic ciders), fructose, glucose and sucrose (0 to 102 g/L in total), malic, lactic, succinic, acetic, citric and quinic acids, glycerol, sorbitol, 1,3-propanediol and a series of minor constituents were determined. The measured alcohol contents agree with the declared ones to −0.06 ± 0.71 % v/v (n = 71). Malolactic conversion had taken place in 36 ciders, sucrose was the principal sugar in 13 and sorbic acid was found in 24. In nine ciders, among them a natural cider from the Basque Country, lactic acid bacteria had converted glycerol to 1,3-propanediol; the molar balance of the conversion and its relation to acetic acid are documented. On the scale of the American Cider Association 27 ciders are dry, 12 semi-dry, 42 semi-sweet and 15 sweet. The same ciders are placed on the Spanish, French and European wine designations, and seven further scales, for balance, acidity, acid profile, body, aromatic character, alcohol tier and energy, are proposed. Alcohol, carbohydrate and energy of a 12 fl oz serving are presented for all ciders, ranked by brand; the energy of the alcoholic ciders ranges from 115 to 247 kcal, and nutrition panels are calculated from the spectra.

In the chemometric study of the ciders the pre-processing of the spectra was examined systematically; absolute intensities, mean-centering and Pareto scaling were found to be appropriate, and uniform bins and optimized buckets to be equivalent for these data. Principal component analysis, hierarchical clustering, t-SNE and UMAP showed a continuum governed by the amount of residual sugar, its kind and the acid profile, without natural clusters. Six classifiers were compared in nested cross-validation grouped by producer, with permutation tests of the best-of-six statistic. Classes defined by the composition were recognized almost perfectly, mass-market and craft ciders were distinguished with a balanced accuracy above 0.8, also within a single campaign, and declared adjuncts were recognized by their individual markers and not by a common model. PLS regression on the buckets reproduced alcohol, sugars, the principal acids and energy with ratios of performance to deviation above 7 and estimated the total polyols from the native spectrum within 0.6 g/L.

Keywords: ¹H qNMR; cider; hard cider; apple juice; heirloom apples; natural cider; sagardoa; 1,3-propanediol; malolactic conversion; residual sugar; dryness scale; nutrition labeling; chemometrics; pre-processing; principal component analysis; hierarchical clustering; t-SNE; UMAP; support vector machine; random forest; permutation test; PLS regression.

Table of Contents

Abstract 2
1. Introduction 7
2. The Apple: Origin, Classification and Composition 8
2.1 Origin and domestication 8
2.2 Heirloom, dessert and cider apples 8
2.3 Composition of apple juice 9
3. Cider: History, Practice and Chemistry 9
3.1 Cider in Europe through the ages 9
3.2 Cider in North America 11
3.3 The cider-making process 11
3.4 The chemistry of cider making 14
3.5 Natural cider of the Basque Country and Asturias and the bacterial metabolism of glycerol 14
3.6 Dryness and its communication 16
3.7 Nutrition labeling of cider in the United States 17
3.8 ¹H NMR of apple juice, cider and related beverages 17
4. The Juices of 56 Apple Varieties: Composition and Chemometrics 19
4.1 The collection and the origins of the varieties 19
4.2 Samples, spectra and measurement 22
4.3 Sugars, sorbitol and malic acid 28
4.4 Minor constituents 31
4.5 Juices of local fruit pressed in 2015 33
4.6 Implications for cider making 34
4.7 Chemometrics of the juice spectra: choice of pre-processing 35
4.8 Principal components, clustering, t-SNE and UMAP 36
4.9 PLS regression of the composition on the juice spectra 40
5. Materials and Methods 47
5.1 Samples, consolidation and removal of duplicates 47
5.2 Sample preparation 47
5.3 NMR acquisition 47
5.4 Internal standard 48
5.5 Processing, referencing and the common chemical-shift axis 48
5.6 Measurement of peak areas 48
5.7 Calculation of concentrations and bridging of the two tubes 50
5.8 Derived quantities 51
5.9 Chemometric methods 51
6. Spectral Assignment 52
6.1 Pure-component spectra 52
6.2 Assigned spectra of ciders 55
6.3 Additives and the confirmation of sorbic acid 58
7. Performance of the Quantitative Measurement 60
7.1 Line fits and library decomposition 60
7.2 Agreement with declared alcohol content, repeatability and consistency of the two tubes 61
8. Composition of the Ciders 63
8.1 Overview 63
8.2 Ethanol, methanol and higher alcohols 65
8.3 Residual sugars 66
8.4 Organic acids and malolactic conversion 68

8.5 Polyols, 1,3-propanediol and the signature of natural cider....69
8.6 Preservatives, minor markers and polyphenols....72
8.7 Market segments....72
8.8 Variation between lots....73
9. Dryness and Further Scales Derived from the Spectrum....76
9.1 The sugar-based scales of North America....76
9.2 The orchard-based scale....76
9.3 European designations....77
9.4 Further scales that can be read from the same spectrum....78
10. Nutrition Labeling....82
10.1 Calculation of energy and carbohydrate....82
10.2 Alcohol, carbohydrate and energy of a serving, by brand....83
10.3 Comparison with declared values....87
10.4 Label panels calculated from the spectrum....88
11. Chemometrics: Choice of Pre-processing....90
11.1 Design of the comparison....90
11.2 Normalization....91
11.3 Mean-centering and scaling....92
11.4 Representation of the spectrum....92
12. Chemometrics: Unsupervised Analysis....94
12.1 Principal component analysis....94
12.2 Hierarchical clustering....98
12.3 t-SNE and UMAP....99
13. Chemometrics: Classification....101
13.1 Questions, labels and validation....101
13.2 Labels defined by the composition....103
13.3 Labels external to the measurement....104
13.4 The measurement campaign as a control....105
14. Chemometrics: PLS Regression of Composition on the Spectra....107
14.1 Purpose and design....107
14.2 Results....108
14.3 What the models have learned....110
15. Conclusions....113
Appendix A. Ciders Included in the Study....115
Appendix B. Composition of the Ciders by ¹H qNMR....119
Appendix C. Dryness Classes and Nutrition Values....127
Appendix D. Composition of the Juices of 56 Apple Varieties....131
References....137
Summary....151

List of Figures

Figure 1. The cider-making process.... 13
Figure 2. Bacterial metabolism of glycerol in cider.... 15
Figure 3. ¹H NMR spectra (300 MHz, D_2O) of the juices of 56 apple varieties.... 24
Figure 4. Expansions of the juice spectra, varieties A01 to A28.... 25
Figure 5. Expansions of the juice spectra, varieties A29 to A55 and the supermarket Gala and Fuji of 2015.... 26
Figure 6. Assigned ¹H NMR spectrum of an apple juice.... 27
Figure 7. Sugars, malic acid and sorbitol in the 56 apple juices, ordered by the sum of the sugars.... 29
Figure 8. Relations among the principal constituents of the apple juices.... 31
Figure 9. Minor constituents of the apple juices standardized to 110 g/L of sugars.... 32
Figure 10. ¹H NMR spectra (300 MHz, D_2O) of three apple juices measured in 2015.... 34
Figure 11. Choice of pre-processing for the apple juice spectra.... 36
Figure 12. Principal component analysis of the apple juice spectra.... 37
Figure 13. Hierarchical clustering of the apple juice spectra.... 39
Figure 14. Non-linear maps of the apple juice spectra.... 40
Figure 15. Screening of PLS models for the apple juices.... 44
Figure 16. PLS predictions against the ¹H qNMR reference values for twelve constituents of the apple juices.... 45
Figure 17. Regression coefficients of PLS models for six constituents of the apple juices.... 46
Figure 18. Pure-component ¹H NMR spectra (300 MHz, D_2O) of the carbohydrates and polyols found in cider.... 53
Figure 19. Pure-component ¹H NMR spectra (300 MHz, D_2O) of the alcohols and organic acids of cider.... 54
Figure 20. Assigned ¹H NMR spectrum of a cider in the native tube.... 56
Figure 21. Assigned ¹H NMR spectra of two ciders in the freeze-dried tube.... 57
Figure 22. Additives and confirmation of sorbic acid.... 59
Figure 23. Measurement of peak areas.... 61
Figure 24. Performance of the quantitative measurement.... 63
Figure 25. Composition of the 96 ciders by ¹H qNMR, ordered by residual sugar and grouped by dryness class.... 64
Figure 26. Residual sugar of the 96 ciders by ¹H qNMR (fructose, glucose and sucrose, g/L).... 67
Figure 27. Relations among acids and fermentation products in the 96 ciders, by market segment.... 69
Figure 28. Natural cider and the bacterial conversion of glycerol.... 71
Figure 29. Lot-to-lot variation of five Stormalong ciders (fourteen cans from different lots).... 74
Figure 30. Residual sugar against total organic acid for the 96 ciders.... 77
Figure 31. Distribution of the 96 ciders over the classes of the scales in use and of the scales proposed.... 80
Figure 32. Profile cards of six ciders of contrasting style, calculated from the ¹H qNMR composition.... 81
Figure 33. Energy content of the ciders calculated from the ¹H qNMR composition (component basis).... 83
Figure 34. Alcohol content of the 96 ciders by ¹H qNMR (% v/v).... 84
Figure 35. Sugars and polyols of the 96 ciders in a serving of 12 fl oz (g).... 85
Figure 36. Energy of the 96 ciders in a serving of 12 fl oz (kcal, component basis).... 86
Figure 37. Nutrition panels calculated from the ¹H qNMR composition for three ciders.... 89
Figure 38. Comparison of pre-processing options.... 91
Figure 39. Principal component analysis of the freeze-dried tubes of 87 ciders.... 95
Figure 40. Principal component analysis of the native tubes of 96 ciders.... 97
Figure 41. Hierarchical clustering of the freeze-dried tubes of 87 ciders.... 99
Figure 42. t-SNE (top) and UMAP (bottom) embeddings of the freeze-dried tubes of 87 ciders.... 100

Figure 43. Supervised classification of the ciders from optimized buckets....102
Figure 44. Location of the campaign signature....106
Figure 45. PLS regression on optimized buckets: cross-validated predictions against the reference values....109
Figure 46. Regression coefficients of six PLS models against the centers of the buckets (5.8 to 0.7 ppm)....111

List of Tables

Table 1. Classification of cider apples by the composition of the juice (Long Ashton Research Station)....9
Table 2. Definitions of the dryness scales and designations applied to the ciders....16
Table 3. The 56 apple juices: variety, place and date of origin, parentage, principal use and remarks....20
Table 4. Composition of the apple juices: summary statistics and comparison with the literature....30
Table 5. Relative composition of apple juices pressed and measured in 2015....33
Table 6. Mean composition of the three clusters of the apple juice spectra (Ward's method)....38
Table 7. PLS regression of the composition of the apple juices on their spectra....42
Table 8. Measurement campaigns and acquisition parameters....48
Table 9. Signals used for quantification, with analytical windows and the method of area measurement....49
Table 10. Concentrations determined by ¹H qNMR in 96 ciders (mg/L unless stated otherwise)....65
Table 11. Median composition of the ciders by market segment....73
Table 12. Composition of fourteen lots of five Stormalong ciders (g/L unless stated otherwise)....74
Table 13. Classification of the 96 ciders on the three North American dryness scales....76
Table 14. Further scales proposed for the description of cider, all derived from the ¹H qNMR composition....78
Table 15. Energy conversion factors used for the two calculations of energy....82
Table 16. Declared nutrition values and values derived from the ¹H qNMR composition for eight ciders....87
Table 17. Mean composition of the four clusters obtained by Ward's method from the freeze-dried tubes....98
Table 18. Classification of the ciders from optimized buckets....103
Table 19. PLS regression of the composition on optimized buckets....107
Table 20. Polyols predicted from the native spectrum for the nine ciders without a freeze-dried tube....110
Table 21. The 96 ciders included in the study (Appendix A)....115
Table 22. Alcohol, sugars, organic acids and polyols of the 96 ciders (Appendix B)....119
Table 23. Minor constituents and aromatic protons of the 96 ciders (Appendix B)....123
Table 24. Dryness classes and nutrition values of the 96 ciders (Appendix C)....127
Table 25. Sugars, sorbitol and malic acid in the 56 apple juices (Appendix D)....131
Table 26. Minor constituents of the 56 apple juices (Appendix D)....134

1. Introduction

Cider occupies an unusual position among fermented beverages. It is made from a single fruit whose juice is chemically simple, it is fermented much as a white wine is fermented, and it is sold and consumed in the manner of beer. The composition of the finished product is dominated by a short list of compounds: ethanol, the residual sugars fructose, glucose and sucrose, the polyols glycerol and sorbitol, and the organic acids malic, lactic, succinic and acetic. Every one of these compounds gives resolved signals in a one-dimensional ¹H NMR spectrum of the untreated beverage, and the signal area of each is strictly proportional to the number of contributing nuclei. A single spectrum acquired with an internal standard therefore yields, without compound-specific calibration, the quantities on which the sensory character, the dryness designation and the nutrition label of a cider depend.[1–4]

Quantitative ¹H NMR (qNMR) has been applied at Process NMR Associates to commercial and craft ciders since 2013 and to the juices of dessert, heirloom and cider apples since 2015. The results were reported in stages: a technical report on the first sixty ciders,[5] a survey of the juices of fifty-five heirloom apple varieties,[6] and a white paper on the quantitative fingerprinting of cider issued in May 2026.[7] The present document brings the complete body of data together and treats it by a single, documented procedure. The spectra, which had accumulated in eleven partly overlapping workspaces, were traced to their original acquisitions, each cider was included once, and all peak areas were measured with one set of rules. The outcome is a data set of 96 ciders from 36 producers, 87 of which were measured both as received and after freeze-drying, together with the juices of 56 apple varieties, all recorded at 300 MHz. This document replaces the earlier reports as the reference for these data.

Three features distinguish this edition. First, the apples are treated before the ciders. The juices of 56 varieties, 53 of them from a single collection of old varieties, were analyzed quantitatively, the origin and history of every variety were compiled, and the spectra were made the subject of a chemometric study of their own, with principal component analysis, hierarchical clustering, t-SNE, UMAP and PLS regression. Second, the background has been widened. The history of cider from antiquity to the present revival, the cider-making process and the marks that each of its steps leaves in the NMR spectrum, and the natural ciders of the Basque Country and Asturias, in which lactic acid bacteria convert glycerol to 1,3-propanediol, are described before the results. Third, the practical use of the composition is carried further. Each cider is placed on the dryness scales of North America and Europe, a series of additional scales for acidity, acid profile, body, balance, strength and energy is proposed, and the alcohol, carbohydrate and energy values of a serving are presented as ranked charts that carry the brand names and in the form of label panels. The chemometric part, finally, was designed to answer specific questions instead of displaying score plots: which representation of the spectrum (uniform bins, optimized buckets, targeted integrals or fitted concentrations) and which scaling should be used; whether the ciders form natural groups; which properties can be recognized from the spectrum by supervised classifiers under cross-validation that never tests a model on a producer it has been trained on; and how well partial least squares (PLS) regression on bucketed spectra reproduces the quantitative composition.

The document is organized as follows. Sections 2 and 3 give the background on the apple and on cider: history, process, chemistry, the designations of dryness and the regulatory framework for nutrition labeling. Section 4 presents the apple juices, their composition and their chemometrics. Sections 5 to 7 describe the cider samples, the measurement and its performance. Sections 8 to 10 report the composition of the ciders, their placement on the scales of dryness and on the further scales proposed here, and the energy and carbohydrate values that follow from the composition. Sections 11 to 14 contain the chemometric study of the ciders, and Section 15 the conclusions. Brand and product names are retained throughout because the value of the survey lies in the identity of the products; the results describe the individual bottles and cans that were analyzed, in most cases a single container per product, and are not a statement about a producer's output in general.

2. The Apple: Origin, Classification and Composition

2.1 Origin and domestication

The cultivated apple, *Malus domestica* Borkh., is not a native European fruit. Its principal progenitor is *Malus sieversii*, a wild apple of the Tian Shan mountains of Central Asia that still forms fruit forests in present-day Kazakhstan, and the domesticate was carried westward along the trade routes that later became known as the Silk Road.[8,9] Population-genetic studies showed that the journey left a second signature in the genome: extensive introgression from the European crabapple, *Malus sylvestris*, which was a major secondary contributor to the cultivated gene pool, to the extent that many western cultivars are genetically closer to the European crabapple than to their Asian ancestor.[10] The sequencing of the apple genome and the subsequent re-sequencing of more than one hundred accessions refined this two-step picture and traced the increase in fruit size, firmness and palatability that accompanied selection.[11,12]

Two biological properties of the apple shaped everything that followed. The species is self-incompatible and highly heterozygous, so a seedling never reproduces its parent; and a desirable individual can be perpetuated indefinitely by grafting. A named variety is therefore a clone, and the history of the apple is a history of chance seedlings that were noticed, named and propagated.[8,9] Thousands of such varieties were described in Europe and North America. *The Apples of New York* recorded the varieties grown in a single state at the beginning of the twentieth century,[13] and the modern compilation of Bussey documents the many thousands of names that appeared in the North American literature before 1930.[14] Only a small fraction of them survives in collections, and the identification of old trees in abandoned orchards against repository accessions has become a subject of genetic research in its own right.[15–17]

2.2 Heirloom, dessert and cider apples

The word heirloom has no formal definition. It is applied to varieties that were in cultivation before the consolidation of the commercial industry around a handful of dessert apples in the twentieth century, and it carries the implication of regional origin and of uses that the modern trade no longer serves: long keeping without refrigeration, drying, cooking, and cider.[18,19] Roxbury Russet, regarded as the oldest named American variety, Newtown Pippin, Esopus Spitzenburg, Rhode Island Greening, Baldwin, Northern Spy and Golden Russet all belong to this group, and all of them had a reputation for cider as well as for the table.[13,18,20] The settlement of the Ohio valley added an enormous population of seedling trees. The nurseries of John Chapman were sown from seed collected at cider mills, and fruit of unpredictable quality from such orchards was destined largely for the press.[21]

The old varieties have three kinds of origin, and all three are represented in the collection examined in Section 4. The first is the inheritance of European orcharding. Grafting was practiced by the Romans, Pliny described more than twenty kinds of apple, and monastic and manorial orchards carried named varieties through the Middle Ages; the Pearmain, the earliest apple named in an English record, appears soon after 1200 in a Norfolk tenure that was discharged by the yearly payment of two hundred pearmains and four hogsheads of cider made from them.[8,19,22] Varieties of this kind, among them the Calvilles of France, the russets and pippins of England and the Decio of Italy, were described in the pomologies of the sixteenth to nineteenth centuries, and many are still grown. The second origin is the seedling orchard of colonial North America. The colonists brought seed more often than grafted trees, the seedlings were selected by a climate harsher than that of western Europe, and the best of them were named after the farm or the town where they stood; Roxbury Russet, Rhode Island Greening, Newtown Pippin, Baldwin and Esopus Spitzenburg are examples. The first American treatise on the subject, published by William Coxe in 1817, treats the orchard and the cider house together and judges the varieties by their value for the press as well as for the table.[13,18,23] The third origin is deliberate breeding, which began in the nineteenth century and produced the varieties of the experiment stations in the twentieth; several of these, although not old, are grown today by

the same orchards that preserve the heirlooms. The place and date of origin of each of the varieties analyzed in this study are given in Section 4.1.

European cider districts took a different course and selected varieties specifically for fermentation. The classification devised at the Long Ashton Research Station divides cider apples into four classes according to the acidity and the tannin content of the juice (Table 1).[24–27] Bittersweet and bittersharp varieties, rich in procyanidins and other polyphenols, give the body and astringency of traditional English and French ciders; their juice may contain several grams per liter of polyphenols, whereas dessert fruit rarely exceeds a fraction of that amount.[28–30] Such fruit is unpleasant to eat, and hardly any of it was grown in North America when the present revival of cider began. American producers therefore work predominantly with dessert and culinary fruit, in part with heirloom varieties of high acidity, and to a growing extent with European bittersweet varieties that have been planted since.[25,31–33]

Table 1. Classification of cider apples by the composition of the juice (Long Ashton Research Station).[24,25]

Class	Titratable acidity, as malic acid	Tannin	Character	Examples
Sweet	below 0.45 % (4.5 g/L)	below 0.2 %	low acid, low tannin; blending fruit	Sweet Coppin; most dessert varieties of low acidity
Sharp	above 0.45 %	below 0.2 %	acid, low tannin	Bramley's Seedling; many North American heirloom and culinary varieties
Bittersweet	below 0.45 %	above 0.2 %	low acid, astringent	Dabinett, Yarlington Mill, Michelin
Bittersharp	above 0.45 %	above 0.2 %	acid and astringent	Kingston Black, Stoke Red

2.3 Composition of apple juice

Apple juice is an aqueous solution of three sugars, one sugar alcohol and one dominant acid. Fructose is always the most abundant sugar; glucose and sucrose follow in proportions that vary with variety, maturity and storage, sucrose being hydrolyzed progressively after harvest.[34–37] In the survey of Eisele and Drake, which covered 175 varieties, the mean distribution of the three sugars was 57.7 % fructose, 20.4 % glucose and 21.9 % sucrose, and malic acid amounted on average to 8.6 g per 100 g of sugars.[38] Sorbitol, the form in which members of the Rosaceae transport photosynthate, is present at several grams per liter[39] and is considerably more abundant in pears than in apples,[40] a difference that is of direct analytical interest because sorbitol is not fermented by yeast and persists in cider and perry. L-Malic acid accounts for the greater part of the acidity; quinic, citric and succinic acids are minor constituents, and the balance between sugar and acid distinguishes cultivated from wild fruit.[41–45] The phenolic fraction comprises hydroxycinnamic acid esters, chiefly chlorogenic (5-O-caffeoylquinic) acid, the flavan-3-ols and their oligomers, dihydrochalcones and flavonols.[28–30,46]

All of the major constituents and several of the minor ones are observed directly by ¹H NMR, and apple juice was one of the first foods to which the combination of NMR and multivariate statistics was applied.[47–49] Section 3.8 returns to this literature.

3. Cider: History, Practice and Chemistry

3.1 Cider in Europe through the ages

Fermented apple juice is probably as old as the cultivation of the fruit in temperate Europe, but its early history is thinly documented, because the classical authors did not distinguish it clearly from other fruit wines. Pliny, in the first century, records that wine was made from pears and from all kinds of apples, and the Hebrew *shekar* and the Latin *sicera*, from which the word cider descends, denoted any strong drink other

than the wine of the grape.[20,50] The first unambiguous references are Carolingian. The capitulary *De villis*, issued for the royal estates of Charlemagne about the year 800, requires every steward to keep men who know how to make beer, cider and perry, and it names the two fruit drinks separately, as *pomatium* and *pyratium*.[50,51] By the later Middle Ages cider was established wherever the vine did not thrive: on the Atlantic coast of Spain, in Normandy and Brittany, and in the southern and western counties of England, where the Norman connection brought both varieties and practice.

On the Cantabrian coast the record begins with the orchards. A donation of the year 1014, by which Sancho the Great of Navarre endowed the monastery of Leyre, enumerates apple orchards among the properties and is regarded as the earliest written evidence of the Basque cider culture.[52,53] In the sixteenth and seventeenth centuries the Basque whalers and cod fishermen who worked the waters off Newfoundland carried cider as the daily ration of their crews, two to three liters for each man.[53] In neighboring Asturias cider (*sidra*) has remained the everyday drink to the present day, and the customs that surround it, among them the pouring of the cider from above the head into a wide glass held low, which aerates it at the moment of drinking, were inscribed in 2024 on the Representative List of the Intangible Cultural Heritage of Humanity.[54] The natural ciders of both regions are considered in detail in Section 3.5.

In Normandy and Brittany cider gradually displaced beer as the common drink in the later Middle Ages, aided, according to tradition, by better varieties brought from Biscay, whose origin survives in Norman variety names such as Bisquet. The first printed treatise devoted to the subject, *De vino et pomaceo*, was published in 1588 by the Norman physician Julien Le Paulmier, who set cider above wine as a wholesome drink and described the varieties of the province.[50,55] French practice developed the technique of keeving (*défécation*), in which pectin and nutrients are removed from the juice as a floating gel, so that the fermentation proceeds slowly and can be brought to rest with sugar remaining; the bottled, naturally sparkling *cidre bouché* that results is the model of the French style.[24,27,56]

England contributed the most extensive early literature. Cider appears in manorial accounts from the thirteenth century onward as a product and as a form of payment, and it is named in the English Bible of the Wycliffite translation, a copy of which at Hereford Cathedral is known for that reason as the Cider Bible.[50] The seventeenth century was the age of improvement. Lord Scudamore propagated the Redstreak in Herefordshire, cider was bottled in the strong glass that the coal-fired English furnaces had begun to produce, and in 1662 John Beale reported to the Royal Society that a small quantity of sugar added at bottling made the cider sparkle, one of the earliest descriptions of a secondary fermentation in the bottle. John Evelyn collected this knowledge in *Pomona* (1664), and John Worlidge, in *Vinetum Britannicum* (1676), described the mill, the press and the management of the fermentation.[50,57,58] The eighteenth century added two episodes of wider interest. The endemic colic of Devonshire, described by Huxham in the 1730s, was traced by George Baker in 1767 to lead that the acid cider dissolved from the presses and vessels, an early demonstration of a chemical cause of a dietary disease;[59] and the cider excise of 1763 was met by riots in the cider counties and repealed in 1766. Farm cider remained part of the wages of agricultural laborers until the practice was prohibited in 1887, the year in which H. P. Bulmer began production at Hereford and the industrial phase of English cider making opened. The National Fruit and Cider Institute, founded at Long Ashton near Bristol in 1903, placed cider making on a scientific basis and devised the classification of cider fruit that is still in use (Section 2.2).[24,26,50]

The regions developed the styles that still define traditional cider. The naturally fermented, still and markedly acidic ciders of Asturias and the Basque Country, the bottle-conditioned ciders of Normandy and Brittany, whose residual sweetness is retained by a deliberately slow, nutrient-starved fermentation, and the dry, tannic ciders of the English West Country differ from one another in fruit, microbiology and chemistry.[24,56,60–63] Other districts kept traditions of their own, among them the *Apfelwein* of Hesse and the ciders of Ireland, and present-day European production extends from industrial ciders made from

concentrate to products with protected designations of origin, such as the ciders of the Pays d'Auge and of Cornouaille in France and the natural ciders of Asturias and the Basque Country.[51,64]

3.2 Cider in North America

Colonists brought seeds and scions with them, and for two centuries cider was the everyday drink of the rural North. It could be made on any farm from fruit that was otherwise of little value, it kept through the winter, and it was safer than water. An often-quoted estimate puts the consumption in Massachusetts in 1767 at more than a barrel, about 35 gallons, for every inhabitant, and the quantities consumed in the early republic are documented by Rorabaugh.[20,65,66] Cider served as a means of payment and as an article of trade, that of Newark in New Jersey, made from the Harrison and Campfield apples, being the most celebrated,[23] and it entered politics in 1840, when the supporters of William Henry Harrison made the log cabin and the barrel of hard cider the emblems of his campaign for the presidency.[20] The decline during the nineteenth century had several causes that reinforced one another: the movement of the population into cities, the arrival of lager brewing with German immigration, the temperance movement, which was directed with particular force against the farm cider barrel, and finally national Prohibition, after which few orchards of cider fruit remained.[20,21,25] The word itself changed its meaning. In American usage cider came to denote fresh, unfermented juice, and the fermented product had to be distinguished as hard cider, a convention that is followed in federal regulations to this day.[67]

The revival of the past two decades has been rapid. By 2023 about 1,070 producers were counted in the United States, roughly seven times as many as in 2011, cider was being made commercially in every state, and about half of the producers grew some or all of their own fruit.[68] The structure of the market shifted at the same time. National brands, which had driven the first phase of the expansion, lost volume, while regional producers increased their share of off-premise dollar sales from 39 % in 2019 to 54 % in 2022.[69] The scientific literature followed the industry: surveys of producers and of their research needs,[32,33,70] evaluations of cider cultivars under North American conditions,[30,31] studies of consumer perception and willingness to pay,[71,72] and the development of sensory lexicons for American ciders.[73–75] The ciders examined in this study reflect that market. They range from the sweetened, filtered and carbonated products of national brands to spontaneously fermented, bone-dry and sour ciders made by small producers in New York and New England.

3.3 The cider-making process

However different the styles, all ciders pass through the same sequence of operations, and the differences between them arise from a small number of choices made at each step. The sequence is set out in Figure 1, with the choices that define the principal styles at the left and, at the right, the constituents by which each step can be recognized in the ¹H NMR spectrum of the finished cider. The fruit is harvested ripe, in the traditional districts often after it has fallen, and may be stored for some weeks so that the starch is converted to sugar. It is washed, milled to a pulp and pressed; in the French tradition the pulp is left to macerate for some hours before pressing, which promotes the oxidation of the polyphenols and the release of pectin. The juice may then be sulfited to suppress the wild microflora, clarified with pectolytic enzymes, supplemented with nutrients or, in the opposite sense, deprived of them by keeving.[24–27]

The alcoholic fermentation is conducted either with a selected strain of *Saccharomyces* or spontaneously. In a spontaneous fermentation a succession of yeasts takes part: non-*Saccharomyces* yeasts such as *Hanseniaspora* and *Metschnikowia* in the first days, *Saccharomyces* species during the main phase, and frequently *Brettanomyces* during maturation.[76–79] The malolactic fermentation may follow, accompany or be prevented, according to the style. The young cider is racked from its lees and matured in tank or cask, and it is at this stage, in ciders that are kept on their lees without sulfite, that the bacterial conversion of glycerol described in Section 3.5 takes place. Blending adjusts sugar, acid and tannin; sweetening, where it is

practiced, is carried out with juice, concentrate or sugar, and fruit, hops or spices are added to flavored products. The cider is finally stabilized by filtration, pasteurization or the addition of sorbate and sulfite, and it is packaged still, carbonated under pressure, or conditioned in the bottle by a second fermentation.[24,51,64] Each of these operations leaves a record in the composition. The fruit determines the acids, sorbitol and polyphenols; the fermentation leaves ethanol, glycerol, succinic acid and the higher alcohols; the bacteria exchange malic for lactic acid and, in some cases, glycerol for 1,3-propanediol; sweetening is recognized by the amount and the kind of sugar; and the preservative is itself observed. Sections 3.4 and 3.5 give the chemistry of these changes, and Section 8 reads the record in the ciders of this study.



Figure 1. The cider-making process. Center: the sequence of operations. Left: the choices at each step that distinguish the principal styles. Right: the constituents by which the step is recorded in the ¹H NMR spectrum of the juice or of the finished cider.

3.4 The chemistry of cider making

The alcoholic fermentation converts the fermentable sugars of the juice to ethanol and carbon dioxide. A juice containing 110 g/L of sugars yields a cider of about 6.5 % alcohol by volume when fermented to dryness, and several by-products of yeast metabolism accumulate alongside the ethanol. Glycerol is formed in a quantity equal to several percent of the ethanol by mass,[80] succinic acid at some tenths of a gram per liter,[81] and smaller amounts of acetic acid, 2,3-butanediol and the higher (fusel) alcohols derived from amino acid metabolism.[82,83] Methanol does not originate in fermentation; it is released from the methyl-esterified pectin of the fruit by pectin methylesterase, and its concentration depends on the fruit and on its processing.[84] Sorbitol passes through the fermentation unchanged and contributes body and a residual sweetness even to a cider that contains no sugar.[24]

A second, bacterial transformation is at least as important for the character of the product. In the malolactic fermentation, lactic acid bacteria decarboxylate L-malic acid to L-lactic acid, which lowers the titratable acidity and replaces a sharp acid with a softer one.[79,81,85] The conversion proceeds spontaneously in traditional French and Spanish ciders,[60,61] whereas producers who value the fresh acidity of malic acid prevent it by sulfiting, sterile filtration or pasteurization. The same group of organisms is responsible for further changes. Heterofermentative strains raise the acetic acid content, and certain species convert the glycerol formed by the yeast to 1,3-propanediol, a transformation that is characteristic of natural cider and is described in Section 3.5.

Residual sweetness is obtained in several ways, and each leaves a recognizable composition. Fermentation may be arrested, by the keeving and repeated racking of the French tradition or by chilling and filtration, which leaves the natural sugars of the juice, enriched in fructose because yeast consumes glucose preferentially. A dry cider may be sweetened afterward with fresh juice or concentrate, which restores all three sugars together with malic acid and sorbitol, or with sucrose or another refined sweetener, which adds a single compound.[24,51,86,87] Sucrose is hydrolyzed slowly at the pH of cider,[88] so its presence indicates a comparatively recent addition, while its absence proves nothing. A sweetened cider that has not been pasteurized or sterile-filtered must be protected against renewed fermentation, and potassium sorbate is the usual means; sorbic acid is readily determined by ¹H NMR.[89] Polyphenols and organic acids change comparatively little during the alcoholic fermentation,[90–92] and the final composition therefore records, in addition to the process, something of the fruit.

3.5 Natural cider of the Basque Country and Asturias and the bacterial metabolism of glycerol

The Spanish quality standard distinguishes natural cider (*sidra natural*) from cider (*sidra*). Natural cider is made from fresh apples or their must alone, contains only sugar and carbon dioxide of endogenous origin, may not be sweetened with sucrose and has at least 5 % of alcohol by volume; cider, with at least 4 %, may be sweetened and carbonated. For both, the volatile acidity is limited to 2.2 g/L, expressed as acetic acid.[93] The protected designation of origin Euskal Sagardoa, which covers natural cider made in the Basque Country from apples grown there, requires in addition an alcohol content above 5 %, a dry extract above 14 g/L, less than 100 mg/L of total sulfur dioxide and, in the specification published in 2023, no more than 0.2 g/L of glucose and fructose together; the cider is neither filtered nor carbonated.[94]

The process is the simplest of all cider-making methods and the one that depends most completely on its microorganisms. The apples, mostly acidic and bitter-acidic local varieties, are milled and pressed between late September and November, and the must ferments without added yeast in large vats of chestnut, oak or stainless steel in an unheated cellar. The alcoholic fermentation, carried out by a succession of indigenous yeasts, and the malolactic fermentation, carried out by indigenous lactic acid bacteria, proceed together or in close succession, so that the malic acid of the must has disappeared by the time the sugars are exhausted.[61,76,95] The cider remains on its lees until it is tasted directly from the vat in the cider houses

(*sagardotegiak*) during the *txotx* season, from January to April or May, and is then bottled without filtration. The product is still, cloudy, completely dry and sharply acidic, with lactic acid in place of malic acid and with a volatile acidity that is high by the standard of other ciders and belongs to the style.[60,61,63,96]

Because such a cider is neither heavily sulfited nor filtered, lactic acid bacteria remain active after the malolactic fermentation, and some of them attack glycerol, the principal by-product of the yeast. *Lactobacillus collinoides*, now classified as *Secundilactobacillus collinoides*,[97] and related heterofermentative species possess a glycerol dehydratase that converts glycerol to 3-hydroxypropionaldehyde.[98–100] The aldehyde is an intermediate (Figure 2). The cell reduces most of it to 1,3-propanediol, a reaction that regenerates the oxidized coenzyme required by the rest of its metabolism; a smaller part may be oxidized to 3-hydroxypropionic acid; and what escapes both reactions exists in solution in equilibrium with its hydrate and its dimer and can lose water to give acrolein.[100,101] The strain studied by Claisse and Lonvaud-Funel could not grow on glycerol alone. It degraded glycerol only while it was fermenting glucose or fructose, fructose being preferred; 1,3-propanediol was the main product; and the presence of glycerol shifted the products of the bacterial metabolism from lactic acid toward acetic acid, because the reduction of the aldehyde takes over the disposal of the reducing equivalents.[99]

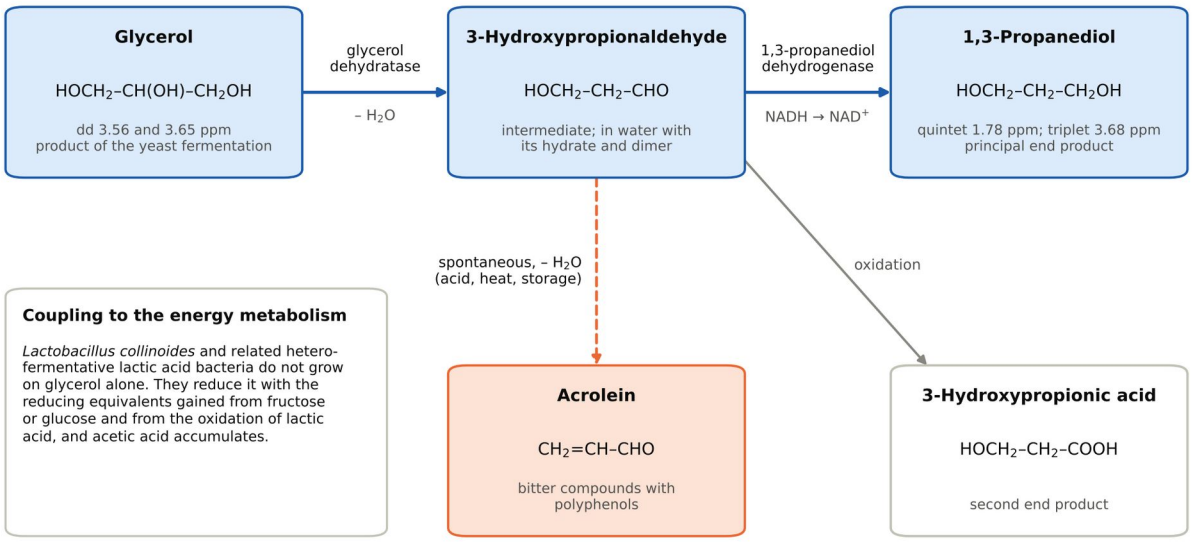

Figure 2. Bacterial metabolism of glycerol in cider. Glycerol, a product of the yeast fermentation, is dehydrated by lactic acid bacteria to 3-hydroxypropionaldehyde, which is reduced to 1,3-propanediol, the principal end product, or oxidized to 3-hydroxypropionic acid; acrolein arises from the aldehyde by spontaneous loss of water. The chemical shifts are those of the signals used in this study.[98–101]

The consequences for the cider are of three kinds. Glycerol, which contributes to the body of a dry cider, is lost, in the extreme case completely. The acetic acid content rises. And acrolein, which is pungent, reacts with the polyphenols of the cider to give bitter products, the fault known in France as *piqûre acroléique* and the cause of the bitterness that is the most feared defect of Basque natural cider.[100,102,103] 1,3-Propanediol itself is without taste at the concentrations concerned, stable and not volatile, so that it remains in the cider as a permanent record of the conversion. In the ¹H NMR spectrum it gives a quintet at 1.78 ppm, in a region where no other constituent of cider interferes, and the spectrum therefore shows at a glance whether, and to what extent, the glycerol of a cider has been metabolized by bacteria. Section 8.5 applies this marker to the ciders of the present study, among them a natural cider from the Basque Country and eight American ciders in which the same conversion had taken place.

3.6 Dryness and its communication

No term on a cider label is used less consistently than dry. To a producer of traditional cider it means the absence of fermentable sugar; on the label of a mass-market product it has been applied to ciders containing 20 g/L of sugar or more. Three North American scales are used in this study. The American Cider Association endorsed, for use in its competition and in communication with consumers, four classes defined by residual sugar alone: dry below 9 g/L, semi-dry from 9 to 18 g/L, semi-sweet from 18 to 45 g/L and sweet above 45 g/L.[104] The cider guidelines of the Beer Judge Certification Program use five classes with boundaries at 4, 9, 20 and 40 g/L.[105] Both scales ignore what every taster knows, namely that acidity and astringency suppress the perception of sweetness.[106–108] The orchard-based dryness scale introduced by the New York Cider Association, commonly called the Merlyn scale, addresses this by classifying on the ratio of residual sugar to titratable acidity, both in grams per liter, with boundaries at ratios of 1, 2 and 3, and by shifting the result toward the dry end for ciders of high polyphenol content.[109,110] Sensory benchmarking has shown that the ratio predicts perceived dryness better than residual sugar alone, although a considerable part of the variance remains unexplained.[111,112]

European regulations also define their terms by residual sugar, but with boundaries that differ from the American ones and from one another. The Spanish standard designates cider as *extra-seca* up to 20 g/L, *seca* from 20 to 30 g/L, *semi-seca* from 30 to 50 g/L and *dulce* above 50 g/L.[93] The French decree of 1953, which was in force when the French ciders examined here were made, defined *brut* as no more than 28 g/L, *demi-sec* as 28 to 42 g/L and *doux* as a higher sugar content combined with less than 3 % of alcohol; it has been replaced, with effect from July 2025, by a decree that defines the same terms by density and alcoholic strength.[113,114] The terms of the European wine regulations, which are sometimes applied to cider, are stricter: dry means no more than 4 g/L, or 9 g/L where the acidity is sufficiently high.[115] A cider with 25 g/L of residual sugar is thus semi-sweet in the United States, *seca* in Spain and *brut* in France, a difference of vocabulary that only a statement of the measured sugar content resolves. The definitions are collected in Table 2; their application to the ciders of this study, together with further scales that can be derived from the same spectrum, is the subject of Section 9.

Table 2. Definitions of the dryness scales and designations applied to the ciders.

Scale	Basis	Classes and boundaries
American Cider Association[104]	residual sugar (g/L)	Dry, below 9; Semi-Dry, 9 to 18; Semi-Sweet, 18 to 45; Sweet, above 45
BJCP cider guidelines[105]	residual sugar (g/L)	Dry, below 4; Semi-dry, 4 to 9; Medium, 9 to 20; Semi-sweet, 20 to 40; Sweet, above 40
Orchard-based (Merlyn) scale of the New York Cider Association[109,110]	residual sugar divided by titratable acidity (both in g/L), adjusted for polyphenols	Dry, ratio below 1; Semi-Dry, 1 to 2; Semi-Sweet, 2 to 3; Sweet, above 3; the ratio is lowered by 0.25, 0.5 or 0.75 for total polyphenols of 500 to 750, 750 to 1000 and above 1000 mg/L
Spain, Real Decreto 72/2017 (*sidra*)[93]	residual sugar (g/L)	Extra-seca, up to 20; Seca, 20 to 30; Semi-seca, 30 to 50; Dulce, above 50
France, decree of 1953[113,114]	residual sugar (g/L)	Brut, up to 28; Demi-sec, 28 to 42; Doux, more than 35 (more than 42 for *cidre bouché*) with less than 3 % of alcohol by volume
European Union, terms for still wine[115]	residual sugar (g/L), with allowance for acidity	Dry, up to 4 (up to 9 if the total acidity, as tartaric acid, is not more than 2 g/L below the sugar content); Medium dry, up to 12 (up to 18 if the acidity is not more than 10 g/L below); Medium, up to 45; Sweet, above 45

3.7 Nutrition labeling of cider in the United States

The labeling of cider in the United States is divided between two agencies according to the alcohol content. A cider of 7 % alcohol by volume or more is a wine within the meaning of the Federal Alcohol Administration Act and is labeled under the rules of the Alcohol and Tobacco Tax and Trade Bureau (TTB). A cider below 7 % is outside that definition, and its label falls under the food labeling regulations of the Food and Drug Administration, which require an ingredient list and, subject to the exemptions available to small businesses, a Nutrition Facts panel.[67,116] The majority of the ciders on the market, and of those examined here, belong to the second group. A separate definition applies for the purposes of taxation, where the hard cider rate covers apple and pear wines of at least 0.5 % and less than 8.5 % alcohol by volume with limited carbonation.[117] For products under its jurisdiction TTB permits voluntary statements of calories, carbohydrate, protein and fat and has published the analytical methods and tolerances that it applies,[118,119] and in January 2025 it proposed a mandatory Alcohol Facts statement for wines, distilled spirits and malt beverages, giving the serving size, the alcohol content, the fluid ounces of pure alcohol, the calories and the carbohydrate, fat and protein of a serving, with a compliance date five years after publication of a final rule.[120] The comment period closed in August 2025, and the rule had not been made final when this document was prepared. The proposal would reach cider of 7 % alcohol or more; cider below that strength would remain under the food labeling regulations.

Energy is calculated in both systems from the composition. The general factors of 4 kcal/g for carbohydrate and 7 kcal/g for alcohol go back to Atwater; the more exact specific factor for ethanol, 7.07 kcal/g, and the derivation of all such factors are given in USDA Agriculture Handbook No. 74.[121] The food labeling regulations assign reduced factors to the sugar alcohols, 2.6 kcal/g in the case of sorbitol,[116,122] and glycerol is counted at 4.32 kcal/g.[123] Organic acids, which are not negligible in cider, are assigned 3 kcal/g in the conventions of the Food and Agriculture Organization and of the European Union.[124,125] The TTB laboratory method for flavored malt beverages calculates the calories of a 12 fl oz serving as $3.55 \times [6.9\,A + 4\,S + 2.4\,(TC - S) + 4\,P + 9\,F]$, in which A, S, TC, P and F are the contents of alcohol, sugars, total carbohydrate, protein and fat in grams per 100 mL; a single factor of 2.4 kcal/g is thus applied to all carbohydrate other than sugars.[119] ¹H qNMR delivers the terms of both calculations from one measurement, which is the reason for the attention given to nutrition labeling in Section 10.

3.8 ¹H NMR of apple juice, cider and related beverages

The assignment of the ¹H NMR spectrum of apple juice at high field was published by Belton and co-workers, who then used principal component and discriminant analysis of the spectra to distinguish juices of different varieties.[47,48] Fruit juice authentication by NMR subsequently developed into a routine commercial service, in which several dozen compounds are quantified and the spectrum as a whole is compared statistically with a reference database of authentic juices,[126–128] and it has been reviewed repeatedly in the wider context of food analysis.[129–132] Related work addressed the detection of adulteration and of juice-to-juice addition,[133–135] untargeted profiling of fruit metabolites,[136,137] and the variety, growing system and storage of apples, studied in juices and, by magic angle spinning, in the intact tissue.[138–143]

Cider itself was investigated chiefly by the group of Berregi and del Campo in the Basque Country, who established ¹H NMR methods for chlorogenic acid,[144] formic acid,[145] malic and citric acids[146] in apple juice, followed the ripening of cider varieties and classified their juices by the signals of polyphenols and acids,[147,148] and later determined ethanol,[149] lactic and acetic acids[150] and acrolein[103] in the finished cider. The fermentation of apple juice by different yeasts has been followed by NMR,[151,152] the degree of polymerization of cider apple procyanidins has been measured,[153] and the composition of ciders has been combined with stable isotope ratios as a guide to authenticity.[154] The closest methodological relatives of the present work are the quantitative NMR procedures for wine[155–159] and beer,[160–163] the automated suppression of the water and ethanol signals that such procedures employ at high field,[164] and recent demonstrations that

the principal constituents of fermented beverages can be quantified at low field.[165,166] The use of chemometrics with NMR data of foods, and the validation that such models require before they are relied upon, are discussed in several reviews.[167–170]

4. The Juices of 56 Apple Varieties: Composition and Chemometrics

4.1 The collection and the origins of the varieties

The fruit for the survey was obtained on a single occasion. In November 2016 Second Chance Farm, a grower of old varieties at East Syracuse, New York, offered its apples at the 22nd Franklin County CiderDays in Shelburne Falls, Massachusetts,[171,172] and 53 lots, each sold under a variety name, were purchased, pressed in the laboratory and analyzed. Fifty of the names belong to described cultivars; one lot was labeled Kingston Black with a query, and two names, Pear Apple and Goof, are reproduced as they were recorded with the samples. A supermarket Granny Smith was prepared alongside, and the juices of supermarket Gala and Fuji that had been measured in 2015 with the same preparation were added for comparison, so that 56 juices are available in all. The varieties, their places and dates of origin, parentage where it is known, and the uses for which they were valued are collected in Table 3.

The collection is a cross-section of five centuries of apple growing on both sides of the Atlantic. Twenty of the varieties arose in the northeastern United States and thirteen elsewhere in the country, eleven in the British Isles, three on the European continent and three in Canada; four, Kidd's Orange Red, Granny Smith, Gala and Fuji, come from New Zealand, Australia and Japan. Fifteen were in cultivation before 1800, twenty-six date from the nineteenth century and thirteen from the twentieth. The oldest are European. Decio, an Italian apple recorded in 1529, is by tradition said to have been brought to the Veneto in the fifth century by the Roman general Aetius, from whose Italian name, Ezio, its own is derived; Calville Blanc d'Hiver, the ribbed cooking apple of French cuisine, was described in 1598; Scarlet Crofton has been associated with the Crofton family of County Sligo since the early seventeenth century; and Golden Harvey, a small Herefordshire russet of the same century, was known as the Brandy Apple for the strength of the cider that it gave. Devonshire Quarrenden, Ashmead's Kernel, raised at Gloucester about 1700, and Pitmaston Pineapple complete the group of old English dessert apples.[19,22,173]

The American varieties of the colonial period are, without exception, chance seedlings that were noticed and propagated (Section 2.1). Roxbury Russet, found at Roxbury in Massachusetts about 1640, is regarded as the oldest named variety of North America, and Rhode Island Greening followed at Newport about 1650. The Newtown Pippin, which arose in Queens County, New York, early in the eighteenth century and is represented here by its green and its yellow strain, was the first American apple to win a reputation in Europe and was grown in Virginia as the Albemarle Pippin. Baldwin, from Wilmington in Massachusetts, became the leading commercial apple of New England and kept that position until the freeze of the winter of 1933 to 1934; Esopus Spitzenburg, from the Hudson valley, was planted by Thomas Jefferson at Monticello; and Sutton Beauty and Wagener close the century.[13,14,18] The nineteenth century added the varieties of an expanding country: Northern Spy, Tompkins King, Fall Pippin and Golden Russet in New York, Ben Davis, Yates and Virginia Beauty in the South, King David in Arkansas, Winter Banana in Indiana, Shiawassee Beauty in Michigan and Opalescent in Ohio, together with Golden Noble and Barnack Beauty from England, Marie-Joseph d'Othée from Belgium and the Peach of Montreal from Quebec.[13,18,173,174]

The remaining varieties are products of deliberate breeding or of selection in the twentieth century. Chieftain, a cross of Jonathan and Delicious made at the Iowa Agricultural Experiment Station in 1917, was introduced in 1966;[175] Kidd's Orange Red, bred in New Zealand from Cox's Orange Pippin and Delicious, became a parent of Gala; Bancroft and Quinte were raised at the Central Experimental Farm in Ottawa; Pink Pearl is a pink-fleshed selection of the Californian breeder Albert Etter; Chestnut Crab was released by the University of Minnesota in 1946; GoldRush, a scab-resistant selection of the cooperative program of Purdue, Rutgers and Illinois, was released in 1993;[176] and Razor Russet is a russeted sport of Golden Delicious found in Kentucky. With regard to cider the set is of mixed standing. Golden Russet and Roxbury Russet were the most esteemed American cider apples of their time, Ashmead's Kernel, Golden Harvey, the

Newtown Pippins, Esopus Spitzenburg, Baldwin, Northern Spy, King David and GoldRush are valued by present-day cider makers for their sugar, acid or aroma, and Kingston Black is the classic bittersharp of Somerset; most of the others were grown for the table, the kitchen or the store room.[18,20,25] Six of the varieties are russets, a group that proves to be distinct in composition (Section 4.3).

Table 3. The 56 apple juices: variety, place and date of origin, parentage, principal use and remarks.[13,14,18,19,22,173–176]

No.	Variety	Origin	Date	Parentage	Use	Remarks
A01	Calville Blanc d'Hiver	France (possibly Germany)	recorded 1598	unknown	culinary and dessert	The classic cooking apple of French cuisine; ribbed, pale yellow
A02	Fall Pippin	United States, place uncertain	before 1806	unknown	culinary and dessert	Large yellow autumn apple, long grown in New York and the Hudson valley
A03	Dr. Matthews	Tennessee and Indiana, USA	listed 1895	unknown	dessert	Attributed to Dr. W. F. Matthews; long grown in Indiana as a keeping dessert apple
A04	Chieftain	Ames, Iowa, USA	cross 1917, introduced 1966	Jonathan × Delicious	dessert and culinary	Cross made by S. A. Beach at the Iowa Agricultural Experiment Station
A05	Pink Pearl	Ettersburg, California, USA	introduced 1944	seedling of Surprise	dessert	Pink-fleshed apple bred by Albert Etter
A06	Oaken Pin	Exmoor, Devon and Somerset, England	nineteenth century or earlier	unknown	dessert and culinary	Oval West Country apple named after the wooden peg it resembles
A07	King David	Washington County, Arkansas, USA	found 1893, introduced 1902	reputedly Jonathan × Winesap or Arkansas Black	dessert and cider	Dark red, sprightly; valued by American cider makers
A08	Benoni	Dedham, Massachusetts, USA	before 1832	unknown	dessert	Early-season New England dessert apple
A09	Green Newtown Pippin	Newtown, Queens County, New York, USA	early eighteenth century; recorded 1759	unknown	dessert, culinary and cider	The green strain of the Newtown Pippin, the first American apple to win a reputation in Europe
A10	Ashmead's Kernel	Gloucester, England	about 1700	unknown	dessert and cider	Russet raised by Dr. Ashmead; sharp, aromatic and high in sugar
A11	Golden Harvey	Herefordshire, England	seventeenth century	unknown	dessert and cider	Small russet also called the Brandy Apple for the strength of its cider
A12	Yates	Fayette County, Georgia, USA	about 1844	unknown	dessert, cider and keeping	Small, late, long-keeping apple of the southern states
A13	Esopus Spitzenburg	Esopus, Ulster County, New York, USA	eighteenth century, before 1790	unknown	dessert, culinary and cider	Hudson valley apple of high flavor, grown by Thomas Jefferson at Monticello
A14	Tompkins King	New Jersey, brought to Tompkins County, New York	before 1804	unknown	dessert and culinary	King of Tompkins County; large, red, triploid
A15	Golden Russet	New York, USA	before 1845	unknown, perhaps a seedling of an English russet	cider, dessert and keeping	The most esteemed American russet for cider; very high sugar
A16	Decio	Veneto and Emilia, Italy	ancient; recorded 1529	unknown	dessert	By tradition carried north by the Roman general Ezio (Aetius) in the fifth century
A17	Edward VII	Worcester, England	introduced 1908	reputedly Blenheim Orange × Golden Noble	culinary	Late-keeping English cooking apple
A18	Rhode Island Greening	Green's End, Newport, Rhode Island, USA	about 1650	unknown	culinary	The standard American pie apple of the nineteenth century

No.	Variety	Origin	Date	Parentage	Use	Remarks
A19	Baldwin	Wilmington, Massachusetts, USA	about 1740	unknown	dessert, culinary and cider	The leading commercial apple of New England until the freeze of 1933-34
A20	Starr	Woodbury, New Jersey, USA	about 1865	unknown	culinary	Large, early, green cooking apple
A21	Razor Russet	Wallingford, Kentucky, USA	1970s	russet sport of Golden Delicious	dessert and cider	Limb mutation found by W. D. Armstrong of the University of Kentucky
A22	GoldRush	West Lafayette, Indiana, USA	selected 1980, released 1993	Golden Delicious × Co-op 17	dessert, keeping and cider	Scab-resistant selection of the Purdue-Rutgers-Illinois program; high sugar and high acid
A23	Jewett Red (Nodhead)	Hollis, New Hampshire, USA	before 1842	unknown	dessert	Old northern New England dessert apple
A24	American Beauty (Sterling)	Sterling, Massachusetts, USA	before 1855	unknown	dessert	Also called Sterling Beauty
A25	Kidd's Orange Red	Greytown, New Zealand	1924	Cox's Orange Pippin × Delicious	dessert	Bred by J. H. Kidd; a parent of Gala
A26	Opalescent	Ohio, USA	introduced 1899	unknown	dessert	Large, dark red, glossy apple introduced by the Dayton Star Nurseries
A27	Ben Davis	Upper South, USA	about 1800	unknown	keeping and shipping	The dominant commercial apple of the South and Midwest in the late nineteenth century
A28	Chestnut Crab	Excelsior, Minnesota, USA	introduced 1946	open-pollinated seedling of Malinda	dessert crab	Large russeted crab apple of the University of Minnesota
A29	Scarlet Crofton	County Sligo, Ireland	early seventeenth century	unknown	dessert	Associated with the Crofton family of Longford House; partly russeted
A30	Quinte	Ottawa, Ontario, Canada	introduced 1964	Crimson Beauty × Red Melba	dessert	Early apple of the Central Experimental Farm
A31	Barnack Beauty	Barnack, Northamptonshire, England	about 1840, introduced about 1870	unknown	dessert and culinary	Late-keeping English apple
A32	Sweet Winter Pennock	Delaware County, Pennsylvania, USA	before 1817	unknown	keeping	Taken to be Pennock (Pennock's Red Winter), first grown by Joseph Pennock
A33	Bancroft	Ottawa, Ontario, Canada	selected 1930, introduced about 1935	Forest × McIntosh	dessert and keeping	Hardy winter apple of the Central Experimental Farm
A34	Titus Pippin	Oyster Bay, Long Island, New York, USA	1839	unknown	dessert and culinary	Also known as Long Island Bellflower
A35	Northern Spy	East Bloomfield, New York, USA	about 1800	unknown	dessert, culinary and cider	Late, long-keeping apple of high acidity
A36	Pear Apple	not established	not established	unknown	not established	Name as written on the sample; not matched to a described cultivar
A37	Golden Noble	Norfolk, England	1820	unknown	culinary	Sharp English cooking apple
A38	Devonshire Quarrenden	Devon, England	recorded in the 1670s	unknown	dessert	Early, crimson dessert apple, possibly of French origin
A39	Virginia Beauty	Piper's Gap, Carroll County, Virginia, USA	seedling of about 1810, named about 1850	unknown	dessert	Raised by Zach Safewright; first distributed as Zach Red

No.	Variety	Origin	Date	Parentage	Use	Remarks
A40	Winter Banana	Cass County, Indiana, USA	1876	unknown	dessert	Pale yellow, waxy apple with a pink blush
A41	Marie-Joseph d'Othée	Othée, Liège, Belgium	about 1870	unknown	dessert, culinary and keeping	Recorded on the sample as Marie Joseph; very late keeping
A42	Mother	Bolton, Massachusetts, USA	recorded 1844	unknown	dessert	American Mother; aromatic dessert apple
A43	Pitmaston Pineapple	Worcester, England	about 1785	reputedly a seedling of Golden Pippin	dessert	Small golden russet with a honeyed flavor
A44	Roxbury Russet	Roxbury, Massachusetts, USA	about 1640	unknown	cider, keeping and dessert	Regarded as the oldest named American variety
A45	Shiawassee Beauty	Shiawassee County, Michigan, USA	about 1850	seedling of Fameuse	dessert	Snow-white flesh inherited from Fameuse
A46	Wagener	Penn Yan, New York, USA	seed planted 1791	unknown	dessert and culinary	A parent of Idared
A47	Sutton Beauty	Sutton, Massachusetts, USA	about 1757	unknown	dessert	Red winter apple of Worcester County
A48	Yellow Newtown Pippin	Newtown, Queens County, New York, USA	early eighteenth century; recorded 1759	unknown	dessert, culinary and cider	The yellow strain, grown in Virginia as the Albemarle Pippin
A49	Collamer	Hilton, Monroe County, New York, USA	about 1900	strain of Twenty Ounce	culinary	Recorded on the sample as Colamar; Collamer Twenty Ounce
A51	Montreal Peach	Montreal, Quebec, Canada	before 1876	unknown	dessert	Peach of Montreal or Pomme Pêche; early, tender apple
A52	Goof	not established	not established	unknown	not established	Name as written on the sample; not matched to a described cultivar
A53	Ellison's Orange	Bracebridge, Lincolnshire, England	recorded 1904, introduced 1911	Cox's Orange Pippin seedling	dessert	Aromatic early dessert apple with an aniseed note
A54	Granny Smith	Eastwood, New South Wales, Australia	1868	chance seedling	dessert and culinary	Supermarket fruit, analyzed for comparison
A55	Kingston Black	Kingston St Mary, Somerset, England	nineteenth century	unknown	cider (bittersharp)	The classic single-variety English cider apple; lot labeled with a query
G15	Gala (2015)	Greytown, New Zealand	cross 1934, named 1965	Kidd's Orange Red × Golden Delicious	dessert	Supermarket fruit of 2015, analyzed for comparison
F15	Fuji (2015)	Fujisaki, Aomori, Japan	cross 1939, named 1962	Ralls Janet × Delicious	dessert	Supermarket fruit of 2015, analyzed for comparison

A01 to A53 and A55, fruit of Second Chance Farm (2016); A54, G15 and F15, supermarket fruit. Dates are those of the first record, of the cross or of the introduction, as stated.

4.2 Samples, spectra and measurement

The apples of each lot were pressed, and 1000 µL of the juice was freeze-dried and redissolved in D_2O together with 100 µL of the maleic acid standard solution, the preparation that is used for the non-volatile constituents of cider (Section 5.2). Two juices, Golden Noble and Marie-Joseph d'Othée, were prepared from 500 µL, which was taken into account in the calculation. The ¹H spectra were acquired at 27 °C on two 300 MHz spectrometers, 45 on the Varian Mercury MVX-300 and 11 on a Varian Mercury Plus 300, with a

one-pulse sequence, an acquisition time of 8.9 s, a relaxation delay of 7 s and, as a rule, 64 transients. All 56 spectra are reproduced in Figure 3, drawn on one scale after standardization of each juice to the same sugar content, and expansions of the regions that carry the information are given for every juice in Figure 4 and Figure 5. To the eye the spectra differ in three respects: the height of the sucrose doublet at 5.40 ppm relative to the remainder of the carbohydrate envelope, the intensity of the malic acid multiplet near 2.9 ppm, and the pattern of small signals between 1.2 and 2.2 ppm and in the aromatic region.

The assignment is illustrated with the juice of Northern Spy in Figure 6. The sugars are measured as in the ciders: sucrose and glucose from their anomeric doublets, fructose and sorbitol by decomposition of the region between 4.30 and 3.18 ppm with the spectra of the pure compounds (Section 5.6). Malic acid required a different treatment from that used for cider. In a juice its methylene multiplet, the AB part of an ABX system, is accompanied by the methylene signals of asparagine near 3.00 ppm and of aspartic acid near 3.09 ppm, whose positions move with the acidity of the sample, and a fixed integration window would collect a varying part of them. The multiplet was therefore fitted as a complete ABX system[177] with the coupling constants determined on the set as a whole (16.6, 4.35 and 7.04 Hz), together with doublets for the two amino acids. Three further assignments are specific to the juices. A singlet at 1.48 ppm, present in 41 of the 53 juices, is accompanied by an AB quartet at 3.08 and 2.75 ppm with a coupling of 16.2 Hz whose intensity follows that of the singlet with the ratio expected for a methylene and a methyl group (correlation coefficient 0.99). The two signals are assigned to citramalic (2-methylmalic) acid, a constituent of apples since its isolation from the peel[178] and now recognized as an intermediate of ester biosynthesis in the ripening fruit.[179,180] A doublet at 1.26 ppm with a coupling of 6.35 Hz is the methyl group of L-rhamnitol, a polyol that NMR profiling identified in apple juice,[140,181] and the doublet of α-xylose is found at 5.18 ppm. Quinic acid, the hydroxycinnamoyl esters, expressed as chlorogenic acid, and the aromatic integral were measured as in the ciders.

Concentrations were calculated from the molar ratio of each signal to the standard as for the freeze-dried tubes of the ciders (Section 5.7) and are reported on two bases. The first is the concentration per liter of the juice as it was measured. The second is the composition relative to the sugars: the shares of fructose, glucose and sucrose in their sum, the other constituents per 100 g of sugars, and, as a form that is easier to read, the concentrations in a juice standardized to 110 g/L of sugars, which corresponds to about 12 °Brix and to a potential alcohol content of about 6.6 % v/v. The relative composition depends only on ratios of signals within one spectrum. It is unaffected by the dry matter of the individual juice, which varied widely among the lots, and it is the basis on which the varieties are compared in the sections that follow. Reporting limits range from 5 mg/L for alanine and 10 mg/L for citramalic acid and L-rhamnitol to 60 mg/L for the amino acids of the malic acid region and 300 mg/L for sorbitol. Four juices, those of Starr, Scarlet Crofton, Benoni and the lot labeled Kingston Black, contained a little ethanol and, in part, 2,3-butanediol, the mark of a fermentation that had begun in the fruit or the juice; they are identified in the figures, and their sugar contents are lower limits.

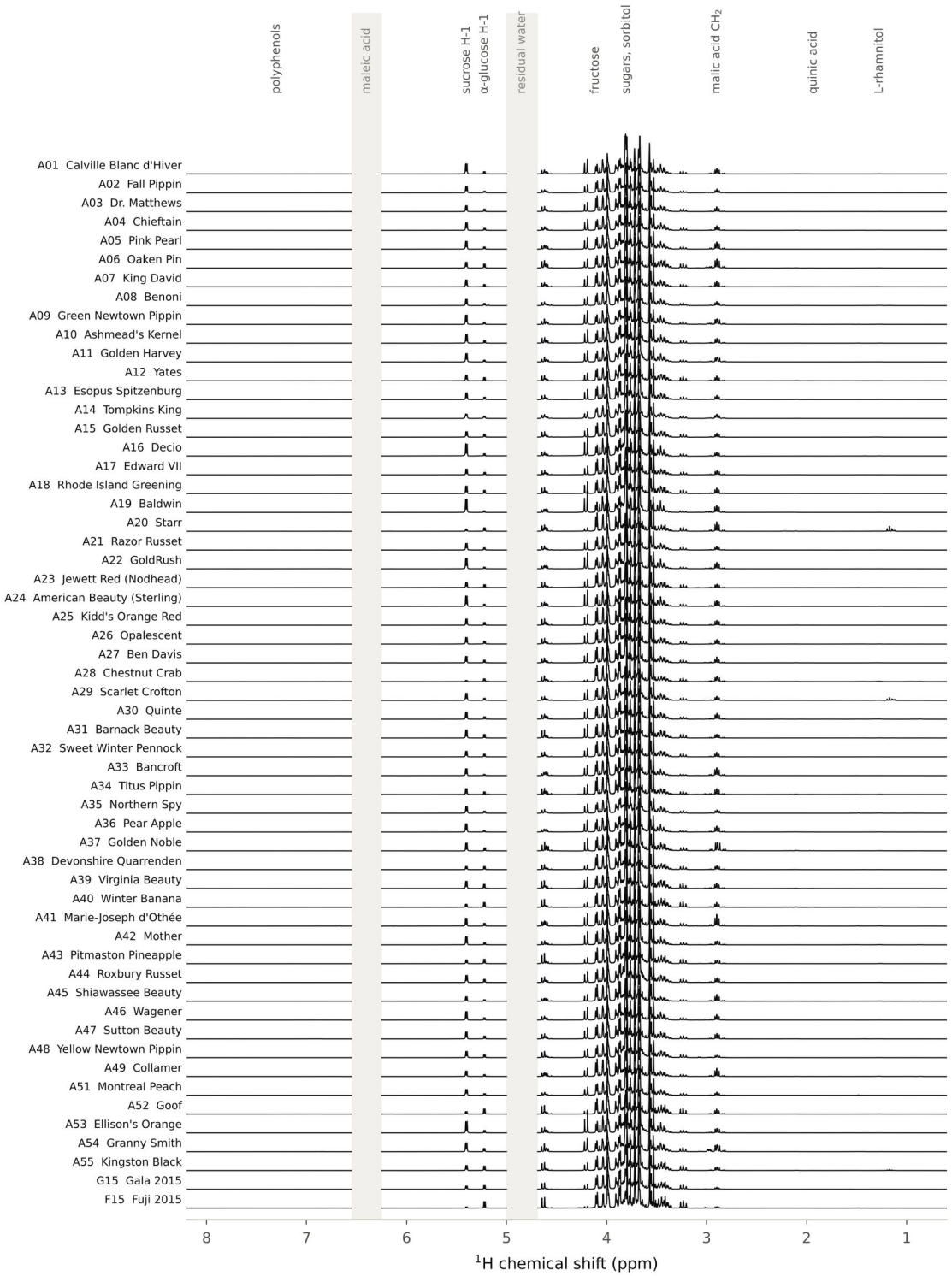

Figure 3. ¹H NMR spectra (300 MHz, D_2O) of the juices of 56 apple varieties, freeze-dried and redissolved with maleic acid as internal standard; 8.2 to 0.6 ppm. The traces are drawn on one vertical scale after standardization of each juice to 110 g/L of sugars. The regions of the standard and of residual water are shaded.

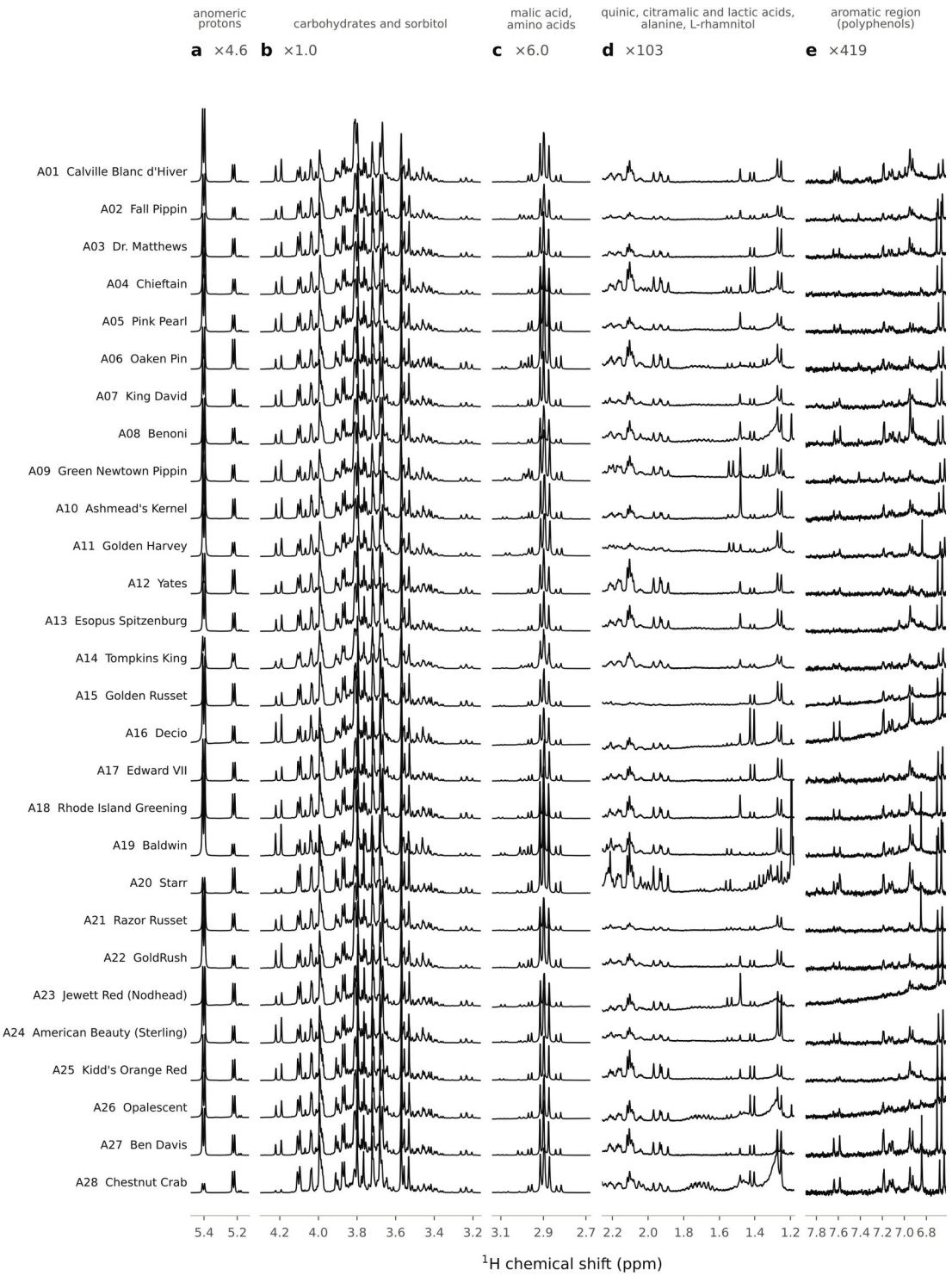

Figure 4. Expansions of the juice spectra, varieties A01 to A28. (a) Anomeric protons of sucrose and α-glucose. (b) Carbohydrates and sorbitol. (c) Malic acid CH_2 with the methylene signals of asparagine and aspartic acid. (d) Quinic, citramalic and lactic acids, alanine and L-rhamnitol. (e) Aromatic region. All traces of one panel share one vertical scale; the factor in the panel heading gives its magnification relative to panel b. Juices standardized to 110 g/L of sugars.

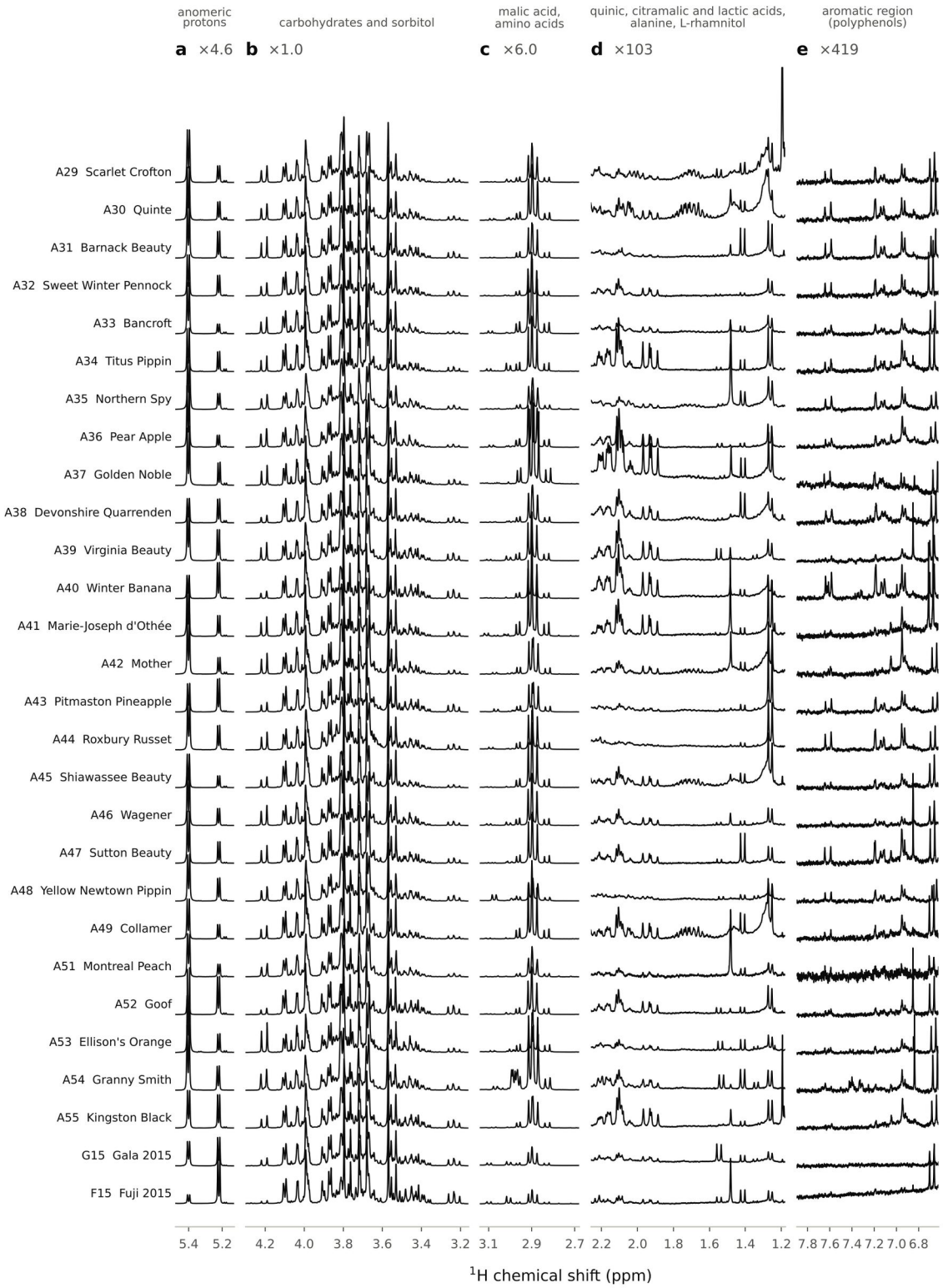

Figure 5. Expansions of the juice spectra, varieties A29 to A55 and the supermarket Gala and Fuji of 2015. Panels and scales as in the preceding figure.

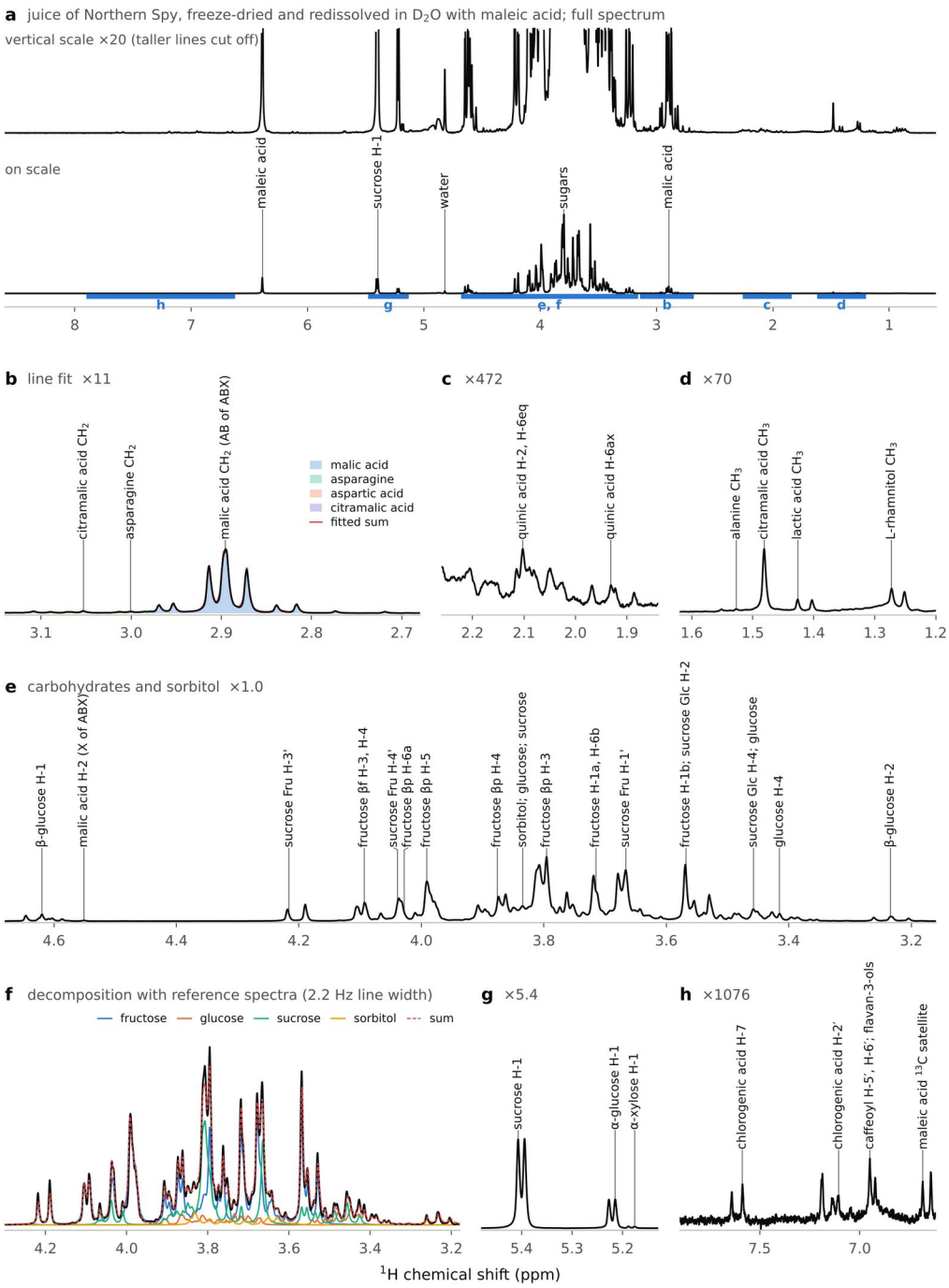

Figure 6. Assigned ¹H NMR spectrum of an apple juice (Northern Spy; freeze-dried and redissolved in D_2O with maleic acid, 300 MHz). (a) Full spectrum on scale and, above it, expanded twentyfold; the bars mark the regions shown below. (b) Malic acid region with the line fit: the ABX multiplet of malic acid and the contributions of asparagine, aspartic acid and citramalic acid. (c) Quinic acid. (d) Methyl signals of alanine, citramalic acid, lactic acid and L-rhamnitol. (e) Carbohydrates and sorbitol. (f) Decomposition of the carbohydrate region into the spectra of fructose, glucose, sucrose and sorbitol. (g) Anomeric protons. (h) Aromatic region. The factor in each heading is the vertical magnification relative to panel a. βp and βf, β-pyranose and β-furanose forms of fructose.

4.3 Sugars, sorbitol and malic acid

The principal results are shown in Figure 7 and summarized in Table 4; the values for every juice are tabulated in Appendix D. Fructose is the most abundant sugar in all but two juices, Baldwin and Decio, and accounts on average for 59.4 % of the three sugars, which agrees with the mean of 57.7 % reported for 175 varieties.[38] The other two sugars depart from that reference in opposite directions. Sucrose averages 27.8 % against 21.9 %, glucose 12.9 % against 20.4 %, and the ratio of fructose to glucose, 5.0 on average, is correspondingly higher than the 2.8 of the literature means. The sucrose share is the most variable property of the set. It extends from 5 % in Chestnut Crab and 8 % in Starr to 40 % or more in Baldwin (52 %), Decio, GoldRush, Ellison's Orange, Calville Blanc d'Hiver, American Beauty and Green Newtown Pippin. Because sucrose is hydrolyzed in stored fruit and, slowly, in the acidified sample, the shares are lower limits; the ranking of the varieties is nevertheless a property of the fruit, since all lots were handled in the same way. The supermarket fruit lies at the edge of the heirloom range. Fuji and Gala contain 5 and 14 % of sucrose and the largest glucose shares of the whole set, 30 and 19 %, the composition of dessert fruit from long storage.

Malic acid ranges from 3.4 to 10.4 g per 100 g of sugars among the lots of Second Chance Farm, with a median of 5.3 and a mean of 5.8, somewhat below the literature mean of 8.6 g per 100 g.[38] Golden Noble, Oaken Pin, Marie-Joseph d'Othée, Bancroft and Starr are the most acidic relative to their sugar, Montreal Peach, Yellow Newtown Pippin and Jewett Red the least. The three supermarket varieties mark the two ends of the scale: Granny Smith, at 10.8 g per 100 g, is more acidic than any of the old varieties, while Gala (2.0) and Fuji (1.4) are less acidic than all of them (Figure 8a). The modern dessert apple has, in other words, been selected to a ratio of sugar to acid that lies outside the range within which the old varieties were grown, which is consistent with the selection for low acidity that the comparison of cultivated and wild apples has documented.[43]

Sorbitol amounts to 2.1 g per 100 g of sugars in the median juice and varies more than tenfold, from 0.6 g in GoldRush to 10.6 g in Roxbury Russet. The russets stand out as a group. Their median, 5.4 g per 100 g, is more than twice that of the other varieties (2.0 g; Mann-Whitney test, $p = 0.010$), and Roxbury Russet, Yellow Newtown Pippin, Golden Harvey, Pitmaston Pineapple and Golden Russet head the list (Figure 8b). The observation is of practical interest for two reasons. Sorbitol is not fermented by yeast, so that it passes into the cider and adds body and a residual sweetness to a product that contains no sugar (Section 3.4); and among the constituents of the apple it is the one that correlates most closely with the sweetness perceived by a trained panel.[182] The reputation of the russets for rich, full ciders has, in part, this chemical basis.

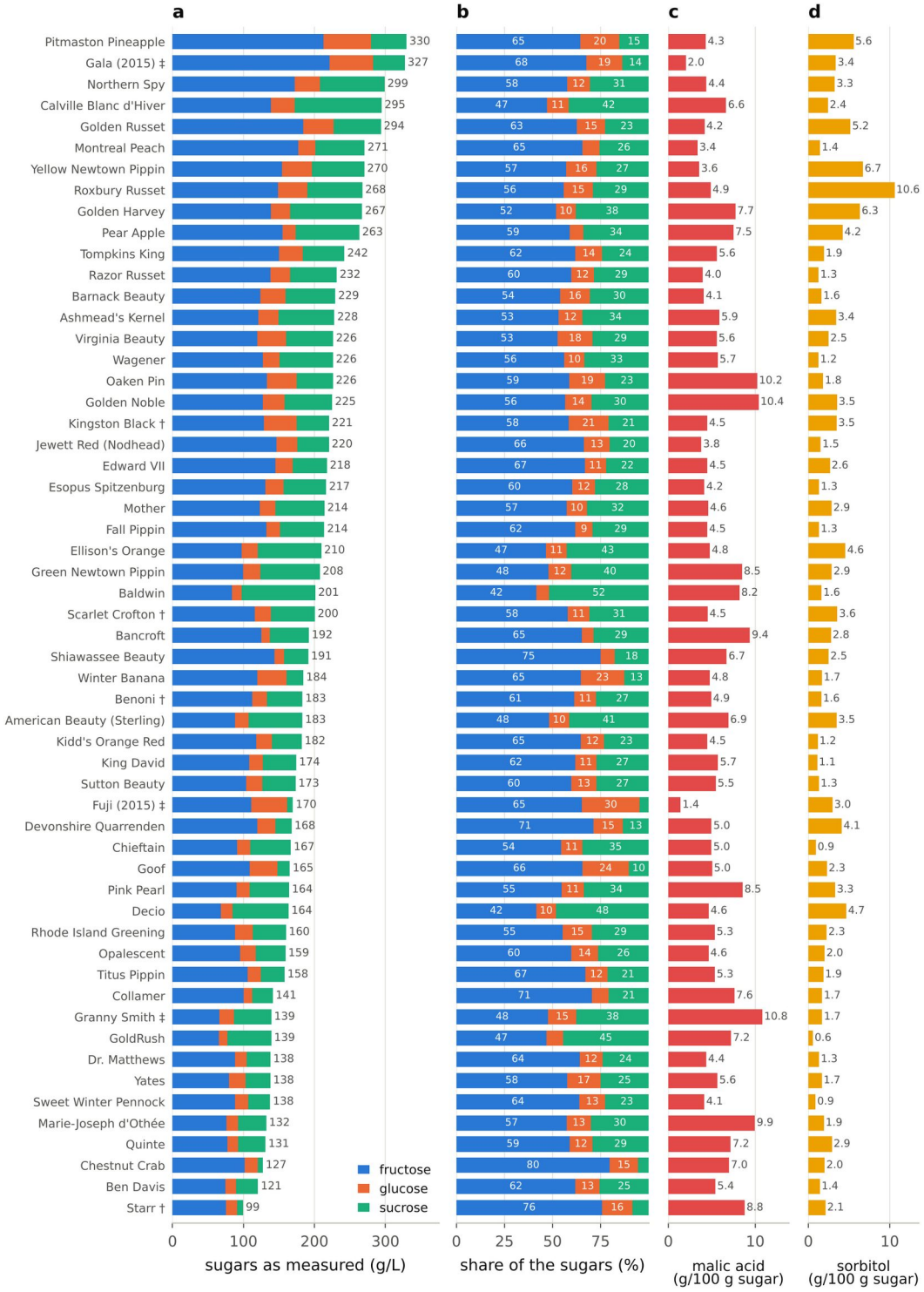

Figure 7. Sugars, malic acid and sorbitol in the 56 apple juices, ordered by the sum of the sugars. (a) Fructose, glucose and sucrose per liter of the juice as measured; the number is their sum. (b) Shares of the three sugars (% by mass). (c) Malic acid and (d) sorbitol per 100 g of sugars. † juices in which fermentation had begun; ‡ supermarket fruit.

Table 4. Composition of the apple juices: the 53 lots of Second Chance Farm (minimum, median, mean and maximum) and the three supermarket juices, compared with the mean of 175 varieties.[38]

Quantity	Minimum	Median	Mean	Maximum	Granny Smith	Gala	Fuji	Mean of 175 varieties
Sucrose (% of the three sugars)	5.5	27.8	27.8	51.8	37.5	13.7	4.7	21.9
Glucose (% of the three sugars)	6.0	11.8	12.9	24.0	14.9	18.6	30.0	20.4
Fructose (% of the three sugars)	41.5	58.9	59.4	79.7	47.6	67.7	65.3	57.7
Fructose / glucose (g/g)	2.7	4.8	5.0	10.8	3.2	3.6	2.2	2.8
Malic acid (g per 100 g of sugars)	3.4	5.3	5.8	10.4	10.8	2.0	1.4	8.6
Sugars / malic acid (g/g)	9.6	18.7	18.6	29.8	9.2	49.9	71.1	11.6
Sorbitol (g per 100 g of sugars)	0.6	2.1	2.7	10.6	1.7	3.4	3.0	–
Malic acid in the standardized juice (g/L)	3.7	5.9	6.4	11.4	11.9	2.2	1.5	–
Quinic acid, standardized juice (mg/L)	37	240	293	1051	306	159	110	–
Citramalic acid, standardized juice (mg/L)	0	18	27	178	8.5	nd	77.3	–
L-Rhamnitol, standardized juice (mg/L)	0	58	68	209	73.6	22.3	37.9	–
Asparagine, standardized juice (mg/L)	0	149	180	1156	3401	283	545	–
Hydroxycinnamates as chlorogenic acid, standardized juice (mg/L)	0	113	135	376	98.6	61.0	46.2	–
Aromatic protons, standardized juice (mmol/L)	0.0	2.5	2.6	5.6	4.2	1.3	0.1	–
Sugars in the juice as measured (g/L)	99	200	200	330	139	327	170	–

Standardized juice: concentrations scaled to 110 g/L of fructose, glucose and sucrose. Sucrose shares are lower limits because of hydrolysis. A minimum of 0 denotes a constituent below the reporting limit in at least one juice. The literature column was calculated from the published mean concentrations.

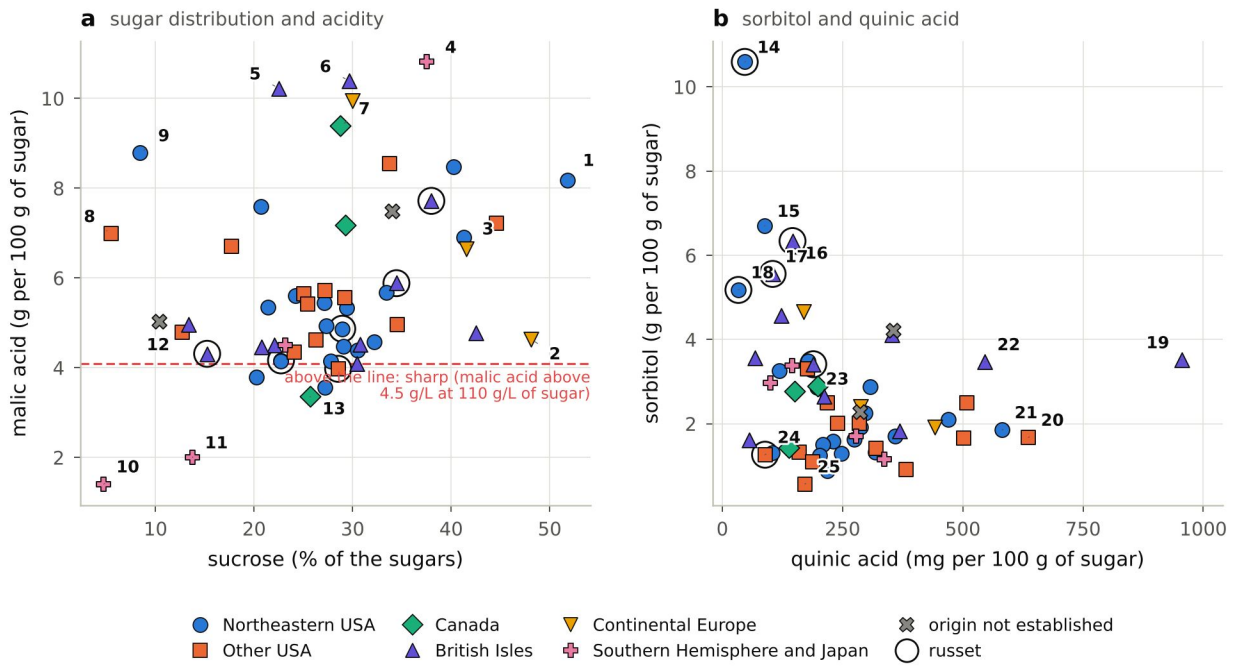

Figure 8. Relations among the principal constituents of the apple juices, by region of origin of the variety. (a) Malic acid against the sucrose share of the sugars; juices above the broken line would be classed as sharp (more than 4.5 g/L of malic acid at 110 g/L of sugars). (b) Sorbitol against quinic acid, both relative to the sugars. Numbered points: 1, Baldwin; 2, Decio; 3, Calville Blanc d'Hiver; 4, Granny Smith; 5, Oaken Pin; 6, Golden Noble; 7, Bancroft; 8, Chestnut Crab; 9, Starr; 10, Fuji (2015); 11, Gala (2015); 12, Winter Banana; 13, Montreal Peach; 14, Roxbury Russet; 15, Yellow Newtown Pippin; 16, Golden Harvey; 17, Pitmaston Pineapple; 18, Golden Russet; 19, Golden Noble; 20, Winter Banana; 21, Titus Pippin; 22, Kingston Black; 23, Ashmead's Kernel; 24, Razor Russet; 25, GoldRush.

4.4 Minor constituents

The minor constituents are displayed in Figure 9 for the standardized juices. Quinic acid, at a median of 240 mg/L, spans a range of almost thirtyfold, from 37 mg/L in Golden Russet to 1051 mg/L in Golden Noble, with Winter Banana, Titus Pippin and the lot labeled Kingston Black following. It is low in the russets (median 97 against 244 mg per 100 g of sugars; $p = 0.001$), which thus combine high sorbitol with low quinic acid. Citramalic acid was found above its reporting limit in 41 juices and reaches 178 mg/L in Northern Spy, followed by Montreal Peach and Marie-Joseph d'Othée; L-rhamnitol is present in all but one juice at a median of 58 mg/L and is highest in Pitmaston Pineapple (209 mg/L), Roxbury Russet, American Beauty and Marie-Joseph d'Othée. The identification of these two compounds adds markers of variety that are obtained at no additional cost.

Asparagine is the principal free amino acid of apple juice[183] and the main source of nitrogen available to the yeast. It was detected in 31 of the 53 lots, at a median of 225 mg/L where it was found and at 1156 mg/L in Green Newtown Pippin, followed by Oaken Pin and Baldwin; the supermarket Granny Smith contains 3.4 g/L, about three times the amount found in any of the old varieties. Aspartic acid accompanies it in 17 juices, most abundantly in Yellow Newtown Pippin and Golden Harvey. The amino acid content matters to the cider maker because a juice that is poor in assimilable nitrogen ferments slowly and is the raw material of the naturally sweet ciders of the French tradition (Section 3.4), whereas a juice rich in it ferments rapidly to dryness.[24] The hydroxycinnamoyl esters, expressed as chlorogenic acid, amount to 113 mg/L in the median juice and to 376 mg/L in Winter Banana, and the aromatic integral is highest in Benoni, Winter Banana, Chestnut Crab and Starr. By the standard of bittersweet cider fruit, whose juice may contain several grams of polyphenols per liter,[28,29] all of these values are modest; old dessert varieties are nevertheless reported to

be richer in polyphenols than modern ones,[184,185] and the supermarket Fuji and Gala are indeed among the lowest of the present set in both measures.

	quinic acid	citramalic acid	L-rhamnitol	asparagine	aspartic acid	alanine	xylose (α)	hydroxy-cinnamates	aromatic H (mmol/L)
Golden Noble	1051	47	108	nd	nd	3	364	196	4.1
Winter Banana	700	32	38	nd	nd	nd	193	376	5.6
Chestnut Crab	264	23	117	nd	nd	nd	367	211	4.5
Starr †	517	nd	48	211	nd	13	238	264	4.3
Devonshire Quarrenden	388	7	40	188	nd	nd	413	228	4.2
Benoni †	254	36	69	nd	nd	nd	563	204	5.6
Titus Pippin	640	39	127	493	158	nd	412	169	4.0
Golden Harvey	160	16	58	212	318	19	57	100	3.0
Yellow Newtown Pippin	97	nd	44	240	552	4	174	109	2.5
Pitmaston Pineapple	115	4	209	131	224	2	116	158	3.1
Roxbury Russet	51	nd	161	nd	128	nd	117	194	3.3
Granny Smith ‡	306	8	74	3401	nd	27	86	99	4.2
Fuji (2015) ‡	110	77	38	545	157	8	136	46	0.1
Jewett Red (Nodhead)	230	58	21	225	214	12	197	107	nd
Gala (2015) ‡	159	nd	22	283	174	32	70	61	1.3
Oaken Pin	405	7	44	790	227	5	59	77	2.9
Goof	315	8	67	137	nd	7	210	73	1.8
Kingston Black †	600	29	68	nd	nd	nd	130	89	3.3
Yates	553	27	66	nd	nd	nd	83	105	2.1
Virginia Beauty	559	21	57	344	92	17	176	68	2.4
Northern Spy	130	178	82	nd	nd	6	159	188	3.2
Montreal Peach	152	137	29	nd	nd	nd	108	147	2.8
Marie-Joseph d'Othée	487	123	146	156	235	nd	141	94	1.1
Ashmead's Kernel	208	83	84	270	202	7	71	134	2.0
Mother	217	63	99	149	140	4	289	91	3.6
Shiawassee Beauty	240	14	90	nd	nd	nd	241	101	1.4
Collamer	395	15	86	nd	nd	nd	348	104	2.3
Quinte	218	41	nd	326	nd	nd	448	223	2.1
Bancroft	165	24	40	nd	219	nd	254	124	2.1
Pink Pearl	195	32	30	nd	nd	nd	240	80	2.2
Pear Apple	391	nd	69	nd	nd	6	143	122	2.5
Dr. Matthews	176	nd	97	nd	nd	nd	181	106	2.7
Edward VII	232	6	69	216	nd	3	152	137	3.1
Ben Davis	351	11	68	326	nd	nd	259	229	3.6
Sweet Winter Pennock	240	nd	45	nd	nd	nd	155	144	3.1
Sutton Beauty	272	nd	47	183	nd	4	103	208	3.7
Wagener	222	21	47	253	144	4	63	69	1.5
Tompkins King	316	21	42	387	nd	nd	117	106	2.2
Rhode Island Greening	326	39	58	177	91	nd	74	114	2.0
King David	206	18	47	118	nd	nd	303	113	1.9
Fall Pippin	114	18	38	546	nd	5	177	112	2.0
Razor Russet	98	nd	39	173	nd	3	149	51	1.4
Esopus Spitzenburg	350	28	48	72	93	nd	219	nd	2.6
Kidd's Orange Red	370	15	31	nd	75	nd	275	49	1.4
Chieftain	420	13	57	nd	nd	9	191	50	0.5
Opalescent	313	7	38	210	nd	4	265	100	nd
Green Newtown Pippin	338	63	81	1156	nd	36	115	135	3.4
Baldwin	302	nd	92	761	237	9	50	111	2.6
GoldRush	189	20	47	454	nd	5	104	54	1.8
Golden Russet	37	nd	71	nd	nd	nd	106	96	0.8
Scarlet Crofton †	74	nd	53	250	nd	10	352	147	2.2
Ellison's Orange	134	nd	27	238	nd	12	222	89	2.4
Decio	186	20	62	nd	nd	nd	116	196	0.5
American Beauty (Sterling)	196	11	146	174	nd	5	143	128	3.3
Calville Blanc d'Hiver	316	24	72	nd	nd	nd	134	232	3.3
Barnack Beauty	62	18	94	nd	nd	nd	161	220	3.2

Figure 9. Minor constituents of the apple juices standardized to 110 g/L of sugars (mg/L; aromatic protons in mmol/L). The rows are ordered by hierarchical clustering of the relative composition, and each column is shaded on its own scale. nd, below the reporting limit. † juices in which fermentation had begun; ‡ supermarket fruit.

4.5 Juices of local fruit pressed in 2015

A second, smaller set of juices was recorded in July and October 2015 and is preserved with the cider spectra (Table 5). It comprises supermarket dessert varieties, juices that had been left exposed to air for five days, and fruit from local trees pressed in the laboratory. The dessert varieties (Golden Delicious, Granny Smith, Gala and Fuji) contain little sucrose after exposure, 1 to 14 % of the sugars, and their glucose share of 24 to 32 % corresponds to ratios of fructose to glucose close to two. The fruit collected from local trees is of a different kind. The juice labeled "Bitter Yard" contains 30 g of malic acid per 100 g of sugars, three to four times the amount found in a dessert apple, together with the largest quinic acid signal of the series; the crab apple combines 11 g of malic acid and 10 g of sorbitol per 100 g of sugars with the strongest aromatic intensity, and Niagara Early is similarly acidic and phenolic. The spectra of three of these juices are compared on a common scale in Figure 10. The expansions show what distinguishes fruit of cider quality from dessert fruit at the level of the spectrum: a malic acid multiplet several times more intense relative to the sugars, the multiplets of quinic acid near 1.9 and 2.1 ppm, and, between 6.9 and 7.7 ppm, the olefinic doublets of the caffeoyl group of chlorogenic acid on a broad envelope that is attributed to flavan-3-ols and procyanidins.[47,144,148]

Table 5. Relative composition of apple juices pressed and measured in 2015.

Juice	Sucrose (%)	Glucose (%)	Fructose (%)	Malic acid (g/100 g sugars)	Sorbitol (g/100 g sugars)	Quinic acid (g/100 g sugars)	Aromatic H (mmol/L)
Golden Delicious, exposed 5 days	3.9	31.0	65.1	2.3	4.0	0.04	1.1
Granny Smith, exposed 5 days	13.3	31.8	54.9	6.3	4.3	0.11	2.3
Gala, exposed 5 days	12.3	24.2	63.5	2.1	3.7	0.12	0.0
Gala — freeze-dried tube	13.7	18.6	67.7	2.0	3.4	0.14	3.9
Fuji, exposed 5 days	1.2	31.3	67.5	nd	4.4	0.10	3.9
Fuji — freeze-dried tube	4.7	30.0	65.3	1.4	3.0	0.10	0.2
"Bitter Yard" apple	22.7	9.4	67.9	30.1	2.2	3.72	10.0
Niagara Early	20.7	12.9	66.4	11.0	1.6	1.37	13.9
Niagara Early — freeze-dried tube	21.7	12.4	65.9	10.8	1.2	1.47	13.5
Crab apple (Bushy Hill Farm, picked 4 Oct, pressed 13 Oct 2015)	18.9	27.9	53.2	11.0	10.2	0.22	18.2
Macoun (Bushy Hill Farm, picked 4 Oct, pressed 13 Oct 2015)	14.4	19.1	66.5	3.7	1.7	0.30	9.0
Garnet Spy (Bushy Hill Farm, picked 4 Oct, pressed 13 Oct 2015)	40.2	10.4	49.4	3.0	0.9	0.08	0.2

Sugar shares in % by mass of fructose + glucose + sucrose. The freeze-dried preparations are separate tubes of the same juice; those of Gala and Fuji are the two juices of 2015 included in the survey of this section (Appendix D). nd, not detected.

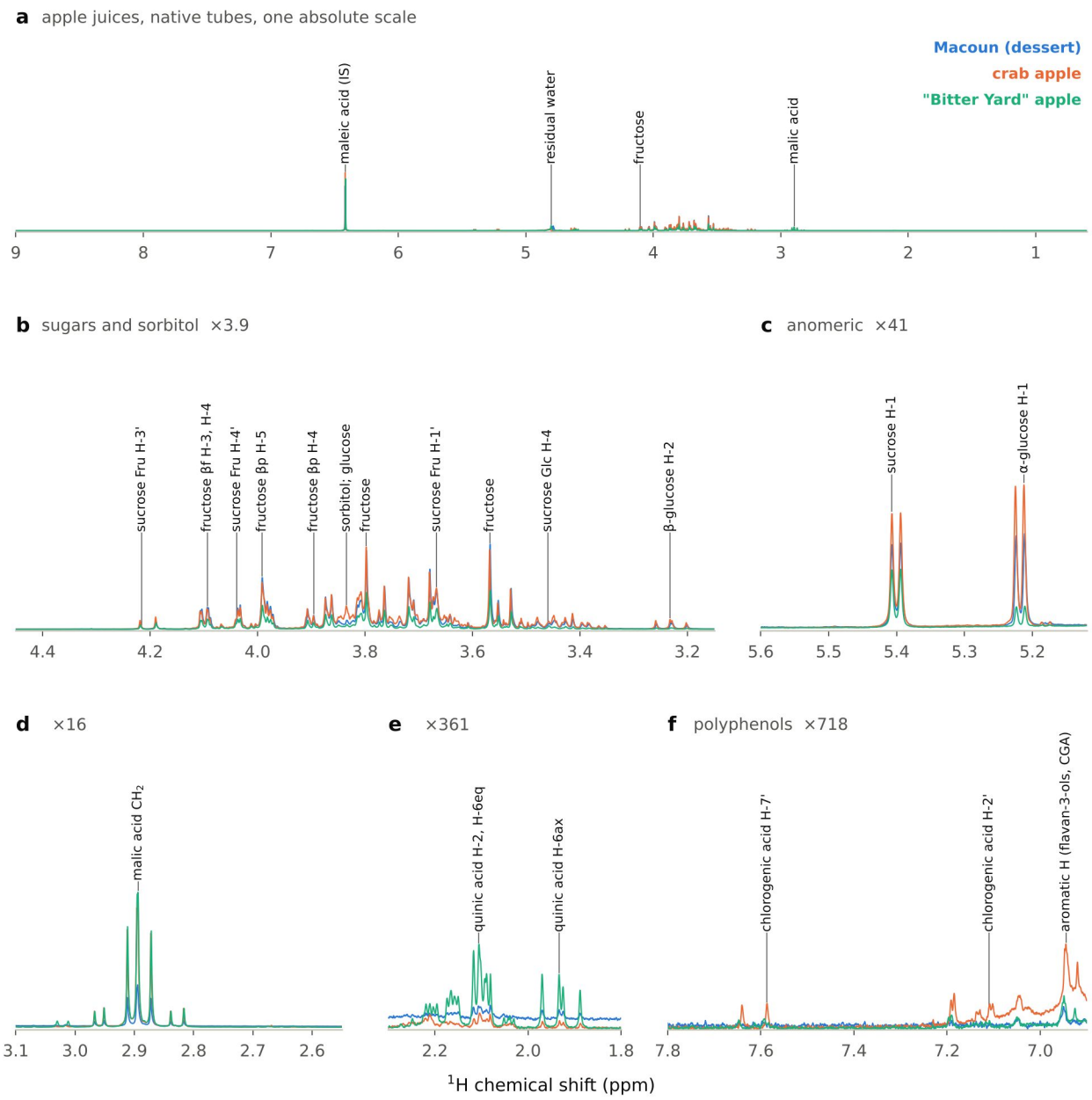

Figure 10. ¹H NMR spectra (300 MHz, D_2O) of three apple juices measured in 2015, drawn on one absolute scale: Macoun, a dessert variety, a crab apple, and the fruit labeled "Bitter Yard". (a) Full spectra on scale. (b) to (f) Expansions with the vertical magnification given in each panel relative to panel a; sugars and sorbitol, anomeric signals, malic acid, quinic acid and the polyphenol region.

4.6 Implications for cider making

For the cider maker the ratio of acid to sugar in the juice is decisive, because the sugar is converted to alcohol while the acid remains. In a juice standardized to 110 g/L of sugars the old varieties contain 3.7 to 11.4 g/L of malic acid, and 48 of the 53 lots lie above the 4.5 g/L that separates sharp from sweet fruit in the Long Ashton classification (Table 1). Their polyphenol signals are those of dessert fruit. Nearly all of these varieties are therefore sharps: fruit that gives a clean, acidic, lightly structured cider of the kind that dominates the regional craft products of Section 8, and that would need bittersweet fruit in the blend, or a malolactic fermentation, to approach the traditional European styles. Three further properties can be read

from the same spectrum. The low glucose share has the consequence that a cider whose fermentation is arrested retains mostly fructose, the sweetest of the three sugars, and the high ratios of fructose to glucose found in several semi-sweet craft ciders (Section 8.3) are consistent with fruit of this composition. The sorbitol content, which differs more than tenfold between varieties, predicts the body that a completely dry cider will retain. And the asparagine signal gives an early indication of the nitrogen available to the yeast, and with it of the speed of the fermentation.

4.7 Chemometrics of the juice spectra: choice of pre-processing

The multivariate analysis of the juices addresses a question that differs from that asked of the ciders. A cider records decisions of its maker; a juice records the variety, and the object is to learn which representation of the spectrum conveys the differences between varieties and how far the composition can be predicted from the spectrum without any assignment. Seven descriptions of each spectrum were compared: uniform bins of 0.04 and 0.01 ppm, optimized buckets of nominal width 0.04 and 0.02 ppm whose boundaries fall on minima of the mean spectrum,[186,187] a table of 31 targeted integrals, the areas of 222 lines obtained by fitting each spectrum with Lorentzian lines at positions common to the whole set,[188,189] and the table of fitted concentrations. Each was combined with four normalizations (absolute intensities, intensities of the juice standardized to 110 g/L of sugars, normalization to total area and probabilistic quotient normalization[190]) and, after mean-centering, with seven scalings: none, Pareto, unit variance, logarithmic transformation, range, variable stability (VAST) and level scaling.[191,192] The regions of water, of the standard and of residual ethanol were excluded.

Two criteria were applied, neither of which uses a class label. The first asks how faithfully the distances between juices in the processed spectra reproduce the distances between them in the table of relative composition, measured by the rank correlation of the two distance matrices in the manner of the Mantel test.[193] The second is a control: whether the spectrometer on which a juice was measured can be recognized from its nearest neighbor more often than the share of the larger group, 0.80, would allow. The results are collected in Figure 11. The control is negative for every pre-processing of practical interest. With centering alone and with Pareto, unit variance and VAST scaling the spectrometer is recognized no more often than the base rate, so that the two instruments leave no signature in the data and their spectra can be analyzed together; the base rate is exceeded, by 0.09 at most, in 11 of the 182 combinations, nearly all of which use the logarithmic transformation or level scaling, the two treatments that give the greatest weight to the smallest variables. The fidelity to composition depends strongly on the normalization. With absolute intensities it does not exceed 0.42 for any bucketed description, because the first component of such data is the total amount of sugar; after normalization to total area it rises to 0.62 for optimized buckets with Pareto scaling and to 0.65 with VAST scaling. The fitted lines reach 0.70 to 0.71 when every line is given equal weight, the highest values among the descriptions of the spectrum, while the logarithmic transformation and level scaling perform worst throughout. Uniform bins and optimized buckets are again practically equivalent at 300 MHz, as they were for the ciders (Section 11.4). On this evidence optimized buckets, normalization to total area, mean-centering and Pareto scaling were adopted for the unsupervised analysis of the relative composition, and absolute intensities where the amount of sugar itself is of interest.

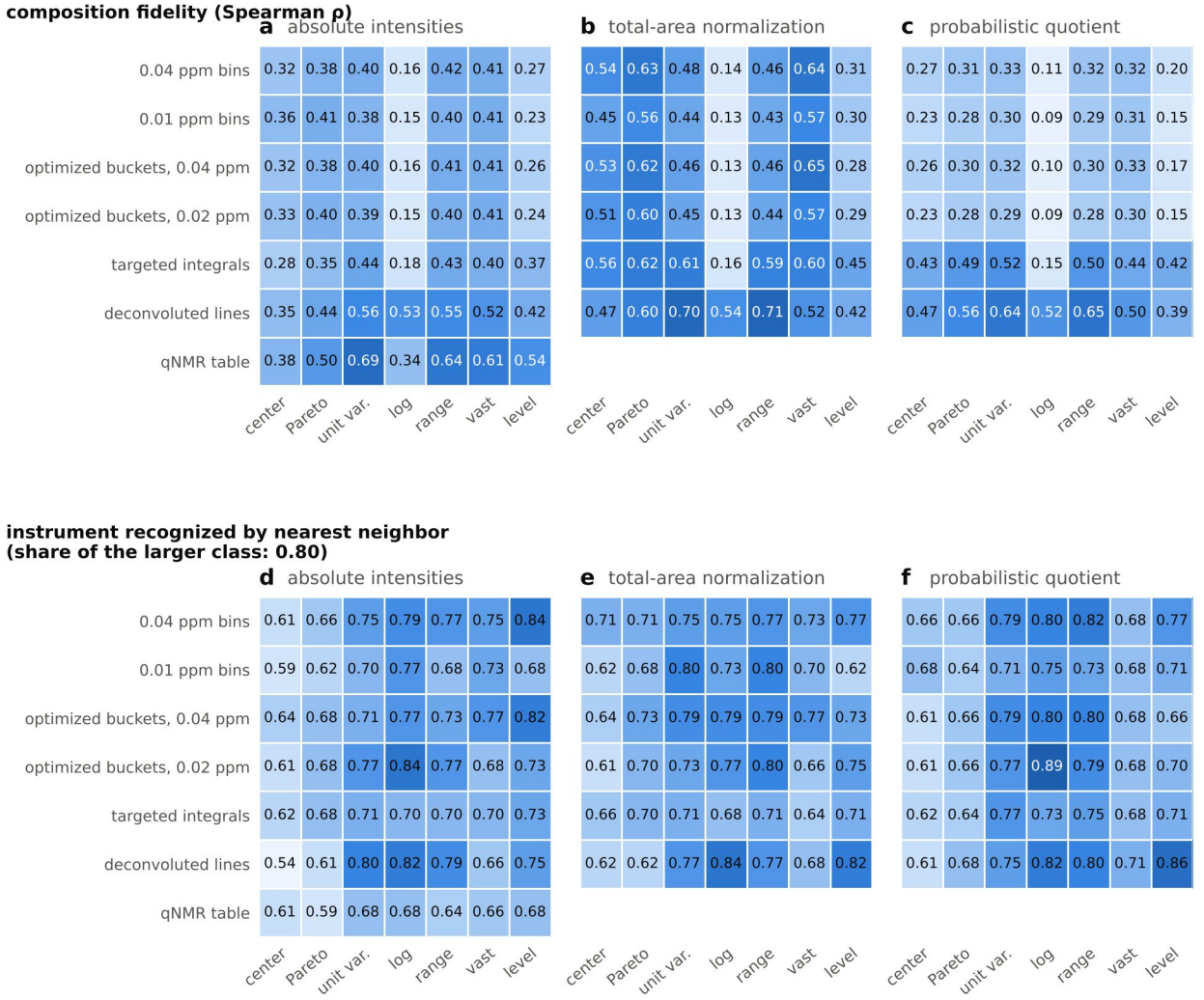

Figure 11. Choice of pre-processing for the apple juice spectra. Upper row: fidelity of the processed spectra to the relative composition (rank correlation between the distance matrix of the spectra and that of the composition table) for seven descriptions of the spectrum and seven scalings, with (a) absolute intensities, (b) normalization to total area and (c) probabilistic quotient normalization. Lower row, (d) to (f): share of the juices whose nearest neighbor was measured on the same spectrometer; values at or below the share of the larger group (0.80) indicate that no instrumental signature is present.

4.8 Principal components, clustering, t-SNE and UMAP

The first four principal components of the normalized, Pareto-scaled buckets carry 49.8, 20.2, 11.4 and 7.7 % of the variance (Figure 12). Their meaning follows from the loadings and from the correlation of the scores with the quantitative composition. The first component is the sucrose share of the sugars (rank correlation 0.98) set against fructose (−0.90): its loadings are positive at the anomeric doublet of sucrose and at the signals of its fructosyl unit and negative at the multiplets of free fructose. The second component separates glucose (0.61) and sorbitol (0.41) from fructose, and the third is the malic acid multiplet (0.62). The juices are distributed continuously along the first component, from Chestnut Crab, Starr and Fuji at the sucrose-poor end to Baldwin, Decio and Green Newtown Pippin at the other, and the supermarket Fuji and Gala are displaced along the second by their glucose. When the absolute intensities are analyzed instead, a

single component with 77.9 % of the variance reproduces the total sugar content (rank correlation 0.99) and the pattern just described moves to the second component.

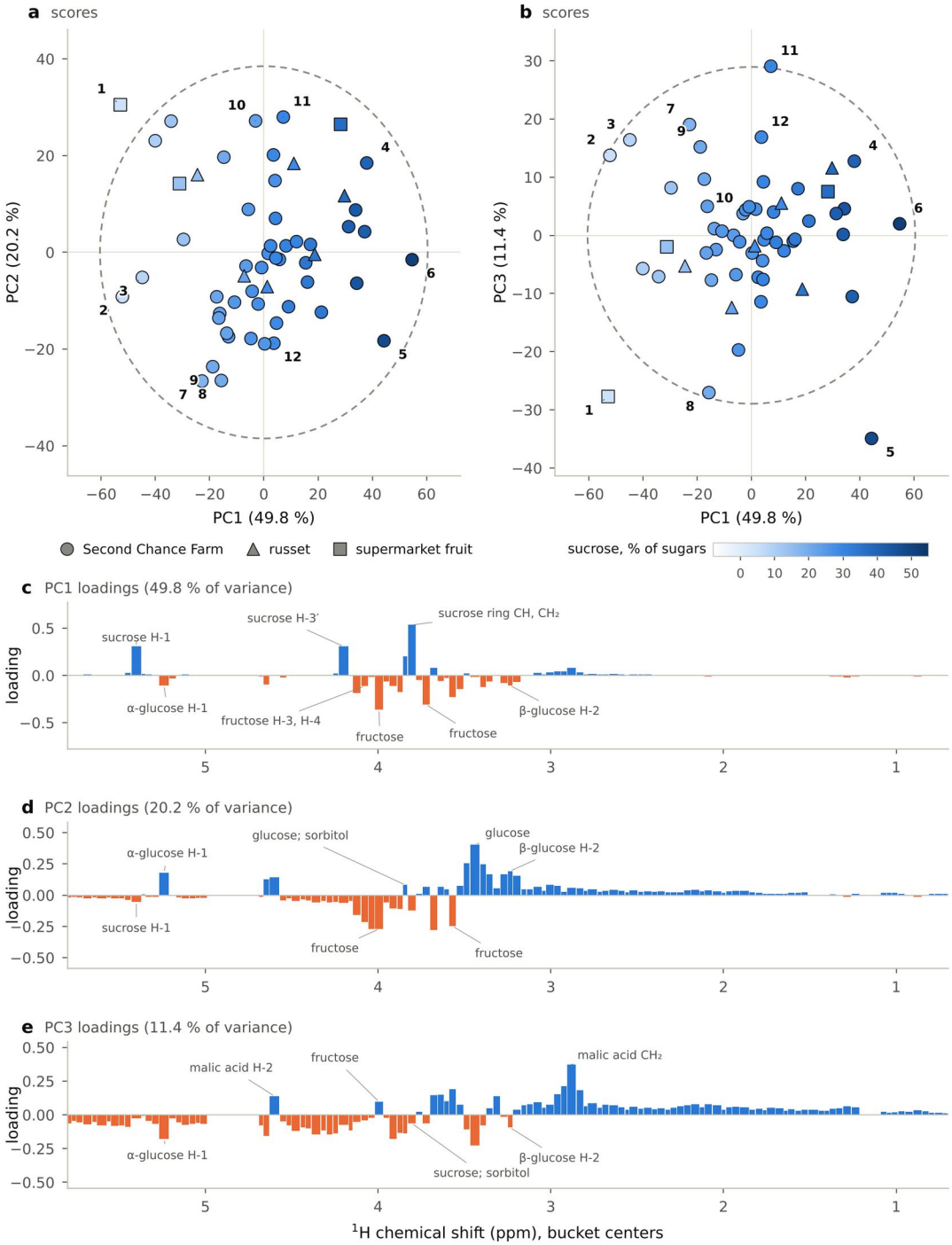

Figure 12. Principal component analysis of the apple juice spectra (optimized buckets, normalized to total area, mean-centered, Pareto-scaled). (a) Scores on the first and second and (b) on the first and third components; color and marker give the region of origin of the variety. (c) to (e) Loadings of the first three components against the chemical shift, with the assignments of the principal features. Numbered points: 1, Fuji (2015); 2, Chestnut Crab; 3, Starr; 4, Green Newtown Pippin; 5, Decio; 6, Baldwin; 7, Shiawassee Beauty; 8, Jewett Red (Nodhead); 9, Collamer; 10, Oaken Pin; 11, Golden Noble; 12, Bancroft.

Hierarchical clustering by Ward's method[194] confirms that the set has little cluster structure (Figure 13). The cophenetic correlation of the dendrogram[195] is 0.64, the mean silhouette width[196] does not exceed 0.32 for any number of clusters between two and eight, and the adjusted Rand index[197] between the clustering of the whole set and clusterings of random subsets of 85 % of the juices, a measure of stability, is 0.44, 0.61 and 0.72 for two, three and four clusters. A division into three is nevertheless instructive, because it is readily interpreted (Table 6). Nine juices form a sucrose-poor, glucose-rich group: Starr, Chestnut Crab, Devonshire Quarrenden, Winter Banana, Pitmaston Pineapple, Goof, the lot labeled Kingston Black, and the supermarket Gala and Fuji. Nine others form a sucrose-rich group of higher acidity and, on average, the highest asparagine content: Calville Blanc d'Hiver, Green Newtown Pippin, Golden Harvey, Decio, Baldwin, GoldRush, American Beauty, Ellison's Orange and Granny Smith. The remaining 38 occupy the middle. Neither the region of origin nor the age of the variety is related to this division, the correlations of the year of origin with the composition being weak throughout (rank correlation below 0.4 in absolute value), and the two spectrometers are distributed evenly over the three groups.

Table 6. Mean composition of the three clusters obtained by Ward's method from the normalized apple juice spectra.

Quantity	Cluster 1 (n = 9)	Cluster 2 (n = 9)	Cluster 3 (n = 38)	p (Kruskal-Wallis)
Sucrose (% of the sugars)	11.6	42.8	27.3	<0.001
Glucose (% of the sugars)	20.3	10.6	12.3	<0.001
Fructose (% of the sugars)	68.1	46.6	60.4	<0.001
Fructose / glucose	3.6	4.6	5.3	0.009
Malic acid (g per 100 g of sugars)	4.7	7.3	5.7	0.034
Sorbitol (g per 100 g of sugars)	3.1	3.1	2.5	0.093
Quinic acid (mg per 100 g of sugars)	320	215	259	0.564
Asparagine (mg per 100 g of sugars)	151	646	141	0.148
Xylose, α anomer (mg per 100 g of sugars)	189	104	189	0.024
Hydroxycinnamates as chlorogenic acid (mg per 100 g of sugars)	152	115	112	0.815
Sugars in the juice as measured (g/L)	199	201	201	0.874

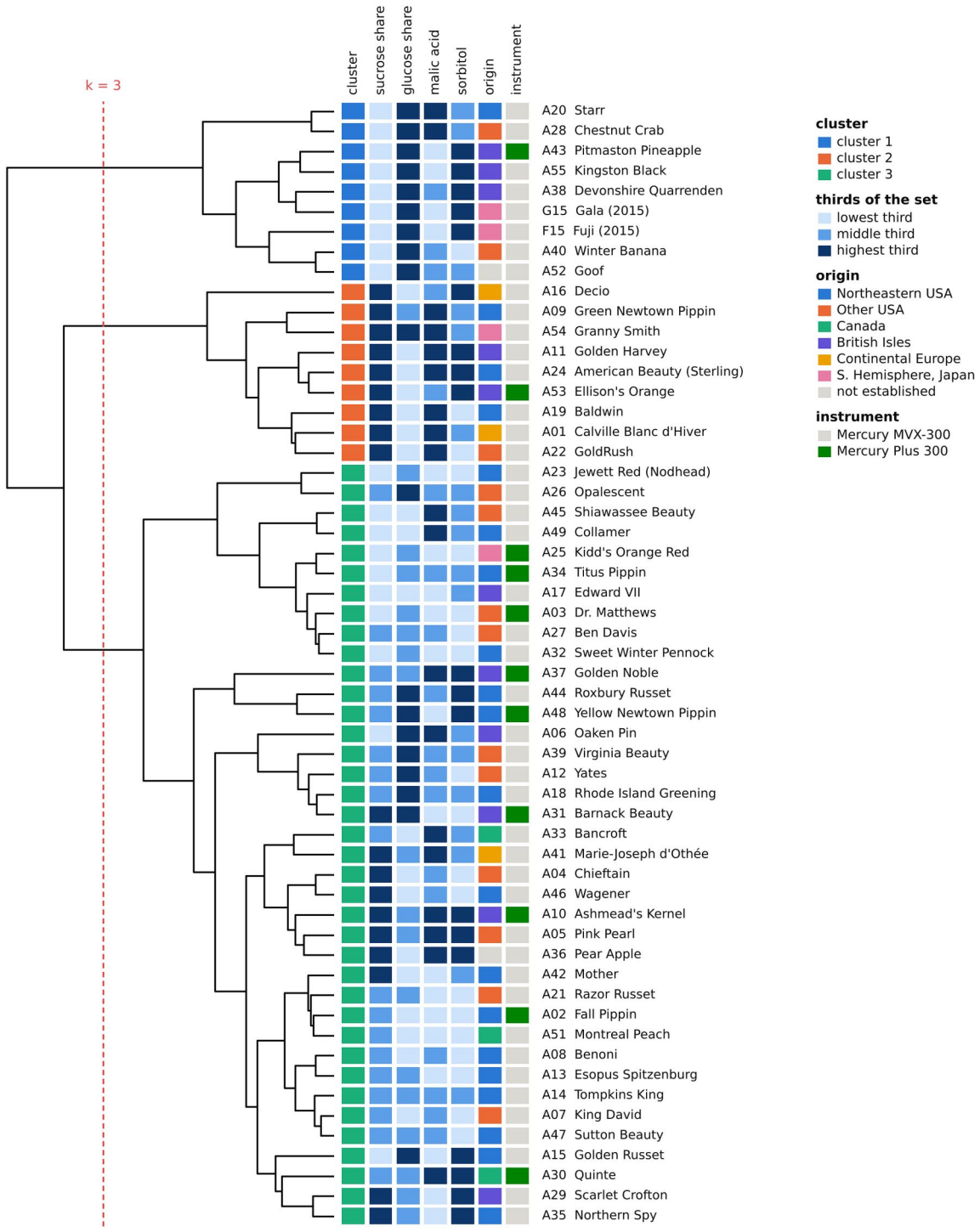

Figure 13. Hierarchical clustering of the apple juice spectra (Ward's method on the scores of the first ten principal components; optimized buckets, normalized to total area, Pareto-scaled). The broken line marks the division into three clusters. The strips give, for each juice, the cluster, the third of the set into which its sucrose share, glucose share, malic acid and sorbitol fall, the region of origin of the variety and the spectrometer.

The same picture is obtained from the non-linear maps (Figure 14). t-distributed stochastic neighbor embedding[198] and uniform manifold approximation and projection[199] preserve the neighborhoods of the

bucket space slightly better than the plane of the first two principal components (trustworthiness[200] 0.94 and 0.91 against 0.91), but they reveal no structure that the principal components had concealed: the juices form one elongated cloud along which the sucrose share changes monotonically, the two small groups of the cluster analysis lie at its ends, and the spectrometers are intermingled. When the table of relative composition is clustered in place of the spectra, with every constituent given equal weight, the minor constituents come to the fore. A group of four varieties rich in sorbitol and aspartic acid separates (Golden Harvey, Pitmaston Pineapple, Roxbury Russet and Yellow Newtown Pippin, three of them russets), as does a group of seven that are rich in quinic acid, xylose and polyphenols and poor in sucrose (Benoni, Starr, Chestnut Crab, Titus Pippin, Golden Noble, Devonshire Quarrenden and Winter Banana). The spectrum as a whole thus classifies the juices by their sugars, while the targeted analysis is needed to bring out the features that are peculiar to a variety.

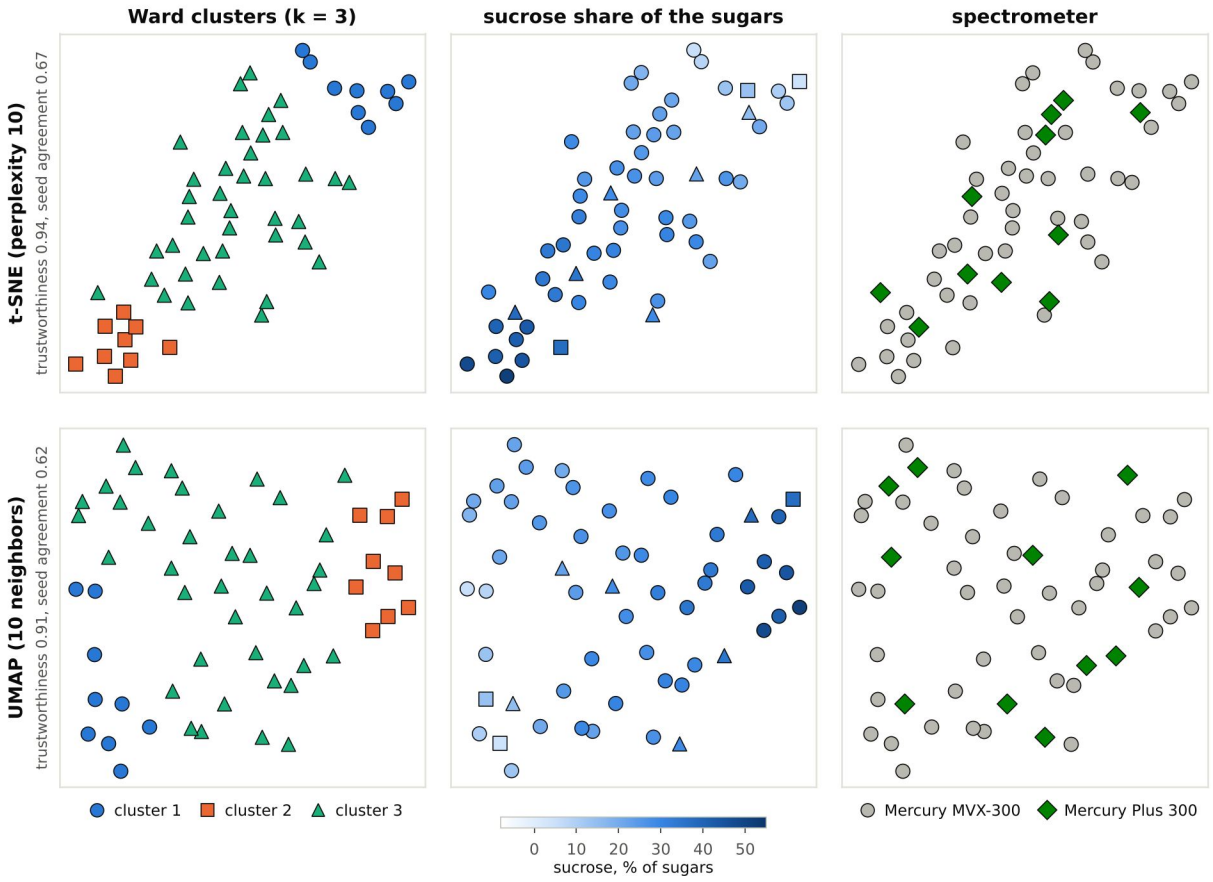

Figure 14. Non-linear maps of the apple juice spectra: t-SNE (upper row, perplexity 10) and UMAP (lower row, 10 neighbors), each shown three times with the juices marked by the Ward cluster, by the sucrose share of the sugars and by the spectrometer. Trustworthiness and the agreement of the neighborhoods between random starts are given at the left.

4.9 PLS regression of the composition on the juice spectra

Partial least squares regression[201,202] was used to predict the composition determined by qNMR from the spectrum treated as an unassigned pattern. The calculation was carried out afresh for the juices and independently of the models for cider (Section 14). Six spectral descriptions (the two kinds of bins, the two kinds of optimized buckets, the targeted integrals and the fitted lines) were combined with seven scalings and with two formulations of the problem: prediction of the concentration per liter of juice from spectra on the absolute scale, and prediction of the composition of the standardized juice from spectra normalized

without reference to the standard, by total area or by probabilistic quotient. Each of the 1680 models was validated by nested cross-validation, with seven outer folds repeated three times and the number of latent variables, up to twelve, chosen in an inner loop;[203,204] scaling parameters were calculated from the training folds only. Performance is expressed as the root mean square error of cross-validation (RMSECV), Q^2 and the ratio of the standard deviation of the reference values to the RMSECV (RPD).[205]

The outcome of the screening is summarized in Figure 15, and the best model for each constituent is listed in Table 7 and shown in Figure 16. The principal constituents are reproduced with a precision that approaches that of the reference analysis: sucrose with an RMSECV of 0.87 g/L (RPD 28), glucose 0.94 g/L (13), fructose 3.5 g/L (10), the sum of the sugars 3.4 g/L (16), sorbitol 0.45 g/L (11) and malic acid 0.32 g/L (14). Citramalic and quinic acids follow with ratios of 6.8 and 4.7; asparagine, xylose and L-rhamnitol, whose signals are small and partly overlapped, are estimated with ratios close to 3, which suffices to rank the juices; the two measures of polyphenols reach only about 2; and aspartic acid, a weak doublet on the flank of other signals, is not predicted at all (RPD 1.3). The regression coefficients of the models for the principal constituents (Figure 17) are concentrated on the signals of the compound concerned, the anomeric doublet and the fructosyl signals for sucrose, the anomeric and H-2 signals for glucose, the methylene multiplet for malic acid, so that the models rest on the chemistry and not on incidental correlations within this set of juices.

Table 7. PLS regression of the composition of the apple juices on their spectra: nested cross-validation (seven outer folds, three repetitions). Best model for each constituent, the single description recommended for all constituents (bins of 0.04 ppm, absolute intensities, VAST scaling), and the prediction of the composition of the standardized juice from spectra normalized to total area (optimized buckets, Pareto scaling).

Constituent	Mean	SD	Best description and scaling	LV	RMSECV	Q^2	RPD	RMSECV, single description	RPD, single description	RPD, standardized juice
Sucrose (g/L)	55.19	24.50	bins of 0.04 ppm, centered	11	0.87	0.999	28.1	1.93	12.7	7.1
Glucose (g/L)	26.75	12.08	bins of 0.01 ppm, centered	7	0.94	0.994	12.8	1.08	11.2	6.3
Fructose (g/L)	118.46	34.40	fitted lines, Pareto	5	3.48	0.990	9.9	6.39	5.4	4.5
Total sugar (g/L)	200.40	54.06	fitted lines, Pareto	4	3.40	0.996	15.9	5.07	10.7	–
Sorbitol (g/L)	5.82	5.13	bins of 0.01 ppm, centered	11	0.45	0.992	11.3	1.25	4.1	6.3
Malic acid (g/L)	11.23	4.35	optimized buckets, centered	8	0.32	0.994	13.6	0.53	8.2	6.3
Quinic acid (mg/L)	498	343	optimized buckets, Pareto	12	73	0.954	4.7	82	4.2	4.4
Citramalic acid (mg/L)	51	82	targeted integrals, level	8	12	0.978	6.8	45	1.8	2.7
L-Rhamnitol (mg/L)	124	95	optimized buckets, 0.02 ppm, VAST	12	34	0.872	2.8	61	1.5	0.9
Asparagine (mg/L)	414	689	optimized buckets, centered	12	221	0.895	3.1	336	2.1	2.4
Aspartic acid (mg/L)	138	254	targeted integrals, level	7	202	0.355	1.3	252	1.0	0.9
Xylose (α anomer) (mg/L)	331	172	optimized buckets, 0.02 ppm, VAST	11	55	0.894	3.1	75	2.3	3.1
Hydroxycinnamates, as chlorogenic acid (mg/L)	235	132	bins of 0.04 ppm, VAST	10	58	0.807	2.3	58	2.3	1.6
Aromatic protons (mmol/L)	4.61	2.45	bins of 0.04 ppm, VAST	10	1.26	0.732	1.9	1.26	1.9	1.3

Mean and SD of the reference values (¹H qNMR, per liter of juice as measured). LV, number of latent variables of the model fitted to all juices. The total sugar content is fixed at 110 g/L in the standardized juice and is not a target on that basis.

Three general conclusions can be drawn from the screening. First, no single description of the spectrum is best for every constituent, but the differences among the bucketed descriptions are small; bins of 0.04 ppm with absolute intensities and VAST scaling gave the best average rank and may be recommended where one model structure is to serve for all constituents, at the cost of a moderate loss for those that a specialized description predicts best (Table 7). The fitted lines are superior for fructose and for the sum of the sugars, whose signals overlap most severely, and the short table of targeted integrals with level scaling for citramalic acid, whose singlet would otherwise be shared with a neighbor. Second, the scaling matters more than the bucketing, and its best choice depends on the size of the signal. For the principal constituents, whose signals dominate the variance, mean-centering without further scaling gives the smallest errors, and centering itself lowers the error by about 3 % in the median. Pareto scaling leaves these models practically unchanged but improves those of the minor constituents considerably, by more than half for citramalic and quinic acids on optimized buckets, and VAST scaling, which favors variables of stable intensity, gives the best models for the weakest signals: L-rhamnitol, xylose, the hydroxycinnamates and the aromatic integral.

Third, the composition relative to the sugars can be predicted without an internal standard: from spectra normalized to total area the shares of sucrose, glucose and fructose and the contents of sorbitol and malic acid in the standardized juice are obtained with ratios of performance to deviation between 4.5 and 7.1. In a test with one hundred permutations of the reference values, 24 of the 27 models of the two recommended configurations were significant at the smallest attainable probability of 0.01; the exceptions are aspartic acid on both bases and L-rhamnitol on the relative basis, for which the reduced description retains no information. For a laboratory that receives juices routinely, such models offer a rapid first estimate of sugars, sorbitol and acidity from a bucketed spectrum, to be followed by the targeted evaluation where the minor constituents are of interest.

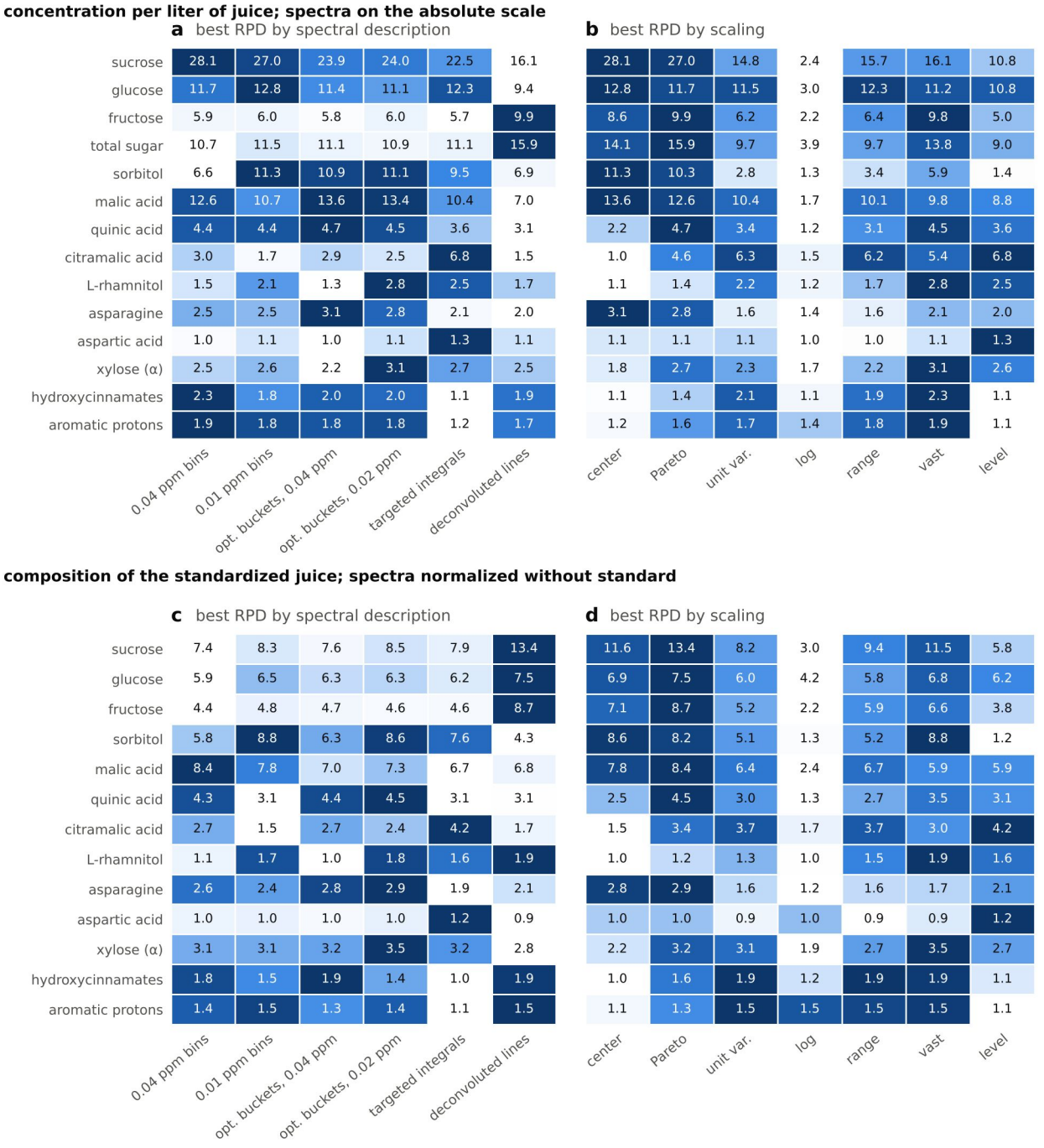

Figure 15. Screening of PLS models for the apple juices: best ratio of performance to deviation (RPD) obtained in nested cross-validation for each constituent, (a, c) by description of the spectrum and (b, d) by scaling. Upper row: concentration per liter of juice from spectra on the absolute scale. Lower row: composition of the standardized juice from spectra normalized without the standard. Each row is shaded on its own scale.

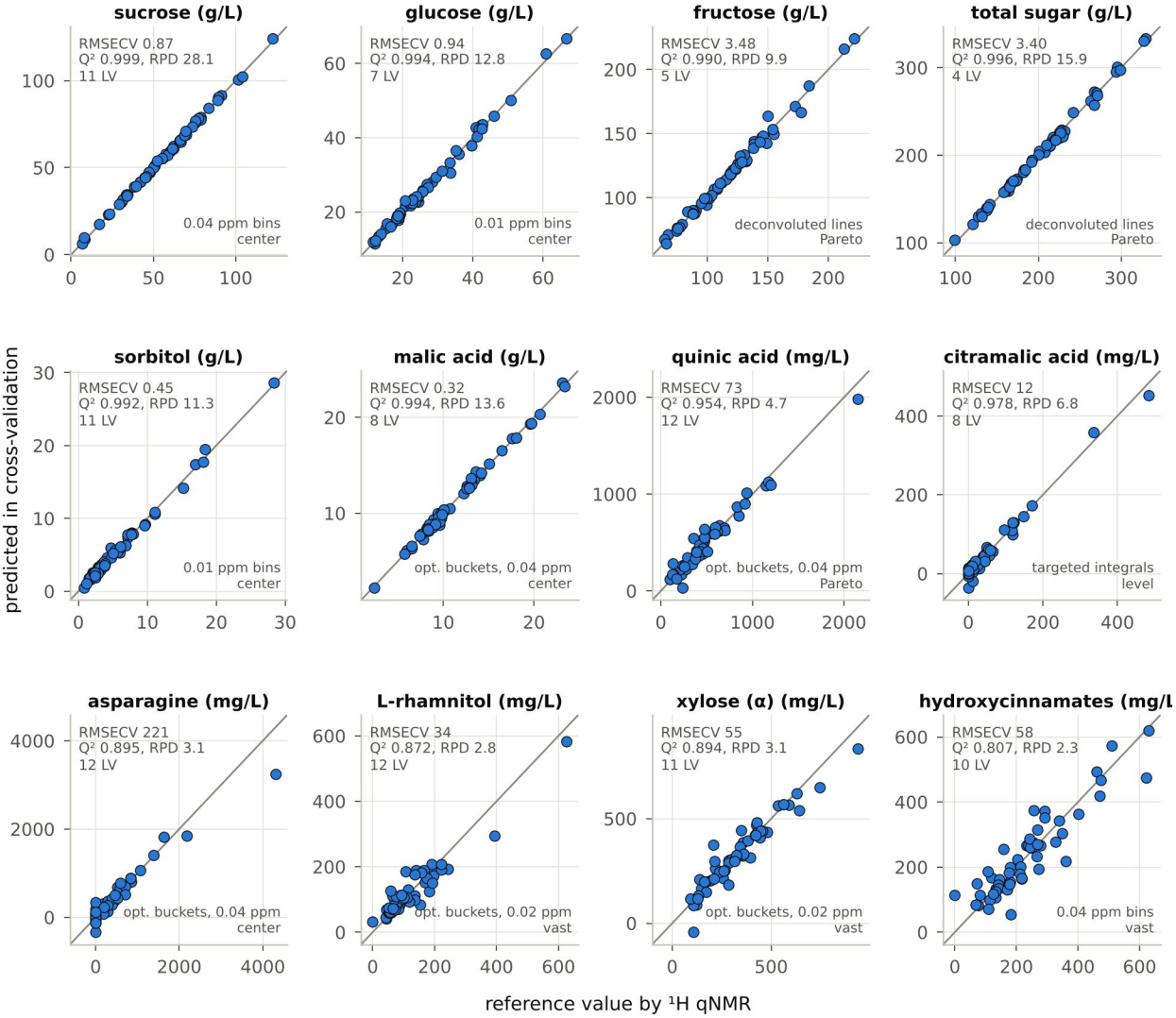

Figure 16. Concentrations predicted by PLS regression in nested cross-validation against the reference values by ¹H qNMR for twelve constituents of the apple juices (per liter of juice as measured). For each constituent the best model of the screening is shown, with its error, Q^2, RPD, number of latent variables, spectral description and scaling. The line is the identity.

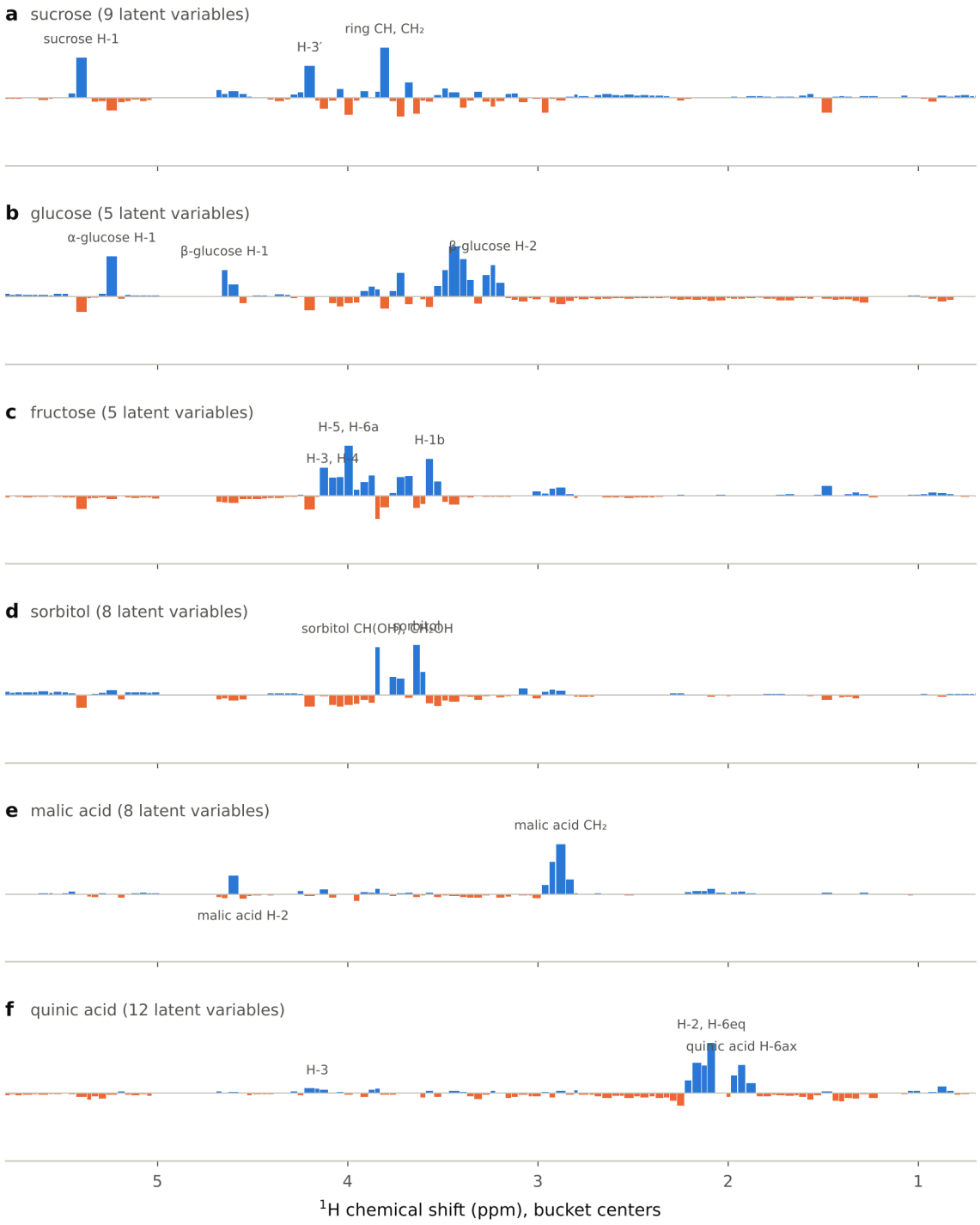

Figure 17. Regression coefficients of PLS models for six constituents of the apple juices (optimized buckets, absolute intensities, mean-centered, Pareto-scaled) against the chemical shift, with the assignments of the largest coefficients. Positive coefficients blue, negative orange.

5. Materials and Methods

5.1 Samples, consolidation and removal of duplicates

The cider spectra are held in eleven workspaces of the PNA NMR Processor (version 2.6.4). Each workspace is an archive that contains the original Mnova documents, the processed spectra and a manifest with the acquisition comments and stored integrals. Together the workspaces hold 553 spectrum records. Because the same Mnova documents had been loaded into several workspaces, most spectra occur more than once; identical acquisitions were recognized by the path of the original free induction decay and by a hash of the data, which reduced the records to 285 distinct spectra. Of these, 212 are spectra of ciders, 23 of the apple juices of 2015 and 50 of other materials (beers, wines and reference compounds), the last of which were used only where they serve as pure-component references. The spectra of the 56 apple juices of Section 4 are held in a twelfth workspace.

The identity of each cider was established from the acquisition comment (producer, product, declared alcohol content and, where given, lot code) and the following rules were applied so that each cider enters the analysis once. One spectrum of the cider as received (the native tube) and one of the freeze-dried cider (the freeze-dried tube) represent each product. Repeated acquisitions of the same tube, sixteen spectra of fifteen ciders, were set aside as replicates for the estimation of precision. Where the same product had been bought and analyzed twice without any indication of a different lot, the later analysis was kept; this applies to Citizen Cider Unified Press, Citizen Cider The Dirty Mayor and Magners Original Irish Cider, and the two analyses of the Citizen Cider products are compared in Section 7.2. Products for which distinct lots were deliberately analyzed, namely the fourteen cans of five Stormalong products from different lots and the batches and cans of Downeast Original Blend, were retained as separate entries because they document the variation between lots. The result is a set of 96 ciders from 36 producers, 87 of them with both tubes and 9 with the native tube only. The complete list, with the identifiers used in the workbooks and in the consolidated workspace, is given in Appendix A.

For the discussion the ciders were assigned to market segments on the basis of the producer: regional craft (63 ciders), mass-market brands of national or international distribution (19), traditional European ciders (7), ciders made by home or experimental cider makers (4), non-alcoholic ciders (2) and one fresh, unpasteurized apple cider that had been held for two months and had begun to ferment. The measurements fall into four campaigns, 2013 to 2015 (59 ciders), 2017 (5), 2020 (14) and 2022 to 2026 (18), which differ in acquisition parameters (Table 8) and, unavoidably, in the products that were on the market.

5.2 Sample preparation

Two NMR samples were prepared from each cider. For the native tube, 175 µL of cider, freed of carbon dioxide by agitation, was mixed with the maleic acid standard solution in D_2O (100 µL until 2020, 50 µL from 2022) and D_2O to a volume suitable for a 5 mm tube. For the freeze-dried tube, 1000 µL of cider was lyophilized and the residue was redissolved in D_2O containing the same volume of standard solution. Freeze-drying removes water and most of the ethanol, concentrates the non-volatile constituents about sixfold relative to the native tube, and uncovers the signals of 2,3-butanediol, glycerol and sorbitol that are hidden by ethanol in the native spectrum; it also removes the volatile constituents, which is why both tubes are needed. In one case (J.K.'s Scrumpy, 2013) the standard was weighed directly into the native tube.

5.3 NMR acquisition

The spectra were acquired at 27 °C with a one-pulse sequence on two 300 MHz spectrometers: a Varian Mercury MVX-300 with a 5 mm ATB probe (¹H frequency 299.67 MHz until 2015 and 299.99 MHz from 2020), on which 74 of the ciders were measured, and a Varian Mercury Plus 300 (299.94 MHz), which was used for one or both tubes of the remaining 22, all of them from the campaigns of 2013 to 2015 and of 2017.

The acquisition parameters of the four campaigns are summarized in Table 8. Excitation pulses of 8 to 12 µs were used for all but three spectra, and the repetition time, the sum of the relaxation delay and the acquisition time, was 14 to 24 s for all but four. Water suppression was applied to the native tubes, as described previously,[7] and the freeze-dried tubes, which contain only the residual water of the solvent, were acquired without it. The free induction decays (64,000 complex points) were zero-filled to 262,144 points.

Table 8. Measurement campaigns: number of ciders, spectrometer, internal standard and principal acquisition parameters.

Campaign	Ciders (native / freeze-dried tubes)	Spectrometer	Standard solution per tube	Maleic acid delivered (µmol)	Pulse width (µs)	Relaxation delay (s)	Acquisition time (s)	Spectral width (Hz)	Transients
2013–2015	59 / 53	MVX-300; Mercury Plus 300	100 µL	79.95	10 (8–12)	7	8.0	8000	32–256
2017	5 / 5	Mercury Plus 300; MVX-300	100 µL	79.95	10	7	8.9	7194	64–128
2020	14 / 14	MVX-300	100 µL	79.95	10	10	4.27	14,993	20–64
2022–2026	18 / 15	MVX-300	50 µL	39.98	8 (8.5)	15–20	4.27	14,993	16–256

Exceptions: J.K.'s Scrumpy (2013), 7.6 mg of maleic acid weighed into the native tube (64.82 µmol); Hudson North Peach Mango (2022), 100 µL of standard solution.

5.4 Internal standard

Maleic acid is a customary internal standard for qNMR in aqueous solution. Its two olefinic protons give a singlet near 6.4 ppm in a region that is free of cider signals, and it is available in certified purity.[206–209] The standard solution used throughout contained 93.7 mg of maleic acid of 99 % purity per milliliter of solution in D_2O. A volume of 100 µL thus delivered 9.37 mg of the standard substance, or 79.95 µmol of maleic acid, to the tube, and 50 µL delivered 39.98 µmol. All spectra were referenced, and all concentrations calculated, with the maleic acid singlet of the same tube.

5.5 Processing, referencing and the common chemical-shift axis

The processed spectra held in the workspaces (Fourier transformation of the zero-filled data and phase correction with the parameters stored at acquisition) were used without further global manipulation; baselines were treated locally in each analytical window. Chemical shifts in cider depend on pH and on the ethanol content, and no shift reference had been added to most samples. Each spectrum was therefore referenced internally by setting the center of the methyl triplet of ethanol to 1.1708 ppm; where ethanol was absent, as in the freeze-dried tubes of dry ciders, the non-alcoholic products and the juices, the singlet of succinic acid (2.671 ppm) or the anomeric doublets of sucrose (5.4005 ppm) and α-glucose (5.218 ppm) served instead. All spectra were then interpolated onto one axis from 10.5 to −0.5 ppm with a point spacing of 0.0002 ppm, on which the targeted measurements and the bucketing for the chemometric analysis were carried out.

5.6 Measurement of peak areas

Three procedures were used, each where it is appropriate, and the same procedure was always applied to the analyte and to the standard with which it is compared (Table 9). Isolated multiplets were integrated numerically between fixed limits after subtraction of a straight baseline drawn between the edges of the window. Signals that overlap with a neighbor were fitted with pseudo-Voigt lines[188] constrained by the known multiplet structure (number of lines, coupling constants and intensity ratios) together with a linear baseline, which separates, for example, the doublet of lactic acid from the ¹³C satellite of ethanol and the

doublet of 2,3-butanediol from the ethanol triplet. The crowded region between 3.18 and 4.30 ppm, in which fructose, glucose, sucrose, glycerol, sorbitol and residual ethanol overlap completely at 300 MHz, was decomposed by non-negative least squares[210] into the spectra of the pure compounds, in the manner of targeted profiling.[189] The reference spectra had been recorded under the same conditions and were shifted rigidly and broadened to match each cider; the contributions of sucrose and glucose were fixed at the values obtained from their resolved anomeric doublets, so that only fructose, glycerol, sorbitol, 1,3-propanediol and residual ethanol were free.

Two constants required by this scheme were measured, not assumed. The α-pyranose fraction of glucose at equilibrium, needed to convert the α-anomeric doublet into total glucose, was determined as 0.380 from the reference spectrum, in agreement with the accepted value,[211] and the fraction of the fructose intensity that falls into the window between 4.080 and 4.128 ppm (0.455 proton equivalents per molecule, the sum over the tautomers present at equilibrium[211,212]) was obtained in the same way. Three analytes are reported only when a second signal confirms the assignment: citric acid when the fitted coupling constant of its doublet lies between 15.3 and 16.5 Hz, sorbic acid when the olefinic doublet of H-2 at 5.83 ppm is present (Section 6.3), and 5-hydroxymethylfurfural when the aldehyde singlet at 9.445 ppm is accompanied by the ring proton at 7.53 ppm.

Table 9. Signals used for quantification, with analytical windows and the method of area measurement.

Compound and signal	δ (ppm) / window	Multiplicity	Protons	Area measurement	Reported from
Maleic acid, CH=CH (internal standard)	≈6.40; ±0.045 (integral), ±0.085 (fit)	s	2	integral and line fit	both tubes
Ethanol, CH_3	1.171; 1.245–1.100	t, 7.1 Hz	3	integral	native
Methanol, CH_3	3.344	s	3	line fit	native
Higher alcohols, CH_3 (isoamyl alcohols, isobutanol, 1-propanol)	0.926–0.840	d, d, t	6, 6, 3	line fit	native
Acetic acid, CH_3	2.083	s	3	line fit	native
Sorbic acid, CH_3 (confirmed by H-2, 5.83 ppm, d, 15.3 Hz)	1.84; 1.872–1.822	d	3	integral	native
Formic acid	8.224	s	1	line fit	native
Lactic acid, CH_3	1.41	d, 7.0 Hz	3	line fit	freeze-dried
Succinic acid, $(CH_2)_2$	2.671	s	4	line fit	freeze-dried
Malic acid, CH_2 (citrate doublet subtracted)	2.985–2.795	AB of ABX	2	integral	freeze-dried
Citric acid, CH_2 (downfield half)	3.041	d, 15.9 Hz	2 of 4	line fit	freeze-dried
Quinic acid, H-6ax	1.932	dd, 13.4 and 11.4 Hz	1	line fit	freeze-dried
Sucrose, H-1 (glucosyl)	5.402	d, 3.9 Hz	1	line fit	freeze-dried
Glucose, H-1 of the α anomer (total = α / 0.380)	5.218	d, 3.75 Hz	1	line fit	freeze-dried
Fructose	4.30–3.18 (library); 4.128–4.080 (window)	multiplets	0.455 in window	library decomposition; integral	freeze-dried

Compound and signal	δ (ppm) / window	Multiplicity	Protons	Area measurement	Reported from
Glycerol, sorbitol, 1,3-propanediol	4.30–3.18	multiplets	—	library decomposition	freeze-dried
1,3-Propanediol, central CH_2 (cross-check)	1.78; 1.830–1.735	quintet	2	integral	both tubes
2,3-Butanediol, CH_3	1.142	d, 6.3 Hz	6	line fit	freeze-dried
Doublet at 5.18 ppm (α-xylose or trehalose; tentative)	5.180	d	1	line fit	freeze-dried
Galacturonic acid, H-1 of the α anomer	5.322	d, 3.8 Hz	1	line fit	freeze-dried
Alanine, CH_3 (tentative)	1.577	d, 7.3 Hz	3	line fit	freeze-dried
5-Hydroxymethylfurfural, CHO (confirmed by H-3, 7.53 ppm)	9.445	s	1	line fit	freeze-dried
Hydroxycinnamoyl esters, H-7 (as chlorogenic acid)	7.685–7.555	d, 16 Hz	1	integral	freeze-dried
Aromatic and olefinic protons (fumaric acid and sorbic acid H-3 removed)	7.80–6.76	envelope	—	integral	freeze-dried

Shifts on the scale ethanol CH_3 = 1.1708 ppm. For the nine ciders without a freeze-dried tube the sugars and the principal acids were taken from the native tube with reporting limits raised 5.7-fold, and the polyols are not reported.

5.7 Calculation of concentrations and bridging of the two tubes

The area of each signal, divided by the number of protons that it represents, was compared with the area of the maleic acid singlet measured by the same procedure, and the concentration in the cider was calculated as

$$C\ (mg/L) = 0.99 \times m(MA) \times [(I / N) / (I(MA) / 2)] \times (M / 116.07) \times (1{,}000{,}000 / 175) \quad \text{(Eq. 1)}$$

in which I is the area of the signal, N the number of protons it represents, I(MA) the area of the maleic acid singlet, with its two protons, M the molar mass of the analyte (g/mol), 116.07 the molar mass of maleic acid, 0.99 the purity of the standard, m(MA) the mass of standard in the tube (9.37 mg in 100 µL of standard solution, 4.69 mg in 50 µL), and the factor 1,000,000 / 175 the conversion from the 175 µL of cider in the native tube to one liter. For the freeze-dried tube, in which 1000 µL of cider is represented, the last factor changes accordingly:

$$C\ (mg/L) = 0.99 \times m(MA) \times [(I / N) / (I(MA) / 2)] \times (M / 116.07) \times (1{,}000{,}000 / 1000) \quad \text{(Eq. 2)}$$

Volatile compounds (ethanol, methanol, the higher alcohols, acetic and formic acids, sorbic acid) were obtained from the native tube with Eq. 1, the non-volatile compounds from the freeze-dried tube with Eq. 2. So that both tubes of a cider stand on exactly the same scale, the freeze-dried spectrum was tied to the native spectrum of the same cider. Non-volatile analytes that are measured reliably in both tubes (succinic, malic, lactic and citric acids and the hexoses) give a set of ratios between the two tubes; their median, the bridging factor B, was used in place of the nominal ratio of the two preparations (1000 / 175 = 5.71 for equal amounts of standard). In this way the freeze-dried tube contributes its superior signal-to-noise ratio and its freedom from ethanol, while the scale of every result rests on the simple preparation of the native tube. The bridging factors are themselves a measure of the consistency of the two preparations and are reported in Section 7.2.

Sucrose is hydrolyzed measurably in the acidified sample between preparation and measurement, to a different extent in the two tubes. Where the native tube contained more sucrose than the freeze-dried tube, the native value was reported and the corresponding amounts of glucose and fructose were deducted, which

was the case for five ciders; the sum of the sugars is not affected by the hydrolysis, but all sucrose values are lower limits. Results below the reporting limits, which were set from the overlap and noise in each window and range from 5 mg/L for the higher alcohols to 300 mg/L for sorbitol, are given as not detected.

5.8 Derived quantities

Alcohol by volume was calculated from the ethanol concentration with the density of ethanol at 20 °C (0.7893 g/mL). Residual sugar is the sum of fructose, glucose and sucrose. Total organic acid is the sum of malic, lactic, succinic, citric, acetic and quinic acids; for use with the dryness scale it was also expressed as malic acid by summing the acid equivalents. This quantity counts the acids whether they are free or neutralized, and so it is an upper bound to the titratable acidity, not a measurement of it. The extent of malolactic conversion is described by the molar fraction of lactic acid in the sum of lactic and malic acids, and a cider is called lactate-dominant when this fraction is 0.5 or more. Energy was calculated on two bases described in Section 10, per liter and per serving of 12 fl oz (354.9 mL).

5.9 Chemometric methods

The multivariate analysis used five representations of each spectrum between 9.5 and 0.6 ppm: uniform bins of 0.04 and of 0.01 ppm, optimized buckets whose boundaries are moved to the nearest local minimum of the mean spectrum so that no peak is divided between two variables,[186,187,213] a table of targeted integrals, and the table of fitted concentrations. The regions of residual water, of the maleic acid standard and its fumaric acid impurity and, in the freeze-dried tubes, of residual ethanol were excluded. Intensities were expressed on the absolute scale provided by the internal standard (millimoles of protons per liter of cider and per ppm) and, for comparison, normalized to total area or by probabilistic quotient normalization.[190,214] Mean-centering was combined with no scaling, Pareto scaling, scaling to unit variance or a logarithmic transformation.[191,215] The choice among these options is itself a result and is reported in Section 11.

Principal component analysis,[216–218] hierarchical clustering by Ward's method,[194] t-distributed stochastic neighbor embedding[198] and uniform manifold approximation and projection[199] were used without reference to any label. Six classifiers were compared for supervised recognition: partial least squares discriminant analysis,[219,220] support vector machines with linear and radial kernels,[221] random forests,[222] k-nearest neighbors and linear discriminant analysis. PLS discriminant analysis and the random forest were applied to the scaled variables themselves, the other four methods to the scores of the first ten principal components of the training data. Composition was predicted from the bucketed spectra by PLS regression.[201] All models were validated by nested cross-validation in which the hyperparameters were selected in an inner loop and the performance was estimated in an outer loop,[203,204,223] and in which all ciders of one producer were always assigned to the same fold, so that no model was tested on a producer represented in its training set. Scaling parameters were calculated from the training folds only. Classification is reported as balanced accuracy and was tested against permuted labels; regression is reported as the root mean square error of cross-validation, Q^2 and the ratio of performance to deviation.[205] The calculations were carried out in Python with NumPy, SciPy,[224] scikit-learn[225] and umap-learn; the code accompanies this document.

6. Spectral Assignment

6.1 Pure-component spectra

The assignments rest on spectra of the pure compounds recorded under the same conditions and on published data for apple juice, wine and beer.[47,155,161,226–228] Figure 18 shows the carbohydrates and polyols. Sucrose is identified by the anomeric doublet of its glucosyl unit at 5.40 ppm and by the doublet of H-3 of the fructosyl unit at 4.21 ppm, both clear of other signals. Glucose at mutarotational equilibrium gives the anomeric doublets of the α form at 5.22 ppm and of the β form at 4.63 ppm; the latter is lost under the water resonance in a native cider, which is why total glucose is derived from the α doublet. The only resolved signal of the β form is the triplet-like H-2 at 3.24 ppm. Fructose, a mixture of the β-pyranose and β-furanose forms with minor contributions from the α forms, has no anomeric proton and is recognized by the multiplets near 4.10 and 4.00 ppm. Sorbitol and glycerol give no signal outside the crowded region between 3.5 and 3.9 ppm at 300 MHz, and only the high-field doublets of doublets of the glycerol methylene groups (3.56 and 3.65 ppm) stand out. 1,3-Propanediol contributes a triplet at 3.68 ppm to the same region and, more usefully, a quintet at 1.78 ppm in an otherwise empty part of the spectrum.

The alcohols and acids are collected in Figure 19. The spectra of the acids were recorded in neutral or weakly acidic solution, and their shifts in cider, where the pH lies between 3 and 4 and is lowered further by the standard, differ from those of the reference solutions; the triangles in the figure mark the positions observed in the ciders. Ethanol supplies the reference of the shift scale. Methanol, acetic acid and succinic acid are singlets, lactic acid is a doublet at 1.41 ppm with its quartet at 4.36 ppm, and malic acid forms an ABX system whose AB part between 2.80 and 2.95 ppm is used for quantification. The two halves of the AB quartet of citric acid lie at 3.04 and 2.89 ppm; the upfield half coincides with malic acid, so that citric acid is measured on its downfield half and its contribution is subtracted from the malic acid window. The methyl signals of the higher alcohols overlap between 0.87 and 0.90 ppm and are separated only approximately, and the methyl doublet of 2,3-butanediol at 1.14 ppm lies on the high-field flank of the ethanol triplet, from which it can be distinguished only in the freeze-dried tube.

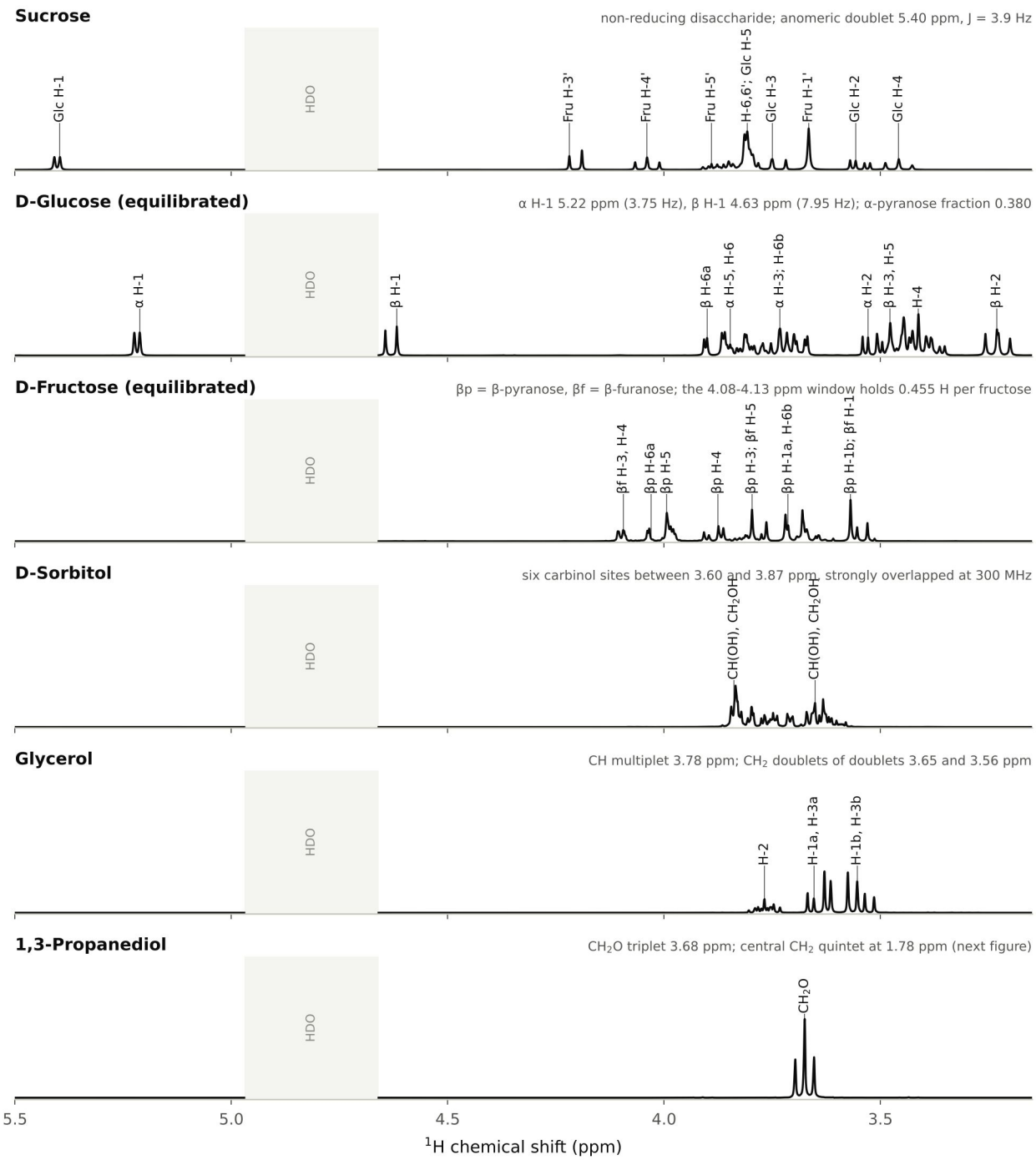

Figure 18. Pure-component ¹H NMR spectra (300 MHz, D_2O) of the carbohydrates and polyols found in cider, 5.5 to 3.15 ppm, each scaled to its own maximum, with assignments. The region of the residual water signal is shaded. βp and βf denote the β-pyranose and β-furanose forms of fructose.

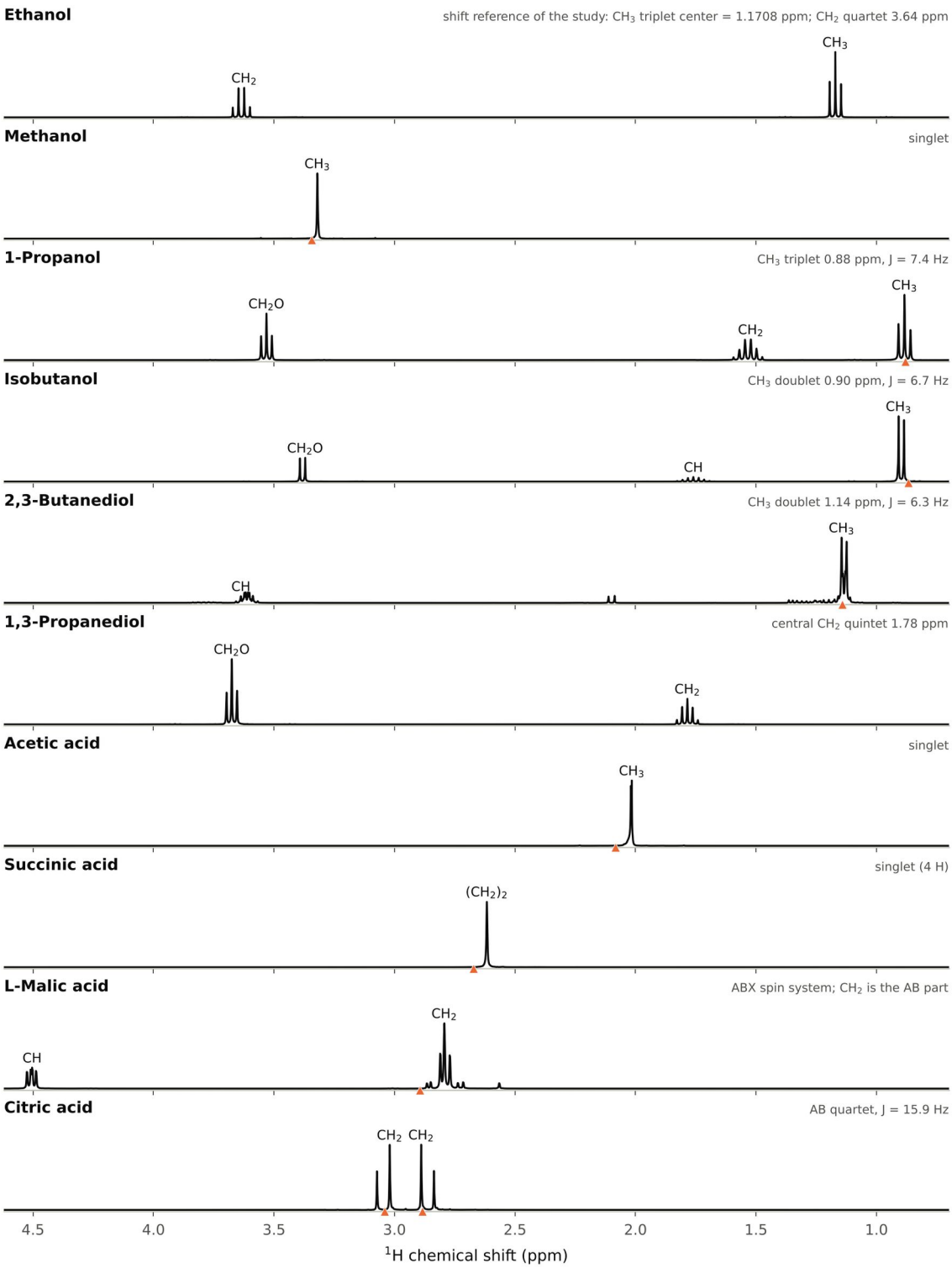

Figure 19. Pure-component ¹H NMR spectra (300 MHz, D_2O) of the alcohols and organic acids of cider, 4.6 to 0.7 ppm, each scaled to its own maximum. Triangles mark the position of the quantified signal in cider, where it differs from that of the reference solution because of pH.

6.2 Assigned spectra of ciders

The native spectrum of a typical craft cider, Doc's Draft Original Hard Apple Cider, is shown in Figure 20. Drawn on scale, it consists of four signals: the methyl and methylene multiplets of ethanol, the residual water resonance and the singlet of the internal standard. Everything else lies two orders of magnitude lower, and the expansions are drawn with the magnification stated in each panel so that the weak signals become visible while their size relative to the full spectrum remains explicit. At a magnification of about 20 the carbohydrate region and the anomeric signals appear; at 100 to 200 the organic acids, with malic acid dominant in this cider and lactic acid small, together with the methyl doublet of sorbic acid, the ^{13}C satellites of ethanol and the methyl signals of the higher alcohols; at 2000 to 3000 the olefinic protons of sorbic acid and the singlet of fumaric acid. The last of these is a trace companion of the maleic acid standard and was excluded from all aromatic integrals.

Figure 21 compares the freeze-dried tubes of two ciders of opposite character on one absolute scale: Stormalong Boston Heirloom, a sweet, malate-dominant craft cider, and Sarasola Sagardoa, a dry Basque cider that has undergone complete malolactic conversion. Removal of the ethanol and sixfold concentration bring the non-volatile constituents to the height of the standard. In the American cider, fructose, glucose, sucrose, sorbitol and glycerol fill the region between 3.2 and 4.3 ppm and malic acid is the only major acid. In the Basque cider the sugars are absent, lactic and acetic acids replace malic acid, and the triplet and quintet of 1,3-propanediol appear where the glycerol signals would be expected, the signature of bacterial glycerol degradation described in Section 3.5. The aromatic region, at a magnification of several hundred, shows the hydroxycinnamate doublet at 7.6 ppm and a broad envelope of polyphenol signals that is more intense in the traditional cider.

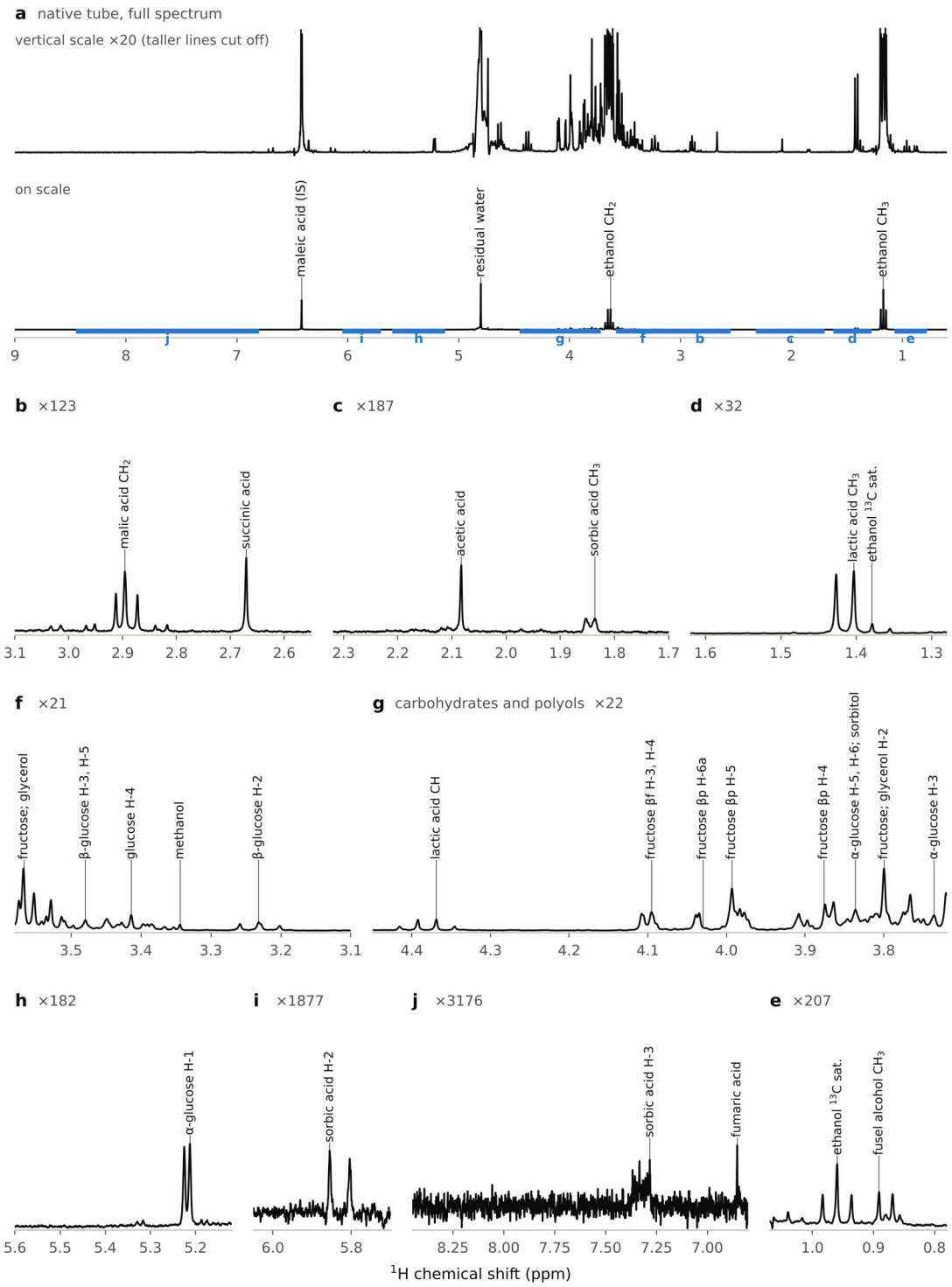

Figure 20. Assigned ¹H NMR spectrum of a cider in the native tube (Doc's Draft Original Hard Apple Cider; 175 µL of cider with maleic acid in D_2O, 300 MHz). (a) Full spectrum, drawn on scale (lower trace) and with the vertical scale expanded twentyfold and the tallest lines cut off (upper trace); the bars mark the regions expanded below. (b) to (j) Expansions, each with its vertical magnification relative to the trace on scale. IS, internal standard; sat., ¹³C satellite.

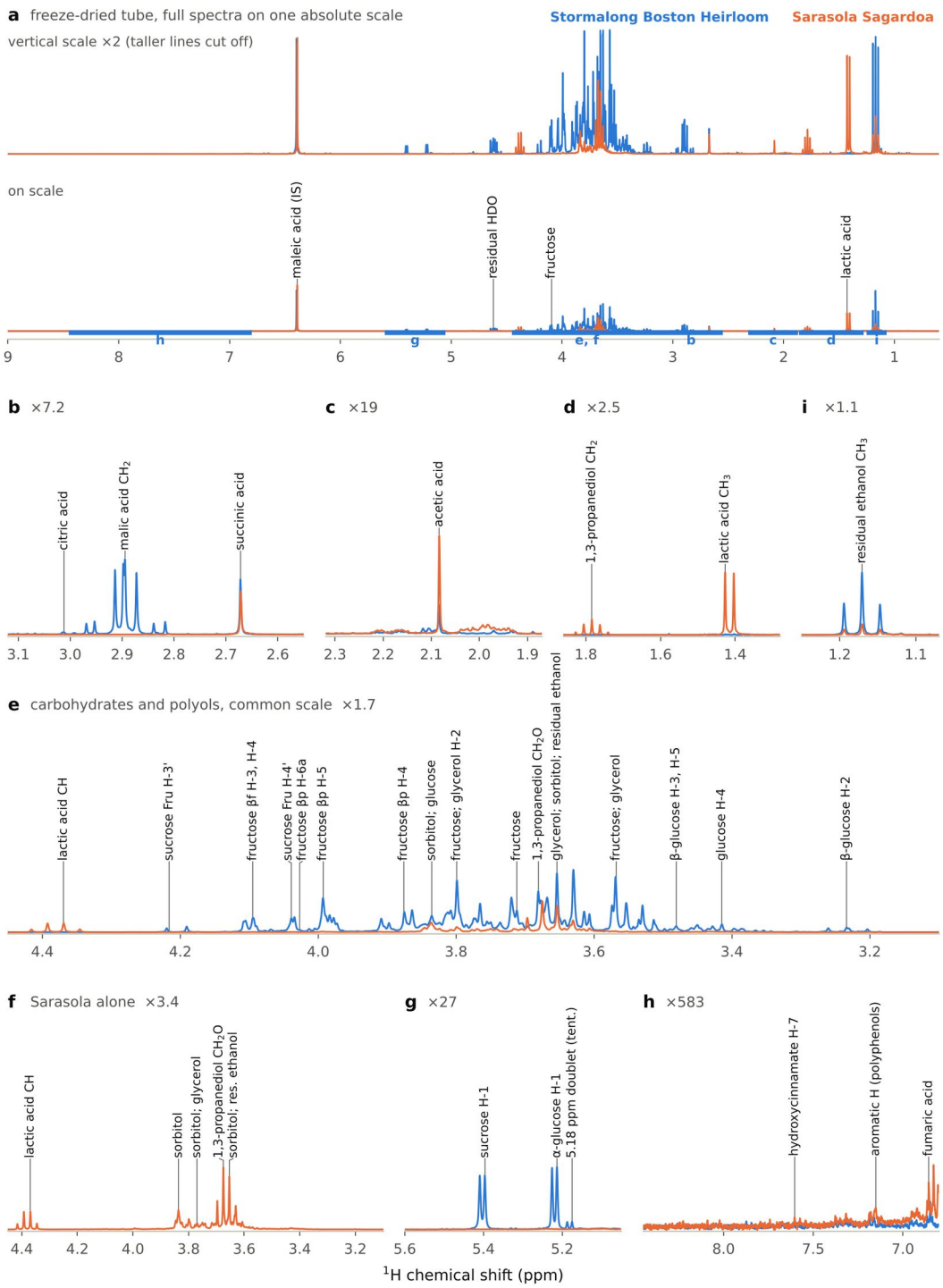

Figure 21. Assigned ¹H NMR spectra of two ciders in the freeze-dried tube (1000 µL of cider, lyophilized and redissolved in D_2O with maleic acid), drawn on one absolute scale: Stormalong Boston Heirloom (blue) and Sarasola Sagardoa (orange). (a) Full spectra, drawn on scale (lower traces) and with the vertical scale doubled and the tallest lines cut off (upper traces). (b) to (i) Expansions with their magnification relative to the traces on scale; panel f shows the Basque cider alone. tent., tentative assignment.

6.3 Additives and the confirmation of sorbic acid

Reference spectra of the additives that may be encountered in cider (potassium sorbate, sodium benzoate, ascorbic acid, propylene glycol) and of chlorogenic acid are given in Figure 22. Benzoate, ascorbic acid and propylene glycol were not detected in any cider of this study. Sorbic acid was found in 24 ciders, and its determination from a fixed integration window requires a safeguard. The methyl doublet at 1.84 ppm, which is used for quantification, lies 0.06 ppm from the center of the 1,3-propanediol quintet, and the outer line of that quintet falls inside the integration window. In ciders that contain 1,3-propanediol the window therefore registers an apparent sorbic acid content where there is none (panels g and h of the figure). Sorbic acid was accepted only when the doublet of H-2 at 5.83 ppm, with its coupling of 15.3 Hz, was found with an area consistent with the methyl signal. By this criterion 11 apparent findings were rejected, all of them in sour or traditional ciders that are, by their nature, unlikely to be preserved.

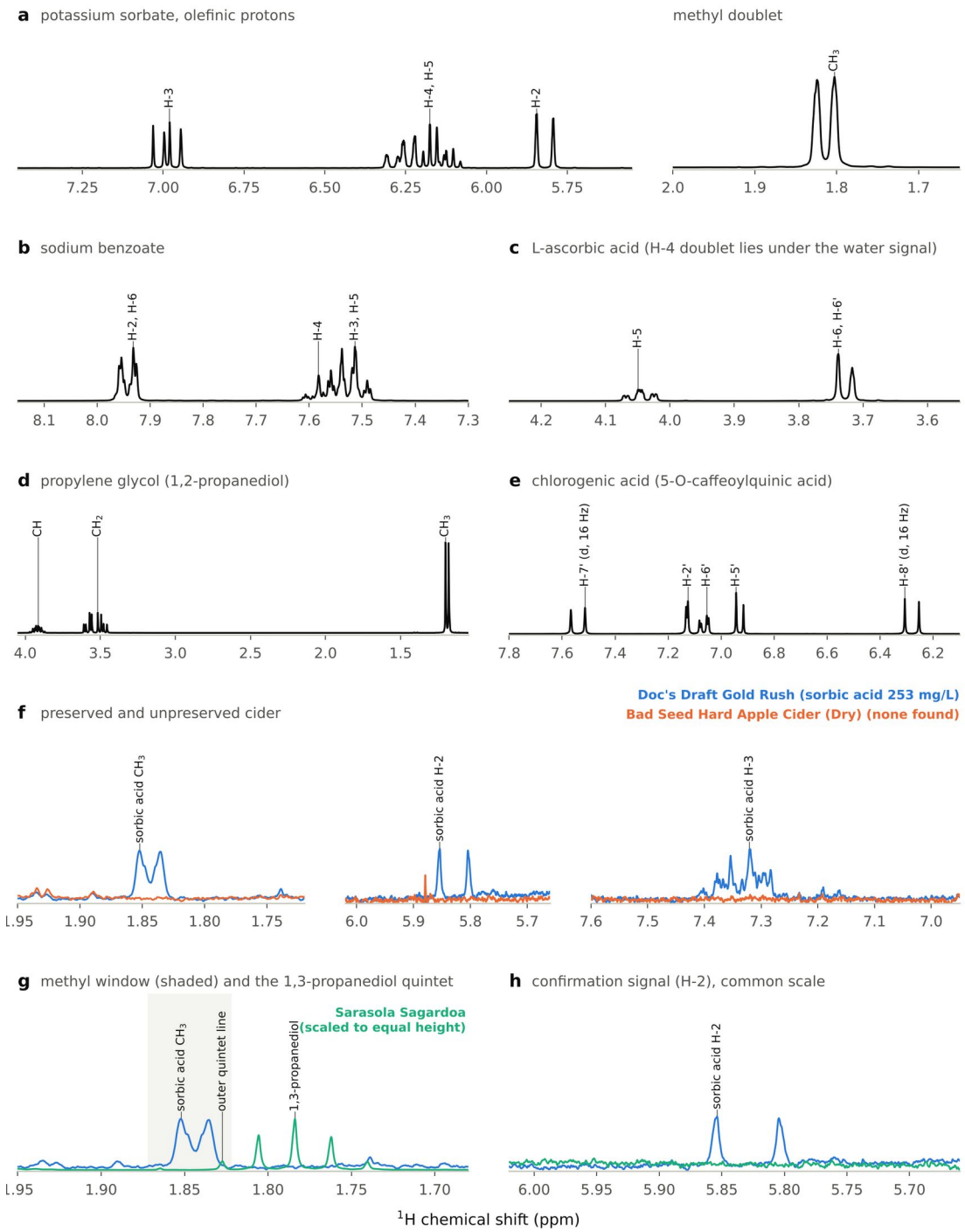

Figure 22. Additives and confirmation of sorbic acid. (a) to (e) Reference spectra of potassium sorbate, sodium benzoate, L-ascorbic acid, propylene glycol and chlorogenic acid (300 MHz, D_2O). (f) Native spectra of a preserved cider (Doc's Draft Gold Rush) and an unpreserved cider (Bad Seed Hard Apple Cider, Dry) in the three regions in which sorbic acid gives signals. (g) The methyl window of sorbic acid (shaded) and the 1,3-propanediol quintet of Sarasola Sagardoa, whose outer line falls into the window. (h) The confirmation signal of H-2 on a common scale: present in the preserved cider, absent in the Basque cider.

7. Performance of the Quantitative Measurement

7.1 Line fits and library decomposition

Representative fits are reproduced in Figure 23. For well-resolved signals (maleic acid, the ethanol triplet, the anomeric doublets) the residuals are at the level of the noise apart from the small dispersive features that result from imperfect phasing and from line shapes that are not exactly pseudo-Voigt. The fit of quinic acid illustrates the limit of the method: the doublet of doublets of H-6ax rides on a sloping background of unassigned signals, and although the coupling pattern identifies the compound, its area carries an uncertainty that is considerably larger than for the other acids. The decomposition of the carbohydrate region reproduces the experimental spectrum closely; the coefficient of determination has a median of 0.989 over all ciders and exceeds 0.986 in three quarters of them. The residual is largest between 3.6 and 3.9 ppm, where sorbitol, glycerol and the methylene quartet of residual ethanol coincide, and the values for glycerol and sorbitol, which are taken from this interval, carry a correspondingly larger uncertainty than those obtained from resolved signals.

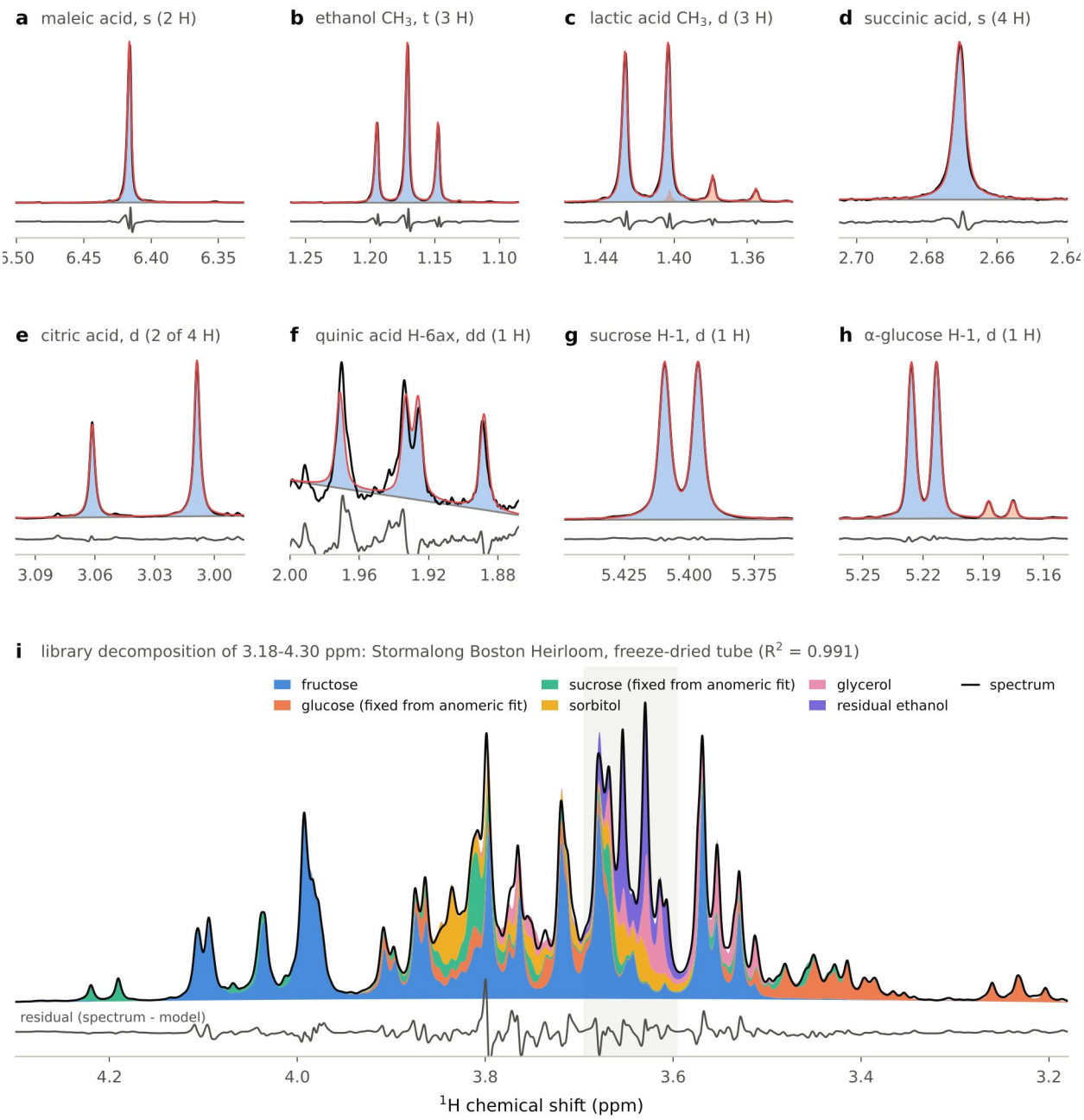

Figure 23. Measurement of peak areas. (a) to (h) Line fits of individual signals: spectrum (black), fitted model (red), components (shaded) and residual (below). (i) Decomposition of the carbohydrate region of Stormalong Boston Heirloom (freeze-dried tube) into the spectra of the pure compounds by non-negative least squares; glucose and sucrose were fixed at the values from their anomeric doublets. The shaded band marks the interval in which sorbitol, glycerol and residual ethanol overlap.

7.2 Agreement with declared alcohol content, repeatability and consistency of the two tubes

The only quantity for which an independent value exists for most ciders is the alcohol content declared on the label. For the 71 ciders with a declaration (Figure 24a) the mean difference between qNMR and label is −0.06 % v/v and the standard deviation of the differences is 0.71 % v/v; 48 ciders agree within 0.5 % v/v and 61 within 1.0 % v/v. The scatter is not an analytical property. Labels are approved before a batch is made and

are allowed a tolerance, and the largest differences concern craft products whose alcohol content varies from batch to batch.

Sixteen repeated acquisitions provide an estimate of the repeatability of the measurement on one tube (Figure 24c). The median absolute difference between repeat and original is below 1 % for ethanol, malic, lactic and succinic acids, 1 to 2 % for glycerol, sorbitol, fructose and glucose, 3.5 % for quinic acid and 5 to 13 % for the minor volatile compounds (isoamyl alcohols, acetic acid, isobutanol, methanol), whose signals are small and overlapped in the native tube. Over all pairs with concentrations above 200 mg/L the median difference is 1.4 % and the ninetieth percentile is 10 %.

Two products that were purchased and analyzed on two occasions show the variation that lies beyond repeatability, that between batches. For Citizen Cider The Dirty Mayor, ethanol, malic and succinic acids, glycerol and sorbitol agree within 5 %, while glucose and fructose differ by more than 20 %; the fermentation was the same, the sweetening was not. The two purchases of Citizen Cider Unified Press differ more widely, in ethanol (4.5 against 6.4 % v/v) as well as in malic acid, glycerol and sorbitol (16 to 20 %), and evidently came from different batches. The variation between lots of one product is examined further in Section 8.8.

The bridging factors measure the consistency of the two preparations of a cider (Figure 24b). For the 87 ciders with both tubes the ratio of the observed factor to its nominal value has a median of 1.026 and an interquartile range of 0.98 to 1.06, which means that the two tubes of a typical cider agree within a few percent. Six freeze-dried tubes give ratios between 0.47 and 0.60 and thus represent about 500 µL of cider; the bridging places their results on the same scale as all others. Within one cider the ratios obtained from the individual bridging analytes span 8 % in the median, and this spread is the principal contribution to the uncertainty of the non-volatile constituents.

Three further checks were made. First, because the suppression of the water signal could attenuate the neighboring anomeric doublet of α-glucose, the fraction of the α form was calculated in the cider spectra themselves from that doublet and from H-2 of the β form, which lies far from water. It was 0.37 to 0.39 in every campaign, the equilibrium value, so that the doublet is measured without loss. Second, the small singlet of fumaric acid, 0.45 ppm downfield of maleic acid, accompanies the standard at no more than 0.14 % of its area; it was excluded from the aromatic integrals and is not counted among the constituents of the ciders. Third, the amount of ethanol remaining after freeze-drying was determined (median 16 % of the original, with a range from almost nothing to 64 %), which confirms that volatile compounds cannot be quantified in the freeze-dried tube and that residual ethanol must be included in the library decomposition.

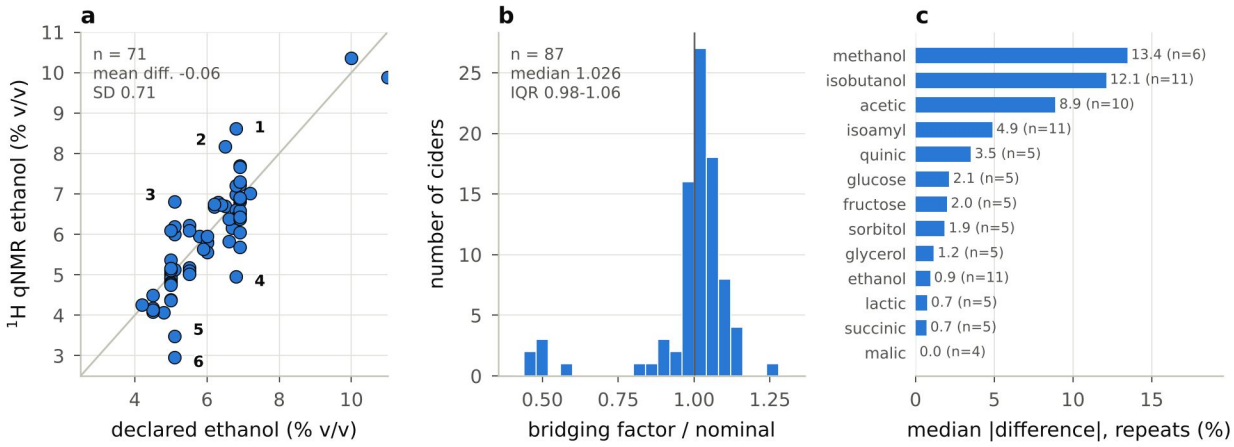

Figure 24. Performance of the quantitative measurement. (a) Ethanol by ¹H qNMR against the declared alcohol content; the line is the identity. Numbered points: 1, Hudson Valley Farmhouse Bourbon Barrel Aged; 2, Bad Seed Bourbon Barrel Reserve; 3, Downeast Original Blend, overpressurized can; 4, Naked Flock Lemon and Ginger; 5, Downeast Hard Honey Cider; 6, Downeast Cranberry Blend. (b) Distribution of the bridging factor between the freeze-dried and native tubes relative to its nominal value. (c) Median absolute difference between repeated acquisitions of the same tube, by analyte.

8. Composition of the Ciders

8.1 Overview

The composition of all 96 ciders is tabulated in Appendix B and displayed as a heat map in Figure 25; summary statistics are given in Table 10. Among the 93 alcoholic ciders, ethanol ranges from 3.0 to 10.4 % v/v with a median of 5.9 % v/v. Residual sugar extends from nil to 102 g/L (median 21.7 g/L), the sum of the organic acids from 1.9 to 13.9 g/L (median 5.4 g/L) and the sum of the polyols, where the freeze-dried tube was available, from 1.8 to 25.8 g/L (median 7.3 g/L). The picture that emerges from the heat map is one of two largely independent axes of variation: the amount and kind of residual sugar, which is a decision of the producer, and the state of the acids and of glycerol, which records the microbiology of the fermentation.

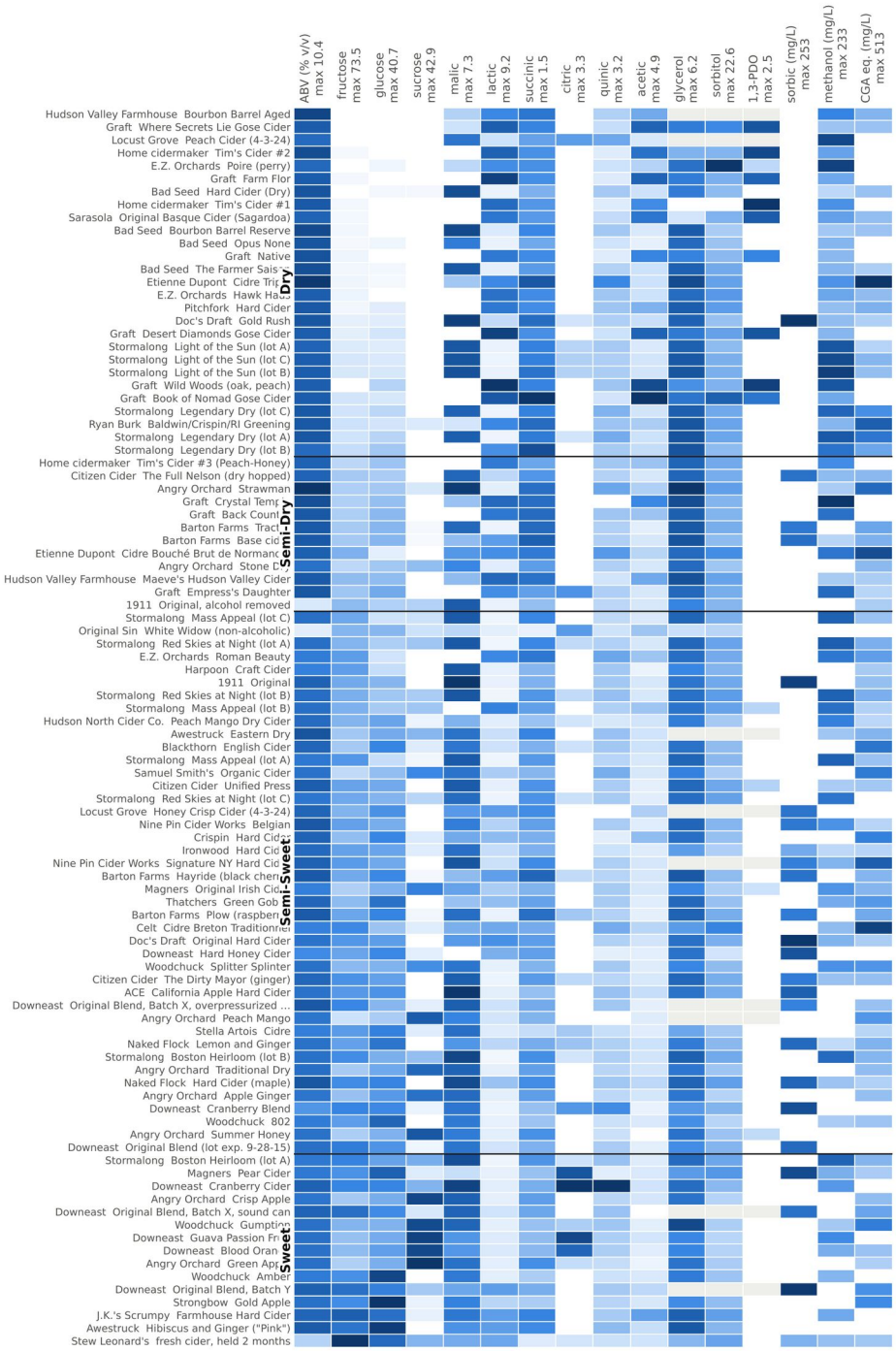

Figure 25. Composition of the 96 ciders by ¹H qNMR, ordered by residual sugar and grouped by dryness class (American Cider Association). Each column is scaled to its own maximum, which is given in the column heading (g/L unless stated otherwise); the color scale follows the square root of the scaled value so that small amounts remain visible. Gray cells: not determined (no freeze-dried tube). 1,3-PDO, 1,3-propanediol; CGA eq., hydroxycinnamoyl esters as chlorogenic acid.

Table 10. Concentrations determined by ¹H qNMR in 96 ciders (mg/L unless stated otherwise): number of ciders measured, number in which the compound was found above the reporting limit, and the minimum, median and maximum of the findings.

Compound	Measured	Found	Minimum	Median	Maximum	Cider with the maximum
Ethanol (g/L)	96	96	1.1	46.5	81.7	Angry Orchard Strawman
Methanol	96	68	12.8	52.0	233	Graft Book of Nomad - Crystal Temple (Cabernet Franc,…
Isoamyl alcohols	96	93	33.1	126	232	Angry Orchard Strawman
Isobutanol	96	93	6.3	27.4	79.4	Graft Desert Diamonds Gose Cider
1-Propanol	96	79	5.1	15.9	281	Graft Wild Woods - Grim Wilderness Vol. 1 (oak aged p…
Acetic acid	96	94	28.4	145	4863	Graft Book of Nomad Gose Cider
Formic acid	96	21	1.9	4.6	53.8	Graft Back Country (sour cider, citrus and bourbon)
Sorbic acid	96	24	44.1	125	253	Doc's Draft Gold Rush
Malic acid	96	79	213	3452	7335	1911 Original
Lactic acid	96	96	28.2	361	9162	Graft Wild Woods - Grim Wilderness Vol. 1 (oak aged p…
Succinic acid	96	96	12.9	424	1540	Graft Book of Nomad Gose Cider
Citric acid	96	28	101	231	3304	Downeast Cranberry Cider
Quinic acid	96	93	37.5	257	3168	Downeast Cranberry Cider
Fructose	96	90	127	10433	73478	Stew Leonard's Fresh apple cider, unpasteurized, held…
Glucose	96	86	64.5	6914	40692	Strongbow Gold Apple
Sucrose	96	48	56.8	2268	42931	Angry Orchard Green Apple
Glycerol	87	86	186	3143	6230	Angry Orchard Strawman
Sorbitol	87	86	743	3754	22580	E.Z. Orchards Poire (Forelle / Comice / Bosc perry)
1,3-Propanediol	87	14	103	1304	2492	Home cidermaker (Tim) Tim's Cider #1
2,3-Butanediol	87	84	25.8	94.8	652	Stella Artois Cidre
Doublet at 5.18 ppm, as α-xylose (tentative)	96	73	36.6	133	380	E.Z. Orchards Poire (Forelle / Comice / Bosc perry)
Galacturonic acid, α anomer	96	64	30.5	105	545	Woodchuck 802
Alanine (tentative)	96	44	5.5	25.5	79.4	Locust Grove Peach Cider (4-3-24)
5-Hydroxymethylfurfural	96	10	1.4	2.9	6.2	Barton Farms Hayride (black cherry)
Hydroxycinnamoyl esters, as chlorogenic acid	96	70	31.4	75.7	513	Etienne Dupont Cidre Triple

Polyols and 2,3-butanediol require the freeze-dried tube (87 ciders). Minimum and median refer to the ciders in which the compound was found.

8.2 Ethanol, methanol and higher alcohols

Most of the alcoholic ciders contain between 4.5 and 7 % v/v of ethanol. The exceptions at the upper end are two products that are sold as strong ciders, Angry Orchard Strawman (10.4 % v/v) and Etienne Dupont Cidre Triple (9.9 % v/v), and a group of barrel-aged craft ciders. A point of regulatory interest is the boundary at 7 % v/v that separates the jurisdictions of the two labeling agencies (Section 3.7). Only 3 ciders

declare 7 % v/v or more, whereas 12 were found at or above that value. Declarations of 6.8 or 6.9 % v/v are conspicuously frequent, 20 ciders in all. The alcohol content measured in this group ranges from 5.0 to 8.6 % v/v with a mean of 6.8 % v/v, so the declarations are accurate on average; but 6 of the 20 ciders were found at or above 7 % v/v. Given the uncertainty of a single measurement and the tolerance allowed to a label, no individual product can be said to be mislabeled on this evidence. What the data show is that the ordinary variation between batches carries a craft cider declared just below the boundary across it in a substantial fraction of cases.

Methanol was found in 68 ciders at 13 to 233 mg/L. The highest values belong to ciders with added fruit or grape skins, to the perry and to the Stormalong ciders, and methanol was below the reporting limit in twelve of the nineteen mass-market products, which is consistent with its origin in the pectin of the fruit and with the lower juice content or the clarification of industrial ciders.[84] The methyl signals of the higher alcohols correspond to 126 mg/L of isoamyl alcohols (median), with much smaller amounts of isobutanol and 1-propanol. The division of this overlapped multiplet among the three alcohols is approximate, and their sum is the more reliable quantity. The signal assigned to 1-propanol is strongly elevated in the sour ciders (up to 281 mg/L in Graft Wild Woods), which is consistent with a bacterial origin.

8.3 Residual sugars

Residual sugar is the property in which the ciders differ most (Figure 26). Sixteen ciders contain less than 1 g/L and are dry in the strict sense; at the other extreme fifteen exceed 45 g/L. The composition of the sugar is as informative as its amount. Sucrose at 2 g/L or more was found in 24 ciders and reaches 43 g/L in Angry Orchard Green Apple. It is the principal sugar in 13 ciders: ten of the nineteen mass-market products, two fruit-flavored ciders of Downeast and Samuel Smith's Organic Cider. Because apples themselves contain sucrose and because sucrose is inverted on storage, neither its presence nor its absence is proof of the kind of sweetener; but yeast hydrolyzes and consumes sucrose early in the fermentation, and a cider in which sucrose exceeds the sum of glucose and fructose, as it does in 11 of these products, is best explained by the addition of refined sugar or a sucrose syrup after fermentation. Ciders sweetened with juice or by arrested fermentation show the opposite pattern, fructose in excess of glucose and little sucrose. The ratio of fructose to glucose has a median of 1.5 and exceeds 4 in 7 ciders, among them Etienne Dupont Cidre Bouché Brut (25), E.Z. Orchards Roman Beauty, two lots of Stormalong Mass Appeal and Harpoon Craft Cider. Such ratios are the mark of a fermentation that stopped, or was stopped, after the yeast had consumed nearly all of the glucose.

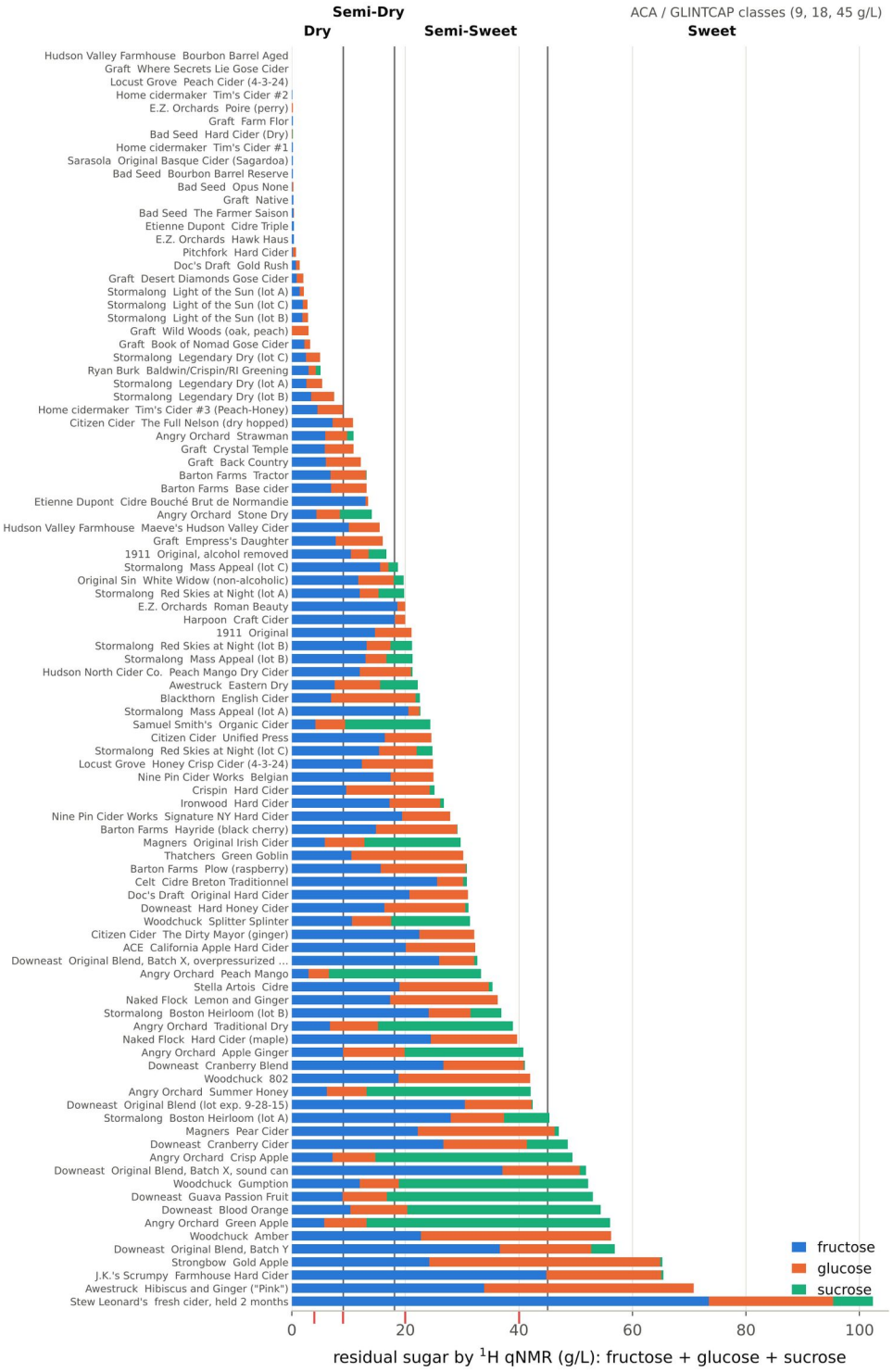

Figure 26. Residual sugar of the 96 ciders by ¹H qNMR (fructose, glucose and sucrose, g/L), in ascending order. The vertical lines and the class names at the top give the classes of the American Cider Association (boundaries at 9, 18 and 45 g/L); the red ticks at the bottom mark the boundaries of the BJCP guidelines (4, 9, 20 and 40 g/L).

8.4 Organic acids and malolactic conversion

Malic and lactic acids are complementary (Figure 27a). In 39 ciders less than a tenth of the sum of the two is lactic acid, in 20 more than nine tenths, and malic acid was not detected at all in 17. By the criterion of Section 5.8, 36 of the 96 ciders are lactate-dominant. The distribution over the market segments is uneven: one of the nineteen mass-market ciders, five of the seven traditional European ciders, all four home-made ciders and 25 of the 63 regional craft ciders. Malolactic conversion is thus avoided almost without exception by the large producers and occurs, by intention or not, in two fifths of the craft products. The amounts are considerable. Lactic acid exceeds 1 g/L in 37 ciders and reaches 9.2 g/L in Graft Wild Woods, far more than the conversion of the malic acid of an apple juice can supply, so that in the sour ciders lactic acid must also have been formed from sugar.

Acetic acid follows lactic acid (Figure 27b). Its median is 145 mg/L, but 12 ciders contain 1 g/L or more, up to 4.9 g/L in Graft Book of Nomad Gose. All twelve are lactate-dominant ciders with more than 2 g/L of lactic acid: seven ciders of Graft, the Basque cider, the two home-made ciders and two farmhouse-style craft ciders. At such concentrations acetic acid is a defining element of the style, and the products concerned are presented as sour. Succinic acid, a product of the yeast, increases with the alcohol content (Figure 27c; correlation coefficient 0.57) and lies between 0.2 and 0.8 g/L in nearly all ciders. Citric acid was confirmed in 28 ciders. Above 0.8 g/L it occurs exclusively in products with declared fruit other than apple (cranberry, blood orange, guava and passion fruit, peach, strawberry, blackberry) and in Magners Pear Cider, and it may derive from the fruit or from acidification. Quinic acid, at a median of 257 mg/L, is a normal constituent of the apple; the two cranberry ciders stand apart with 3.2 and 1.0 g/L (Figure 27d), as expected from the composition of cranberry juice,[229] and the three traditional French ciders contain 0.6 to 1.1 g/L.

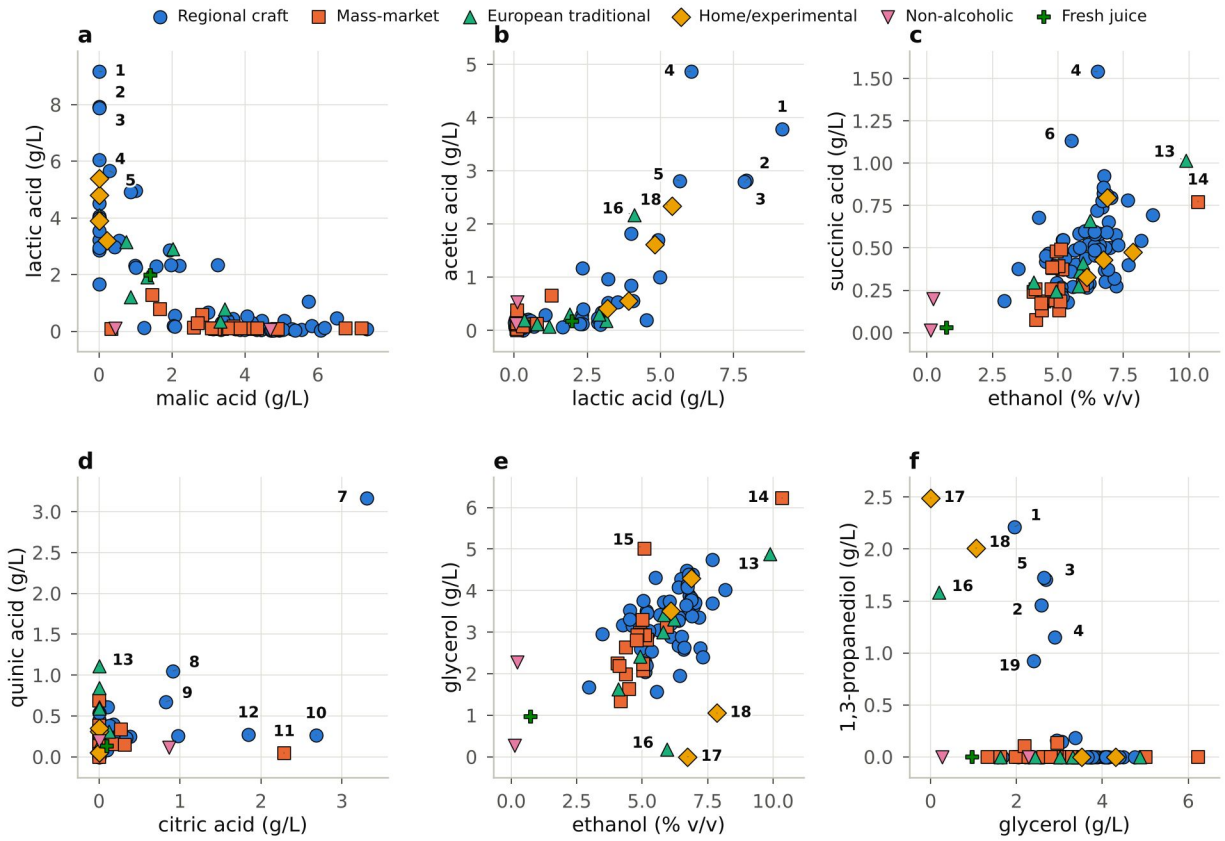

Figure 27. Relations among acids and fermentation products in the 96 ciders, by market segment. (a) Lactic against malic acid. (b) Acetic against lactic acid. (c) Succinic acid against ethanol. (d) Quinic against citric acid. (e) Glycerol against ethanol. (f) 1,3-Propanediol against glycerol. Numbered points: 1, Graft Wild Woods; 2, Graft Farm Flor; 3, Graft Desert Diamonds Gose; 4, Graft Book of Nomad Gose; 5, Graft Where Secrets Lie Gose; 6, Stormalong Legendary Dry (lot B); 7, Downeast Cranberry Cider; 8, Downeast Cranberry Blend; 9, Locust Grove Peach Cider; 10, Downeast Guava Passion Fruit; 11, Magners Pear Cider; 12, Downeast Blood Orange; 13, Etienne Dupont Cidre Triple; 14, Angry Orchard Strawman; 15, Woodchuck Gumption; 16, Sarasola Sagardoa; 17, Tim's Cider #1; 18, Tim's Cider #2; 19, Graft Native.

8.5 Polyols, 1,3-propanediol and the signature of natural cider

Glycerol amounts to 3.1 g/L in the median cider, and its mass ratio to ethanol has a median of 0.069 with an interquartile range of 0.060 to 0.078, the range expected for the alcoholic fermentation (Figure 27e).[80] The ciders that fall far below this relation are the instructive ones. In nine ciders 1,3-propanediol was found at 0.9 to 2.5 g/L (Figure 27f): Sarasola Sagardoa, the two home-made ciders of one cider maker, and six ciders of Graft. All nine are lactate-dominant and all contain more than 1.5 g/L of acetic acid. The combination of exhausted malic acid, high lactic and acetic acids and conversion of glycerol to 1,3-propanediol is the chemical signature of the heterofermentative lactic acid bacteria of natural cider (Section 3.5),[98–100] and it is found here not only in the Basque cider but also in American ciders made by spontaneous or mixed fermentation. The determination of 1,3-propanediol rests on two independent signals, the quintet at 1.78 ppm and the triplet in the library decomposition, which agree.

Sarasola Sagardoa, the one cider of the collection that was made in the Basque Country by the traditional method, shows the style in its purest form (Figure 28). It contains 5.9 % v/v of ethanol and 0.2 g/L of residual sugar. Malic acid is absent, lactic acid amounts to 4.1 g/L, and acetic acid, at 2.2 g/L, corresponds to the upper end of the volatile acidity that the standard admits for natural cider. Of the glycerol, 0.2 g/L remains, against some 3.4 g/L that a fermentation to this strength may be expected to produce, and 1,3-

propanediol is present at 1.6 g/L. In the spectrum the quintet of the central methylene group of 1,3-propanediol stands alone at 1.78 ppm, and between 3.5 and 3.7 ppm its triplet takes the place of the doublets of doublets of glycerol, which are prominent in a cider of equal dryness that was fermented without the participation of these bacteria (Figure 28a and b). The same pattern is found, in various degrees, in the eight American ciders. In Tim's Cider #1 no glycerol is left at all; in the ciders of Graft glycerol and 1,3-propanediol are present side by side.

A balance in moles shows how far the conversion has gone (Figure 28c). In the 70 alcoholic ciders that are free of 1,3-propanediol the molar ratio of glycerol to ethanol has a median of 0.036, with an interquartile range of 0.031 to 0.040, and this ratio, applied to the ethanol content, gives the amount of glycerol that the yeast may be expected to have formed. In Sarasola Sagardoa the glycerol and the 1,3-propanediol that were found account together for 62 % of that amount, in the two home-made ciders for 78 %, and in Graft Native for 84 %; the remainder has gone to other products of the pathway, among which 3-hydroxypropionic acid and the adducts of acrolein with polyphenols are known (Section 3.5). In the other five ciders of Graft the sum exceeds the expectation by 15 to 30 %, which suggests that these mixed fermentations formed more glycerol than the typical fermentation and that only a part of it was converted. Acetic acid rises with 1,3-propanediol across the whole collection (correlation coefficient 0.83; Figure 28d), as the coupling of the two reactions in the metabolism of the bacteria requires. The limits that regulations place on volatile acidity, 1.4 g/L for fruit wine in the United States[230] and 2.2 g/L for cider in Spain,[93] are drawn in the figure for orientation; the acetic acid content determined by NMR is not identical with the volatile acidity of the official methods, which is obtained by distillation. For the producer the practical value of the measurement lies in its timing. 1,3-Propanediol is detectable at about 0.1 g/L, long before the glycerol is exhausted or bitterness develops, so that a spectrum of the maturing cider reveals the activity of glycerol-degrading bacteria at a stage at which racking, sulfiting or bottling can still be decided upon.

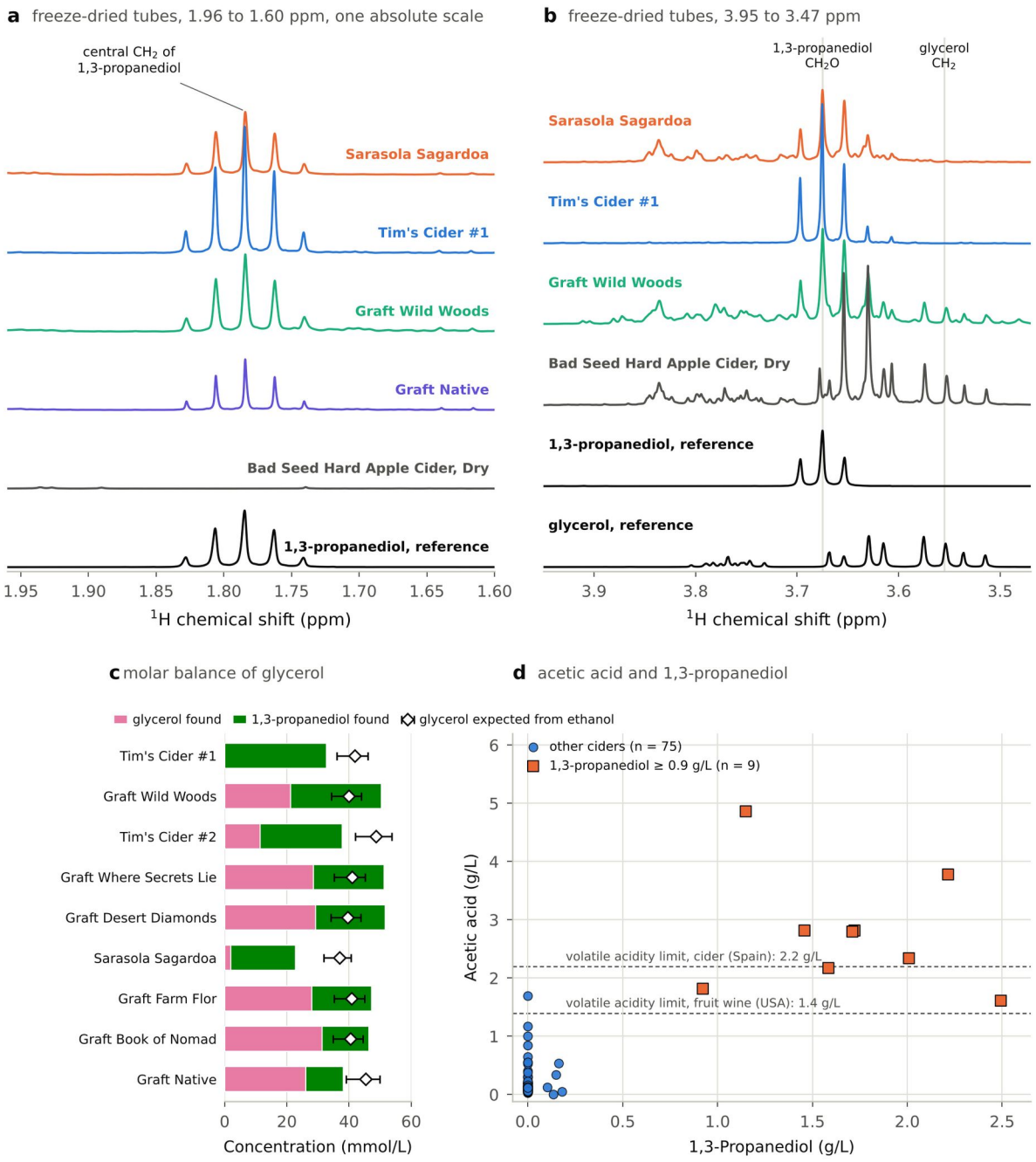

Figure 28. Natural cider and the bacterial conversion of glycerol. (a) The quintet of 1,3-propanediol at 1.78 ppm in the freeze-dried tubes of Sarasola Sagardoa, Tim's Cider #1, Graft Wild Woods and Graft Native, compared with a dry cider that contains none (Bad Seed Hard Apple Cider, Dry) and with the reference compound; one absolute scale for the ciders. (b) The region of the glycerol and 1,3-propanediol methylene signals with the two reference compounds; the strong lines near 3.64 ppm in the trace of the Bad Seed cider belong to residual ethanol. (c) Molar balance of glycerol in the nine ciders that contain 1,3-propanediol: glycerol and 1,3-propanediol found, and the glycerol expected from the ethanol content (diamond, with the interquartile range of the ratio in ciders free of 1,3-propanediol). (d) Acetic acid against 1,3-propanediol in 84 ciders; the broken lines mark regulatory limits for volatile acidity.

Sorbitol has a median of 3.7 g/L. It is by far highest in the perry of E.Z. Orchards (22.6 g/L), as the composition of pears demands,[40] and elevated in Magners Pear Cider, in the traditional French ciders and in

two of the Graft gose ciders, 13.6 g/L in Book of Nomad Gose. A sorbitol content well above 6 g/L in an apple cider is consistent with the presence of pear or of another sorbitol-rich fruit among the ingredients. Because sorbitol is not fermented, it also contributes to the difference between a cider that is dry by its sugar content and one that tastes dry.

8.6 Preservatives, minor markers and polyphenols

Sorbic acid was confirmed in 24 ciders at 44 to 253 mg/L (median 125 mg/L). Twenty-one of them are regional craft ciders, two are mass-market products and one is the fresh juice; twenty of the 24 are semi-sweet or sweet. This is the expected pattern. A large producer stabilizes a sweetened cider by pasteurization or sterile filtration, a small one by adding sorbate. 5-Hydroxymethylfurfural, a product of the thermal degradation of fructose, was detected in 10 ciders at 1.4 to 6.2 mg/L, including all four ciders of Barton Farms and four mass-market products; at these levels it indicates the use of concentrate, of a cooked sweetener such as caramel, or pasteurization. Formic acid was found in 21 ciders, mostly between 2 and 10 mg/L and at 54 mg/L in Graft Back Country.[145]

Polyphenols cannot be determined individually at 300 MHz in a matrix of this kind, but two integral measures are available. The doublet of the olefinic proton of the hydroxycinnamoyl esters, expressed as chlorogenic acid, corresponds to a median of 76 mg/L where it was detected, and the integral of the whole aromatic region, after removal of fumaric and sorbic acids, to a median of 3.4 mmol of protons per liter. Both measures single out the traditional European ciders: their mean hydroxycinnamate content is 287 mg/L against 57 mg/L in the regional craft ciders and 91 mg/L in the mass-market ciders, and their aromatic intensity is twice as high. The ciders of Etienne Dupont and the Breton cider of Celt lead the list, followed by a few American ciders made from fruit of high tannin content. This is the spectroscopic expression of the bittersweet apple, and it is the property in which most American ciders, made from dessert and heirloom fruit, differ from the European tradition.

8.7 Market segments

The medians of the principal constituents in the four larger segments are compared in Table 11. The mass-market ciders form the most uniform group: about 5 % v/v of ethanol, close to 40 g/L of sugar of which a substantial part is sucrose, malic acid intact, lactic and acetic acids low. The regional craft ciders contain more alcohol and about half as much sugar in the median, but the median conceals a range that reaches from completely dry to the sweetest products of the study. The traditional European ciders are distinguished by lactic acid and polyphenols, the home-made ciders by complete dryness and by the products of bacterial metabolism. The comparison describes the products that were collected; the segments are represented unequally, and the assignment of a producer to a segment is a matter of judgment.

Table 11. Median composition of the ciders by market segment.

Quantity	Regional craft	Mass-market	European traditional	Home/experimental
Number of ciders	63	19	7	4
Ethanol (% v/v)	6.4	4.9	5.8	6.8
Residual sugar (g/L)	19.9	39.0	22.6	2.6
Sucrose (g/L)	0.0	5.6	0.0	0.0
Malic acid (g/L)	3.34	3.37	1.32	0.00
Lactic acid (g/L)	0.40	0.10	1.90	4.36
Acetic acid (g/L)	0.14	0.10	0.20	1.08
Total organic acids (g/L)	5.6	4.1	5.2	6.0
Glycerol (g/L)	3.38	2.81	3.01	2.28
Sorbitol (g/L)	4.14	2.48	3.83	4.16
Hydroxycinnamoyl esters, as chlorogenic acid (mg/L)	59	68	235	0
Aromatic protons (mmol/L)	3.2	3.6	5.7	4.0
Energy (kcal per 12 fl oz)	161	165	159	154
Lactate-dominant ciders	25 of 63	1 of 19	5 of 7	4 of 4
Sucrose ≥ 2 g/L	11 of 63	10 of 19	1 of 7	0 of 4
Sorbic acid found	21 of 63	2 of 19	0 of 7	0 of 4

The non-alcoholic ciders (2) and the fresh juice (1) are omitted. Polyols from the ciders with a freeze-dried tube.

8.8 Variation between lots

The fourteen cans of Stormalong, two or three lots each of five products, all measured in 2020, show how reproducible a craft cider is from lot to lot (Figure 29, Table 12). Ethanol, residual sugar, glycerol and sorbitol of a given product agree within the range that might be expected for a blended, back-sweetened product: Light of the Sun, the driest, contains 2 to 3 g/L of sugar in all three lots, Boston Heirloom 37 and 45 g/L. The exception is of the kind that matters. One lot of Legendary Dry (lot B) and one lot of Mass Appeal (lot B) contain no malic acid and 2.9 and 3.5 g/L of lactic acid, whereas the other lots of the same products contain 4.3 to 4.7 g/L of malic acid and virtually no lactic acid. These two lots underwent a complete malolactic fermentation, in tank or in the can, which the others did not. The acidity perceived by the consumer and the dryness index of Section 9 differ accordingly between cans that carry the same label. The products can also be told apart by minor constituents that reflect their recipes: citric acid, at 0.14 to 0.26 g/L, is found in every lot of Red Skies at Night and Light of the Sun, which are made with passion fruit and hibiscus and with citrus zest and hops respectively,[231] whereas it is at or below the reporting limit of 0.1 g/L in the other three products.

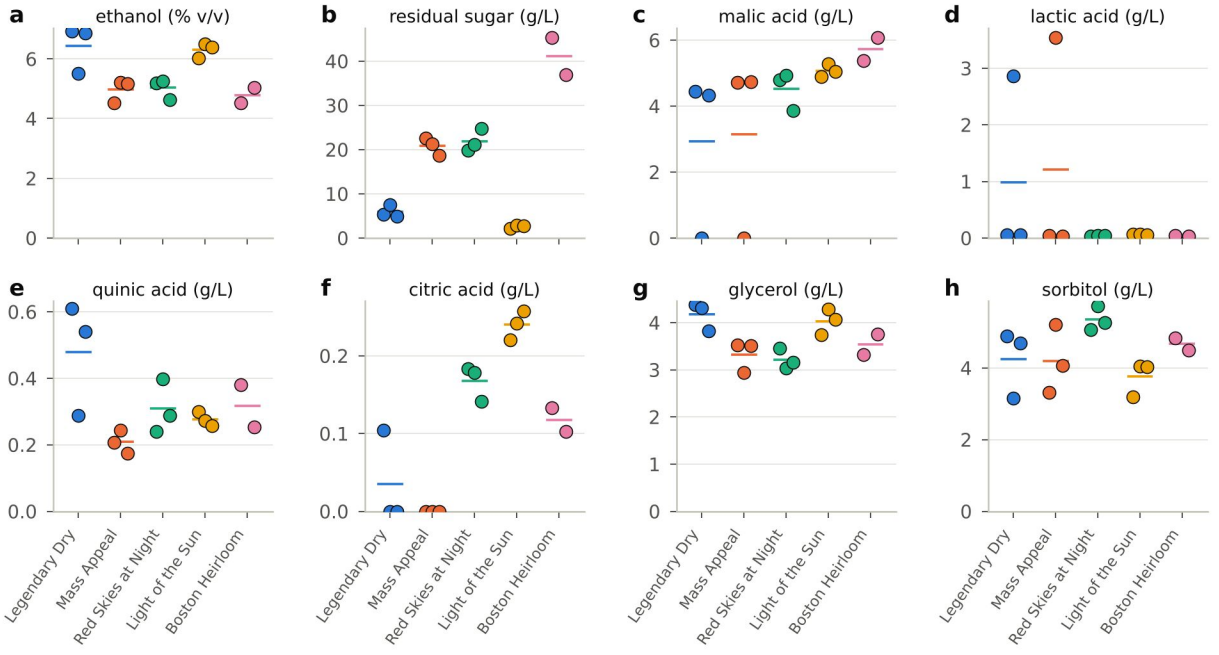

Figure 29. Lot-to-lot variation of five Stormalong ciders (fourteen cans from different lots). Each point is one lot; the bar is the mean of the product. (a) Ethanol. (b) Residual sugar. (c) Malic acid. (d) Lactic acid. (e) Quinic acid. (f) Citric acid. (g) Glycerol. (h) Sorbitol.

Table 12. Composition of fourteen lots of five Stormalong ciders (g/L unless stated otherwise).

Product	Lot	Ethanol (% v/v)	Sugar	Sucrose	Fructose	Malic	Lactic	Succinic	Citric	Glycerol	Sorbitol	Methanol (mg/L)	kcal / 12 fl oz
Legendary Dry	A	6.9	5.3	nd	2.56	4.45	0.05	0.65	0.10	4.38	4.89	152	162
Legendary Dry	B	5.5	7.4	nd	3.45	nd	2.86	1.13	nd	4.32	3.16	110	134
Legendary Dry	C	6.9	5.0	nd	2.45	4.32	0.05	0.59	nd	3.83	4.69	131	159
Mass Appeal	A	4.5	22.6	0.13	20.5	4.71	0.04	0.42	nd	3.53	3.32	137	136
Mass Appeal	B	5.2	21.2	4.56	13.0	nd	3.54	0.37	nd	2.94	5.21	116	148
Mass Appeal	C	5.2	18.7	1.69	15.5	4.74	0.04	0.55	nd	3.50	4.07	137	144
Red Skies at Night	A	5.2	19.8	4.59	11.9	4.80	0.03	0.55	0.18	3.46	5.06	142	147
Red Skies at Night	B	5.3	21.1	3.79	13.1	4.93	0.04	0.44	0.18	3.03	5.73	136	151
Red Skies at Night	C	4.6	24.8	2.77	15.4	3.86	0.05	0.47	0.14	3.16	5.26	108	142
Light of the Sun	A	6.0	2.1	nd	1.41	4.89	0.06	0.60	0.22	3.74	3.19	158	138
Light of the Sun	B	6.5	2.8	nd	1.85	5.29	0.07	0.72	0.24	4.28	4.05	207	150
Light of the Sun	C	6.4	2.8	nd	1.89	5.05	0.05	0.60	0.26	4.07	4.03	180	147
Boston Heirloom	A	4.5	45.4	8.04	28.0	5.38	0.04	0.45	0.13	3.31	4.84	132	171
Boston Heirloom	B	5.0	36.9	5.43	24.2	6.07	0.03	0.49	0.10	3.75	4.49	124	170

nd, not detected.

A second observation concerns two cans of the same batch of Downeast Original Blend, one of which had become overpressurized. The sound can contained 6.0 % v/v of ethanol and 51.8 g/L of sugar, the swollen can 6.8 % v/v and 32.7 g/L. Glucose had fallen by 55 % and fructose by 30 %, the order in which yeast

consumes the two sugars, and the loss of 19 g/L of sugar is accompanied by a gain of 6.4 g/L of ethanol. The apparent yield of 0.33 g of ethanol per gram of sugar is below the 0.46 to 0.51 g/g of a fermentation, which is not surprising for two cans that need not have been identical when they were filled and one of which lost gas and vapor on opening. The can had refermented. The product is unfiltered and was, in the samples of 2015, preserved with sorbic acid at about 0.1 g/L, which evidently did not suffice in this can.

9. Dryness and Further Scales Derived from the Spectrum

9.1 The sugar-based scales of North America

The residual sugar determined by qNMR places the ciders on the two sugar-based scales directly (Table 2, Figure 26). By the classes of the American Cider Association, 27 ciders are dry, 12 semi-dry, 42 semi-sweet and 15 sweet. The BJCP scale, with its lower boundaries, gives 23 dry, 4 semi-dry, 16 medium, 33 semi-sweet and 20 sweet. The two scales agree on the ordering and differ where they are meant to differ: four ciders that are dry for the association (4 to 9 g/L) are semi-dry for the judge, four between 18 and 20 g/L are medium instead of semi-sweet, and five between 40 and 45 g/L change from semi-sweet to sweet (Table 13). The distribution over the segments shows that no mass-market cider of this collection is dry by either scale, while 22 of the 63 regional craft ciders are.

The word dry on the label is a poor guide. Of the products whose name contains it, the three lots of Stormalong Legendary Dry (5 to 7 g/L) and Bad Seed Dry (0.2 g/L) are dry on the association's scale. Angry Orchard Stone Dry (14 g/L) and Citizen Cider The Full Nelson, which is dry-hopped and makes no claim about sugar, are semi-dry. Awestruck Eastern Dry (22 g/L), Hudson North Peach Mango Dry Cider (21 g/L) and Angry Orchard Traditional Dry (39 g/L) are semi-sweet, the last of them close to the boundary of sweet. The French designation brut, on Etienne Dupont Cidre Bouché Brut (13 g/L), corresponds to semi-dry.

Table 13. Classification of the 96 ciders on the three North American dryness scales: classes of the American Cider Association (rows) against the BJCP classes, the unadjusted ratio classes of the orchard-based scale, and the market segments.

ACA class	n	BJCP: Dry	Semi-dry	Medium	Semi-sweet	Sweet	Ratio: Dry	Semi-Dry	Semi-Sweet	Sweet	Craft	Mass-market	European	Home	Non-alc.	Juice
Dry	27	23	4	0	0	0	25	2	0	0	22	0	2	3	0	0
Semi-Dry	12	0	0	12	0	0	0	2	6	4	7	2	1	1	1	0
Semi-Sweet	42	0	0	4	33	5	0	0	1	41	26	11	4	0	1	0
Sweet	15	0	0	0	0	15	0	0	0	15	8	6	0	0	0	1
Total	96	23	4	16	33	20	25	4	7	60	63	19	7	4	2	1

9.2 The orchard-based scale

The orchard-based scale requires the titratable acidity and the total polyphenol content, neither of which is measured by NMR. The ratio was therefore calculated with the sum of the organic acids expressed as malic acid in place of the titratable acidity, and no adjustment for polyphenols was made. Both simplifications bias the result. The sum of the acids includes the fraction that is neutralized by the potassium of the juice, so it exceeds the titratable acidity and the ratio is too low; and the omitted adjustment would move ciders rich in tannin further toward dry. The numbers of Table 13 are consequently an approximation that errs, for most ciders, toward the dry side. Even so, the picture differs sharply from that of the sugar scales (Figure 30). Of the 96 ciders, 60 have a ratio above 3 and would be called sweet, 25 are dry, and only 11 fall into the two intermediate classes. Of the 42 ciders that are semi-sweet by their sugar content, 41 are sweet by the ratio. The median ratio of the whole set is 4.5, and a quarter of the ciders exceed 8.6.

This is not a contradiction but a difference of purpose. The orchard-based scale was devised for ciders in which sugar and acid are both those of the fruit, and within that population a ratio of 3 does mark a sweet cider. Applied to the market as a whole it has no resolving power above 20 g/L of sugar, because the acid content of commercial ciders is confined to a narrow band of 4 to 6 g/L and hardly any of them could reach a

ratio below 3 at that sugar level. Its value lies at the dry end, where it separates ciders that the sugar scales lump together. Stormalong Legendary Dry, with 5 g/L of sugar, has a ratio of 0.9 in the lots that retain their malic acid and of 2.0 in the lot that underwent malolactic conversion; the sour ciders of Graft, with up to 12 g/L of acid, remain dry on any scale; and Downeast Cranberry Cider, sweet by its 49 g/L of sugar, has by virtue of its exceptional acidity a lower ratio (4.3) than many ciders with half as much sugar. For a full application of the scale the NMR measurement is combined with a titration and a polyphenol assay.

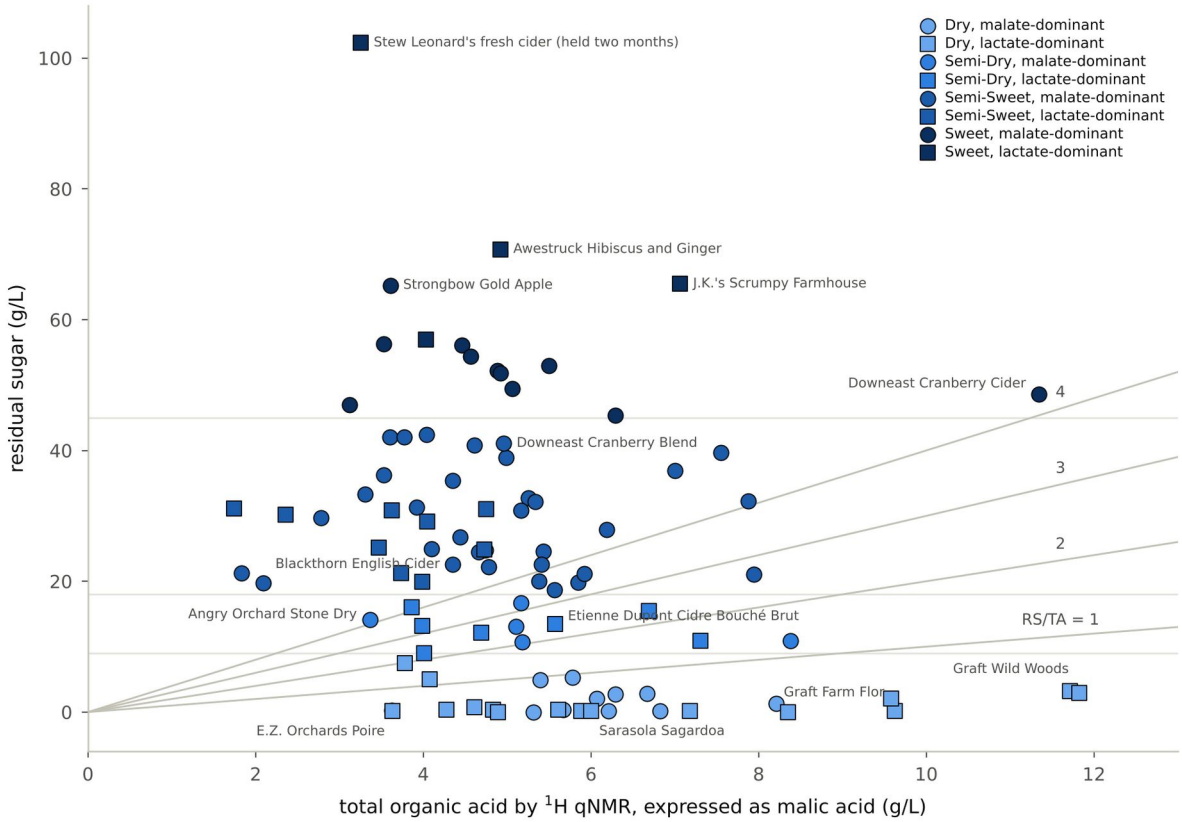

Figure 30. Residual sugar against total organic acid (sum of the acids by ¹H qNMR, expressed as malic acid) for the 96 ciders. Color gives the dryness class of the American Cider Association, marker shape the dominant acid. Horizontal lines: class boundaries at 9, 18 and 45 g/L of sugar. Rays from the origin: ratios of residual sugar to acid of 1, 2, 3 and 4, the boundaries of the orchard-based scale without polyphenol adjustment.

9.3 European designations

Applied to the same residual sugar contents, the European designations of Section 3.6 sort the ciders differently (Figure 31). By the Spanish standard for cider, 43 of the 96 ciders would be *extra-seca*, 18 *seca*, 24 *semi-seca* and 11 *dulce*; every cider that is dry or semi-dry on the American scale, and four of the semi-sweet ones, fall into the first class. By the French thresholds of 1953, 59 would be *brut* and 19 *demi-sec*, and 18 exceed the sugar content of *doux*. Only one of the latter, the fresh juice that had begun to ferment, also satisfies the alcohol limit of that designation; the commercial ciders of this group, with 4.4 to 6.6 % v/v of alcohol, combine a sugar content and an alcoholic strength that the French classification does not provide for. The terms of the European wine regulations agree most closely with the American scale: 26 ciders are dry, 10 medium dry, 45 medium and 15 sweet. Of the 27 ciders with less than 9 g/L of sugar, all but one are dry by the wine terms as well, either because they contain no more than 4 g/L or because their acidity, expressed as tartaric acid, is not more than 2 g/L below their sugar content.

The European ciders of the collection are placed as their labels lead one to expect. Sarasola Sagardoa and Etienne Dupont Cidre Triple, with less than 0.5 g/L of sugar, are dry on every scale. Etienne Dupont Cidre Bouché Brut, with 13.4 g/L, is *brut* in France, *extra-seca* in Spain, medium dry by the wine terms and semi-dry in the United States. Celt Cidre Breton Traditionnel (30.9 g/L) and Thatchers Green Goblin (30.2 g/L) are *demi-sec*, *semi-seca* and semi-sweet, and Blackthorn (22.6 g/L) and Samuel Smith's Organic Cider (24.4 g/L) are *brut* by the French and *seca* by the Spanish threshold while being semi-sweet for the American Cider Association. The comparison makes the point of Section 3.6 with real products: the same cider carries a different word in each market, and only the number is unambiguous.

9.4 Further scales that can be read from the same spectrum

Residual sugar is only one of the properties that a consumer perceives and that a producer may wish to communicate, and the spectrum that yields the sugar content yields the others as well. Seven further scales were therefore derived from the qNMR composition. Their definitions are given in Table 14, the distribution of the 96 ciders over their classes in Figure 31, and the placement of six ciders of contrasting style on all of them in Figure 32. The boundaries are proposals, chosen from the distribution of the present survey and from the anchors named in the table.

Table 14. Further scales proposed for the description of cider, all derived from the ¹H qNMR composition.

Scale	Quantity	Classes and boundaries	Basis
Balance	sweetness equivalent divided by the acids expressed as malic acid (both in g/L); sweetness equivalent = sucrose + 1.2 × fructose + 0.7 × glucose + 0.6 × sorbitol + 0.6 × glycerol	Bone dry, below 1; Dry, 1 to 2; Off-dry, 2 to 4; Medium, 4 to 8; Sweet, above 8	relative sweetness of sugars and polyols;[123,232–234] each class spans a doubling of the index
Acidity	sum of malic, lactic, succinic, citric, acetic and quinic acids, expressed as malic acid (g/L)	Low, below 3.5; Medium, 3.5 to 5; High, 5 to 7; Very high, above 7	distribution of the survey; 4.5 g/L separates sweet from sharp juice in the Long Ashton classification
Acid profile	molar share of lactic acid in the sum of lactic and malic acids; acetic acid (g/L)	Malic, share below 0.1; Mixed, 0.1 to 0.5; Lactic, 0.5 or more; Lactic-acetic, 0.5 or more with at least 1 g/L of acetic acid	extent of malolactic conversion (Section 8.4)
Body	glycerol + sorbitol (g/L)	Light, below 5; Medium, 5 to 8; Full, above 8	non-volatile polyols that add viscosity and sweetness[24,235]
Aromatic index	aromatic protons between 7.80 and 6.76 ppm (mmol/L)	Low, below 2.5; Medium, 2.5 to 5; High, above 5	integral measure of polyphenols (Section 8.6)
Alcohol tier	ethanol (% v/v)	below 0.5; 0.5 to less than 7; 7 to less than 8.5; 8.5 or more	regulatory boundaries in the United States (Section 3.7)[67,117]
Energy band	energy per 12 fl oz, component basis (kcal)	below 140; 140 to 170; 170 to 200; above 200	distribution of the survey (Section 10)

The polyols require the freeze-dried tube, so that balance and body are available for 87 of the 96 ciders.

The balance index carries furthest. It divides a sweetness equivalent by the acidity. The sweetness equivalent weights each sweet constituent by its sweetness relative to sucrose, so that it credits a cider with the greater sweetness of fructose and with its glycerol and sorbitol, which the sugar scales ignore; the division by the acid content expresses, as the orchard-based scale does, that acidity suppresses the perception of sweetness. With boundaries at 1, 2, 4 and 8, 12 of the 87 ciders are bone dry, 10 dry, 11 off-dry, 27 medium and 27 sweet. Unlike the ratio of the orchard-based scale, the index resolves the upper half of the market: the ciders that are semi-sweet by their sugar content divide into 21 that are medium and 15 that are sweet by balance. Thatchers Green Goblin and Magners Original Irish Cider, with about 30 g/L of sugar but

little acid, have indices of 12.8 and 11.2 and are sweet; Naked Flock Hard Cider, with 40 g/L of sugar and 7.6 g/L of acid (6.0), and Downeast Cranberry Cider, with 49 g/L of sugar and 11 g/L of acid (4.7), are medium. At the dry end the index separates the ciders that contain no sugar according to their polyols and their acidity: Sarasola Sagardoa (0.4) is bone dry, and Etienne Dupont Cidre Triple (1.4), with 10 g/L of glycerol and sorbitol, is dry.

The remaining scales describe properties that no dryness scale addresses. Acidity is classed by the sum of the organic acids expressed as malic acid; 10 ciders are low, 41 medium, 30 high and 15 very high, the last class consisting chiefly of the sour ciders and of those made with cranberry. The acid profile states which acid is tasted: malic acid in 39 ciders, a mixture in 21, lactic acid in 24, and lactic together with acetic acid in 12. It is the most direct indication of style that the spectrum offers. Of the nineteen mass-market ciders 14 have a malic profile and none a lactic-acetic one, whereas none of the seven traditional European ciders is malic. Body is represented by the sum of glycerol and sorbitol, the two non-volatile constituents that lend viscosity and a residual sweetness to a dry cider: 16 ciders are light, 43 medium and 28 full, with the perry of E.Z. Orchards and the French ciders at the upper end and the ciders whose glycerol has been metabolized (Section 8.5) at the lower. The aromatic index, the integral of the aromatic region, stands for the polyphenols that carry bitterness and astringency; it is low in 21 ciders, medium in 55 and high in 20, among them six of the seven European ciders. The alcohol tiers follow the regulatory boundaries of the United States: two ciders lie below 0.5 % v/v, 82 between 0.5 and 7 %, nine between 7 and 8.5 %, where the labeling rules for wine apply while the tax rate for hard cider is retained, and three at 8.5 % or above. The energy bands, finally, divide the collection into 14 ciders below 140 kcal per 12 fl oz, 51 between 140 and 170 kcal, 23 between 170 and 200 kcal and eight above.

The cards of Figure 32 show how the scales combine into a description of an individual product. Sarasola Sagardoa and Bad Seed Hard Apple Cider (Dry) are indistinguishable on the dryness scale, both with 0.2 g/L of sugar, and are bone dry by balance; but the first has a lactic-acetic profile, a light body and the markers of bacterial glycerol metabolism, while the second is a malic cider of medium body in the alcohol tier of 7 to 8.5 %. Etienne Dupont Cidre Bouché Brut combines a semi-dry sugar content with a lactic profile, a full body and a high aromatic index, the profile of a keeved cider from tannic fruit. The perry of E.Z. Orchards is dry by its sugar and medium by balance, because its 26 g/L of polyols, nearly all of it sorbitol, enter the sweetness equivalent. Stormalong Boston Heirloom and Angry Orchard Crisp Apple, finally, are both sweet, but the former owes its sweetness chiefly to fructose and glucose in the presence of more than 6 g/L of acid, the latter chiefly to sucrose. A description of this kind is obtained from the one spectrum that also supplies the nutrition values, and it can be printed on a label, used in a specification or followed from batch to batch.

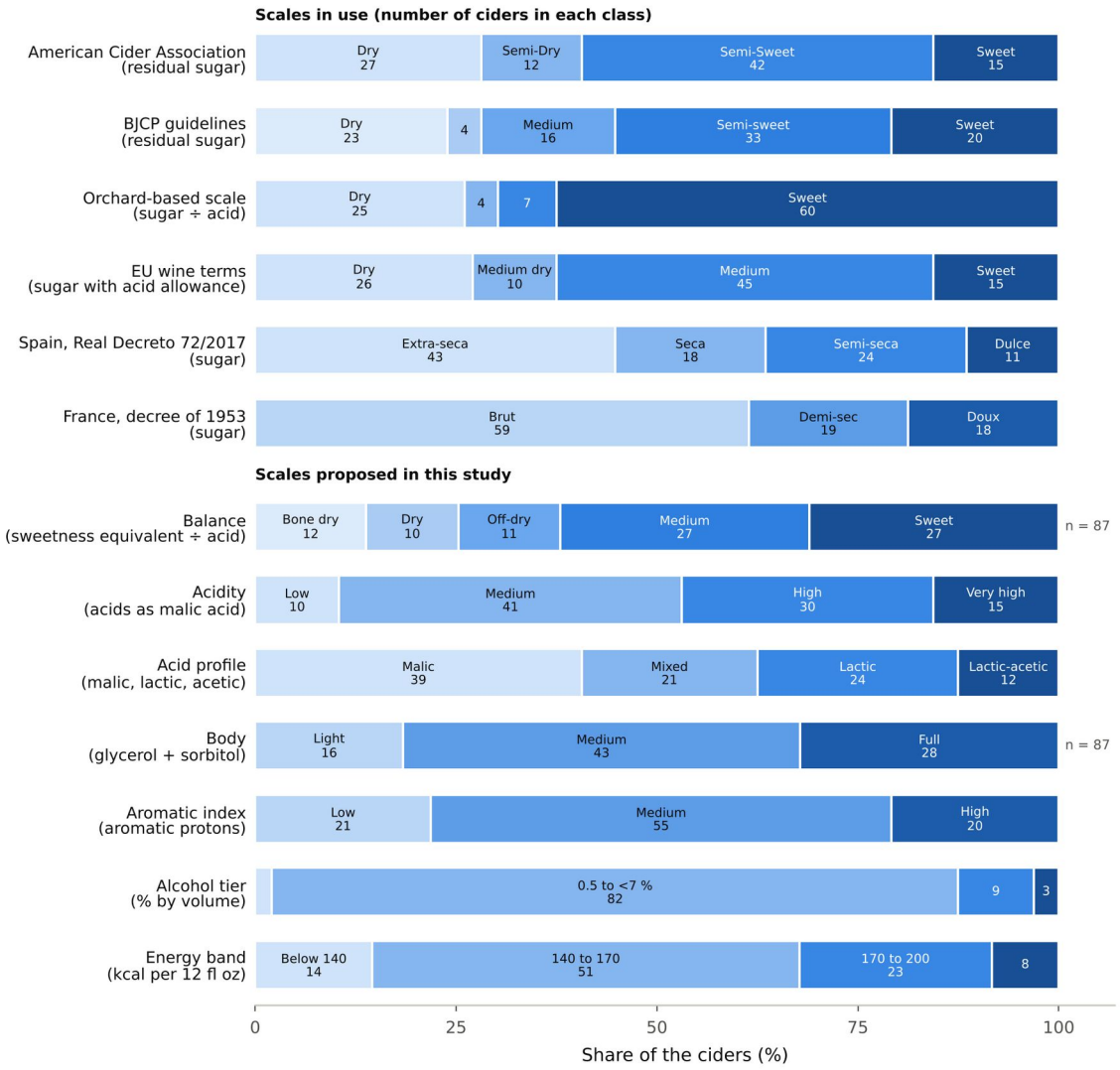

Figure 31. Distribution of the 96 ciders over the classes of the dryness scales and designations in use (upper part) and of the scales proposed in this study (lower part). The numbers are the ciders in each class. Balance and body require the polyols and are given for the 87 ciders with a freeze-dried tube.

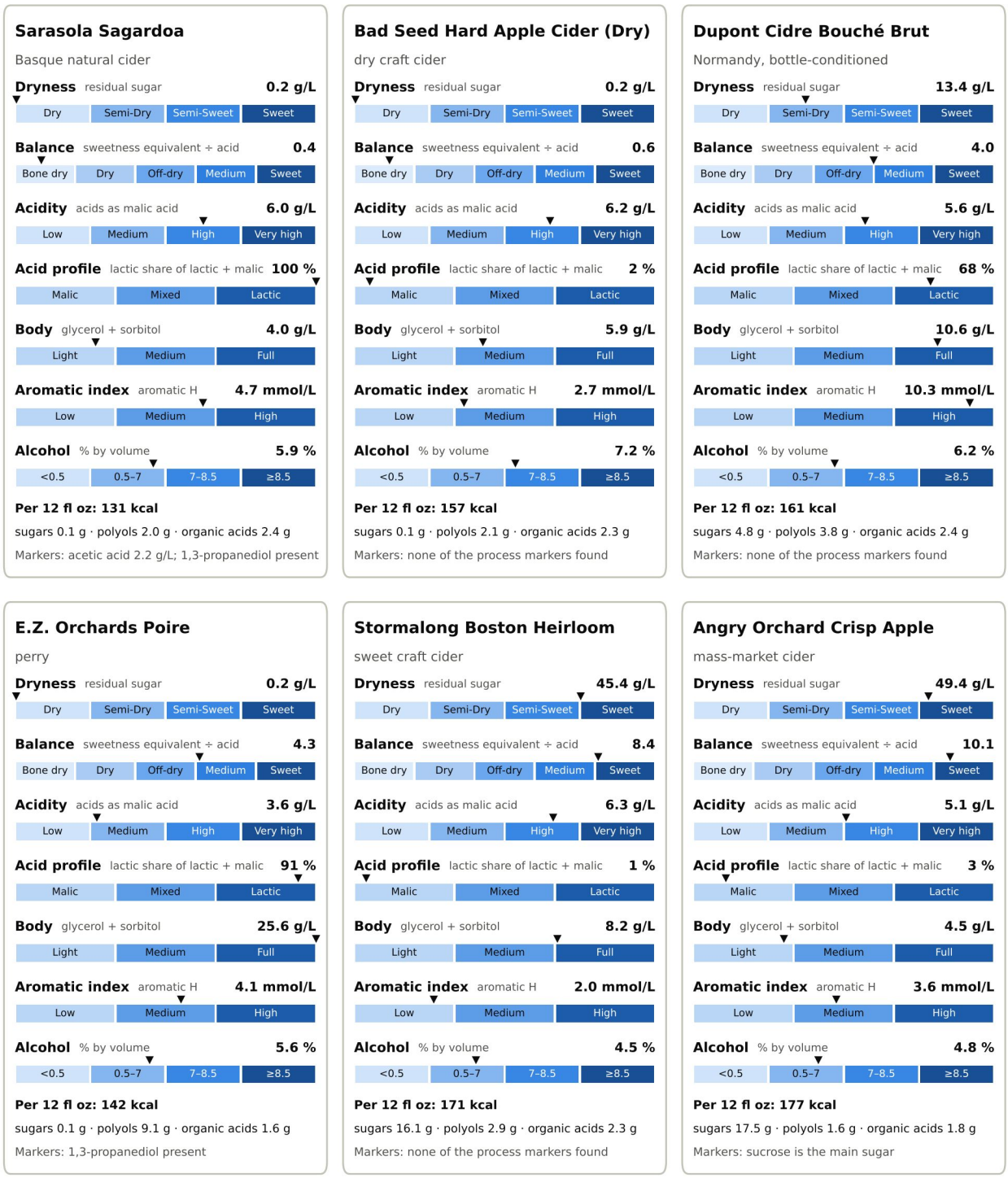

Figure 32. Profile cards of six ciders of contrasting style, calculated from the ¹H qNMR composition. Each card places the cider on the dryness scale of the American Cider Association and on six of the scales proposed here; the triangle marks the measured value within its class. Below: energy, sugars, polyols and organic acids per 12 fl oz, and the process markers found (acetic acid of 1 g/L or more, 1,3-propanediol, sorbic acid, 5-hydroxymethylfurfural, sucrose as the principal sugar).

10. Nutrition Labeling

10.1 Calculation of energy and carbohydrate

Table 15. Energy conversion factors used for the two calculations of energy.

Constituent	Component basis (kcal/g)	Source	TTB method basis (kcal/g)
Ethanol	7.07	specific factor, USDA Handbook No. 74[121]	6.9
Fructose, glucose, sucrose	4	general factor for carbohydrate[116]	4 (sugars)
Sorbitol	2.6	21 CFR 101.9(c)(1)(i)(F)[116]	2.4 (carbohydrate other than sugars)
Glycerol	4.32	factor for glycerin[123]	2.4
1,3-Propanediol	4	assigned; no regulatory factor exists	2.4
Organic acids	3	FAO; Regulation (EU) No 1169/2011[124,125]	2.4
Minor carbohydrates (5.18 ppm doublet, galacturonic acid)	not counted		2.4

Energy was calculated from the measured composition in two ways (Table 15). The component basis applies to each constituent the factor that the food labeling regulations or, where these are silent, international conventions assign to it. The second basis follows the formula of the TTB laboratory method quoted in Section 3.7, in which alcohol is counted at 6.9 kcal/g, sugars at 4 kcal/g and all other carbohydrate at 2.4 kcal/g; total carbohydrate was represented by the sum of sugars, polyols, organic acids and the minor carbohydrates, which is what a determination by difference would include, protein and fat being negligible in cider. The TTB method was written for flavored malt beverages, and its use for cider here is an analogy. Protein, which is not measured, would add at most a few kilocalories per serving.

On the component basis the 93 alcoholic ciders contain 115 to 247 kcal per 12 fl oz with a median of 161 kcal (Figure 33); the two non-alcoholic ciders contain 35 and 41 kcal. Ethanol supplies 71 % of the energy in the median cider (range 45 to 97 %), sugars 20 % (up to 46 %), polyols 5 % and organic acids 4 %. The last two terms are small but not negligible: together they amount to 15 kcal per serving in the median cider and to 32 kcal in the perry, and a calculation from alcohol and sugar alone would understate the energy of a dry cider by 11 % in the median (7 to 21 %). The TTB basis gives values that are lower by 3.8 % in the median, in about equal parts because of the lower factor for alcohol and because of the single factor of 2.4 kcal/g applied to glycerol, sorbitol and the acids.

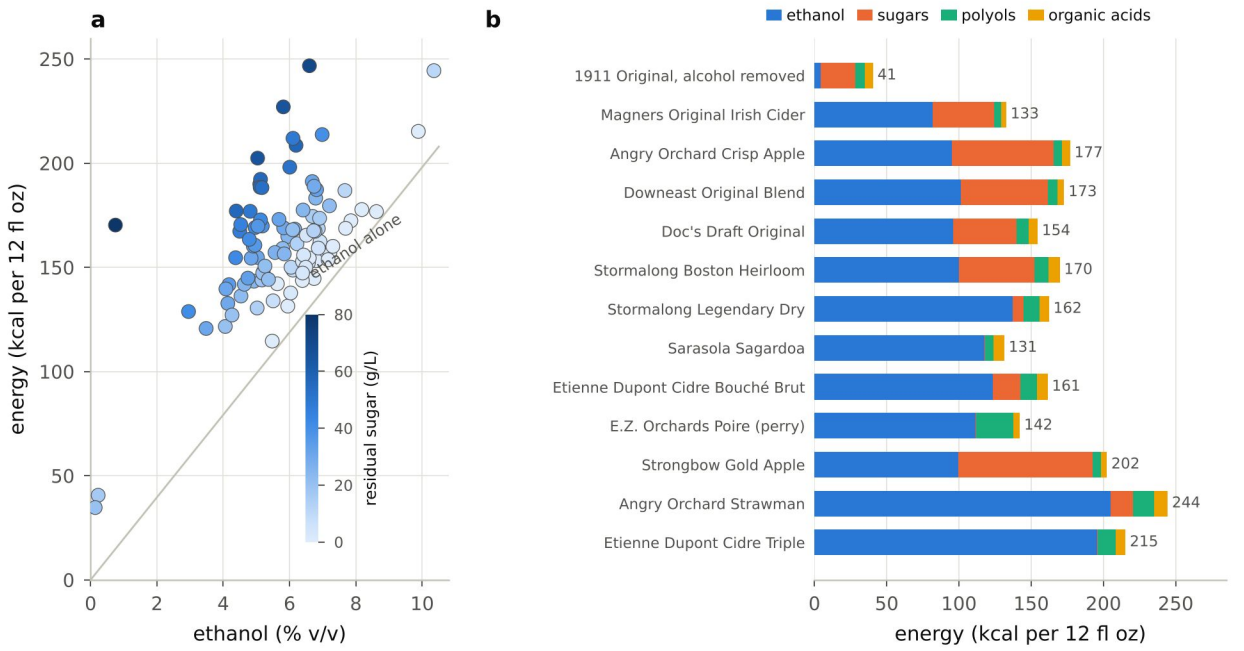

Figure 33. Energy content of the ciders calculated from the ¹H qNMR composition (component basis). (a) Energy per 12 fl oz against ethanol; color gives the residual sugar, and the line is the energy of the ethanol alone. (b) Contributions of ethanol, sugars, polyols and organic acids for thirteen ciders that span the range of composition; the number is the total in kcal per 12 fl oz.

10.2 Alcohol, carbohydrate and energy of a serving, by brand

The three quantities that a label must or may state are compared for all 96 ciders in Figure 34, Figure 35 and Figure 36. In each figure the ciders are ranked from the largest value to the smallest and identified by brand and product. The alcohol contents (Figure 34) lie between 4 and 7 % v/v for 79 of the 96 ciders. Angry Orchard Strawman (10.4 % v/v) and Etienne Dupont Cidre Triple (9.9 % v/v) head the list, followed by the barrel-aged ciders of Hudson Valley Farmhouse Cider and Bad Seed; the fresh juice and the two non-alcoholic ciders close it. The declared alcohol content is marked where a declaration was recorded, and the broken lines show the regulatory tiers of Section 3.7.

Carbohydrate is shown in Figure 35 as the grams of each sugar and polyol in a serving of 12 fl oz. The fresh, partly fermented juice leads with 38 g; among the ciders proper the range extends from 28 g in Awestruck Hibiscus and Ginger, 27 g in J.K.'s Scrumpy and 25 g in Strongbow Gold Apple down to 3 g in the median of the ciders without residual sugar. The colors show at a glance how a cider acquired its sweetness. Sucrose dominates in Angry Orchard Green Apple, Woodchuck Gumption and the fruit ciders of Downeast, fructose and glucose in the ciders sweetened with juice or by arrested fermentation, and in the dry ciders at the foot of the list the bar consists of glycerol and sorbitol alone, an amount that no statement of sugars would reveal. In the perry of E.Z. Orchards nearly the whole of the 9 g is sorbitol, and in the ciders discussed in Section 8.5 1,3-propanediol appears in place of glycerol.

Energy (Figure 36) ranges from 115 to 247 kcal per serving among the alcoholic ciders, and each bar is divided into the contributions of ethanol, sugars, polyols and organic acids. Awestruck Hibiscus and Ginger (247 kcal) and Angry Orchard Strawman (244 kcal) arrive at nearly the same total by opposite routes, the first through sugar and the second through alcohol. The ciders in the middle of the list, around 160 kcal, include sweet ciders of low strength and dry ciders of high strength side by side, so that neither the alcohol content nor the dryness alone predicts the energy of a serving. The energy declared on the label is marked for the eight ciders for which it was recorded (Section 10.3).

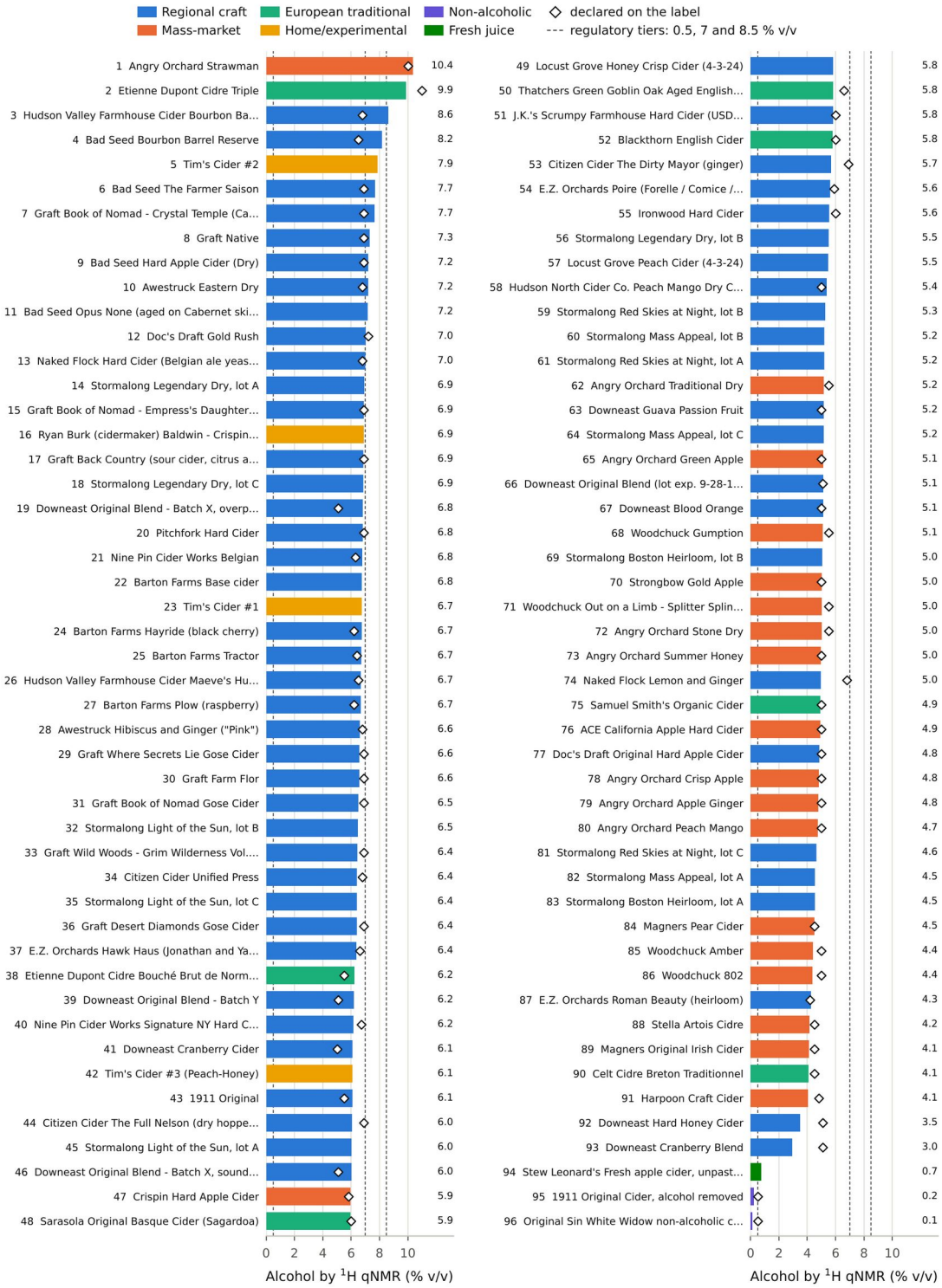

Figure 34. Alcohol content of the 96 ciders by ¹H qNMR (% v/v), ranked from the highest to the lowest, with brand and product names. Color gives the market segment. Diamonds: alcohol content declared on the label. Broken lines: regulatory tiers at 0.5, 7 and 8.5 % v/v.

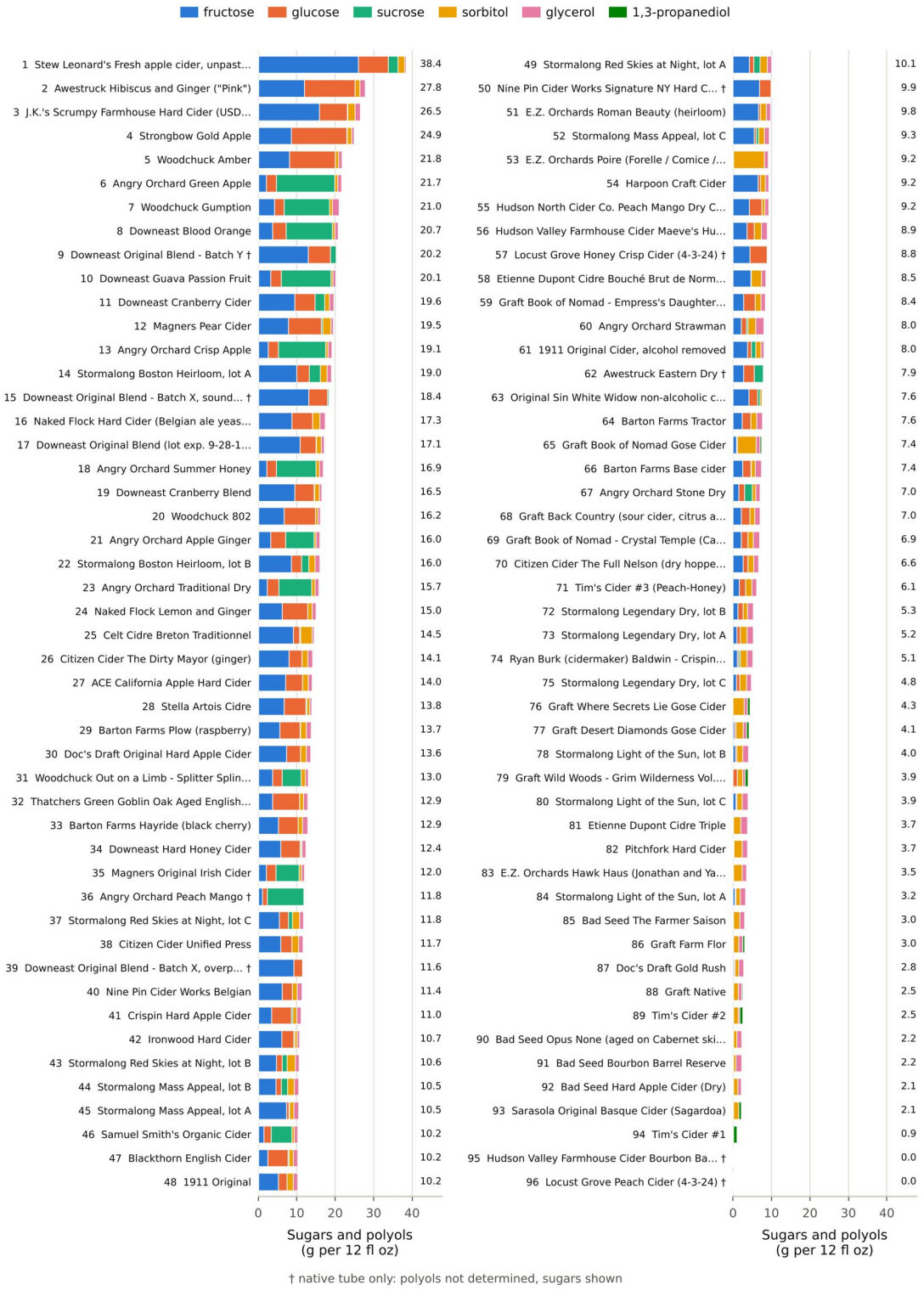

Figure 35. Sugars and polyols of the 96 ciders in a serving of 12 fl oz (g), ranked from the highest to the lowest sum, with brand and product names. The bars are divided into fructose, glucose, sucrose, sorbitol, glycerol and 1,3-propanediol; the number is the sum. † only the native tube was measured: sugars shown, polyols not determined.

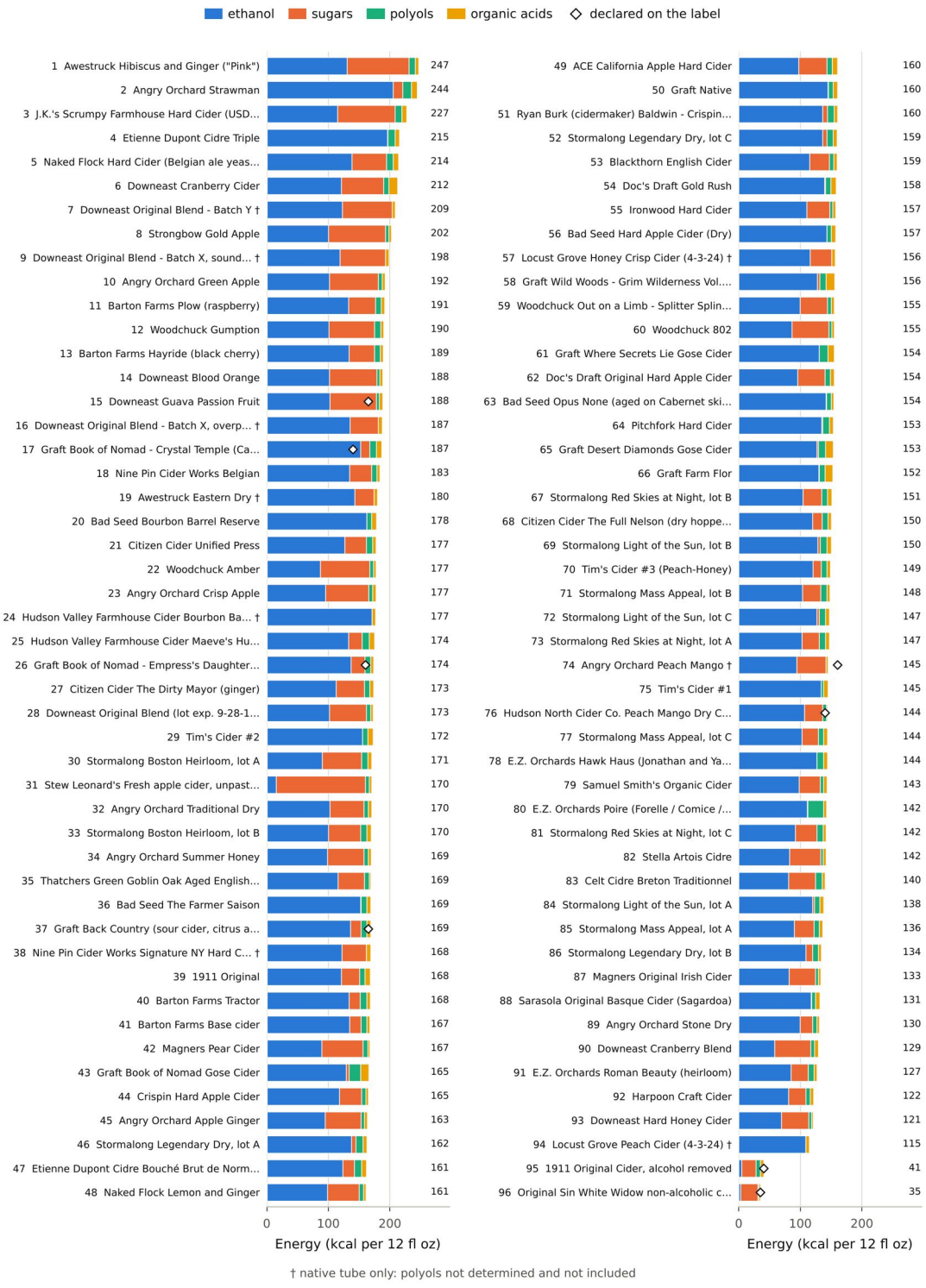

Figure 36. Energy of the 96 ciders in a serving of 12 fl oz (kcal, component basis), ranked from the highest to the lowest, with brand and product names. The bars are divided into the contributions of ethanol, sugars, polyols and organic acids; the number is the total. Diamonds: energy declared on the label. † only the native tube was measured: polyols not determined and not included.

10.3 Comparison with declared values

Eight of the ciders measured since 2022 carried a Nutrition Facts panel whose values were noted at the time of the analysis. They are compared with the values derived from the spectra in Table 16. For the two non-alcoholic ciders the agreement is close in every respect (40 and 35 kcal declared, 41 and 35 kcal found). Among the alcoholic ciders, Hudson North Peach Mango (+3 %) and Graft Back Country (+2 %) agree with their declarations within a few percent, and Graft Empress's Daughter (+8 %) within a tolerable margin, although its sugar content, 5.7 g per serving, is nearly twice the 3 g declared. Downeast Guava Passion Fruit declares 165 kcal and yields 188 kcal (+14 %), a difference that follows from its 53 g/L of sugar. Angry Orchard Peach Mango yields less energy than declared (145 against 160 kcal), in part because no freeze-dried tube was measured and the polyols are missing from the sum. The largest discrepancy is that of Graft Book of Nomad Crystal Temple, which declares 140 kcal, 2 g of carbohydrate and no sugar per serving. The spectrum shows 7.7 % v/v of ethanol against a declared 6.9 % v/v, 3.9 g of sugar and 6.9 g of sugars and polyols per serving, and 187 kcal (+33 %). Even at the declared alcohol content the ethanol of this cider alone accounts for 133 to 137 kcal, depending on the factor, so that the declared energy leaves room for less than 2 g of anything else.

Table 16. Declared nutrition values per 12 fl oz and values derived from the ¹H qNMR composition for eight ciders.

Cider	Ethanol declared (% v/v)	Ethanol found (% v/v)	Energy declared (kcal)	Energy found, component basis	Energy found, TTB basis	Sugars declared (g)	Sugars found (g)	Carb. declared (g)	Sugars + polyols found (g)	Carb. by difference, estimated (g)
Downeast Guava Passion Fruit	5.0	5.2	165	188	183	–	18.8	–	20.1	22.2
1911 Original Cider, alcohol removed	<0.5	0.2	40	41	38	7	5.9	8	8.0	9.9
Original Sin White Widow	<0.5	0.1	35	35	34	–	7.0	9	7.6	8.4
Hudson North Cider Co. Peach Mango Dry Cider	5.0	5.4	140	144	140	8	7.5	12	9.2	9.9
Angry Orchard Peach Mango	5.0	4.7	160	145	142	10	11.8	13	–	–
Graft Back Country	6.9	6.9	165	169	161	4	4.3	7	7.0	9.1
Graft Empress's Daughter	6.9	6.9	160	174	167	3	5.7	6	8.4	10.1
Graft Crystal Temple	6.9	7.7	140	187	178	0	3.9	2	6.9	9.9

Serving of 12 fl oz (354.9 mL). Carb., total carbohydrate. Declared values as recorded with the spectra at the time of analysis. Carbohydrate by difference is estimated as sugars, polyols, organic acids and minor carbohydrates. For Angry Orchard Peach Mango only the native tube was measured.

The comparison exposes an ambiguity of the term carbohydrate that affects every dry cider. Total carbohydrate on a Nutrition Facts panel is determined by difference and therefore includes glycerol, sorbitol and the organic acids, none of which is a sugar. The dry ciders of this study contain 7 to 9 g/L of polyols and 6 to 8 g/L of acids (interquartile ranges), that is about 5 g of carbohydrate by difference per 12 fl oz, even when the sugar content is nil. Declarations of 0 to 2 g of carbohydrate, such as that of Crystal Temple, are not compatible with this composition, while the 6 and 7 g declared for the other two Graft ciders correspond to sugars and polyols without the acids. The tolerance applied by the Food and Drug Administration to declared calories and sugars is that the amount found shall not exceed the declaration by more than 20 %.[116] On the present analysis one of the eight ciders exceeds that margin for energy and two exceed it for sugars. These are results for one container of each product. They illustrate the use of the method for the verification

of declared values, for which a single spectrum supplies alcohol, sugars, polyols, organic acids and energy together.

10.4 Label panels calculated from the spectrum

The quantities of the preceding sections can be assembled into the form in which the consumer meets them. Figure 37 shows such panels for three mass-market ciders of increasing sweetness, Angry Orchard Stone Dry, Angry Orchard Crisp Apple and Strongbow Gold Apple. The panels are illustrations calculated from the ¹H qNMR composition of the containers that were analyzed; they are not the labels of the producers. The serving is 12 fl oz, energy is given on the component basis, and the values are rounded as the food labeling regulations prescribe, energy to the nearest 10 kcal above 50 kcal and carbohydrate to the nearest gram.[116] Total carbohydrate is the carbohydrate by difference of Section 10.1, that is, the sum of sugars, sugar alcohols and organic acids; the sugars are subdivided to show sucrose; and the alcohol is stated in percent by volume, in grams and in fluid ounces of pure alcohol per serving, the form that the Alcohol Facts proposal of 2025 envisages.[120] Fat and protein, which are negligible in cider, and sodium are not determined by ¹H NMR. Below each panel the residual sugar is placed on the dryness scale of the American Cider Association and accompanied by the ratio of sugar to acid.

The three ciders contain 5.0, 4.8 and 5.0 % v/v of alcohol and are thus practically equal in strength, but they yield 130, 180 and 200 kcal per serving, with 5, 18 and 23 g of sugars. Sucrose makes up 40 % of the sugar in the first and 70 % in the second; in the third no sucrose was found, and the sugar consists of glucose and fructose. A single spectrum supplies every entry of the panel that depends on the organic composition, together with the dryness designation, and the same calculation can be carried out for any cider of Appendix B.

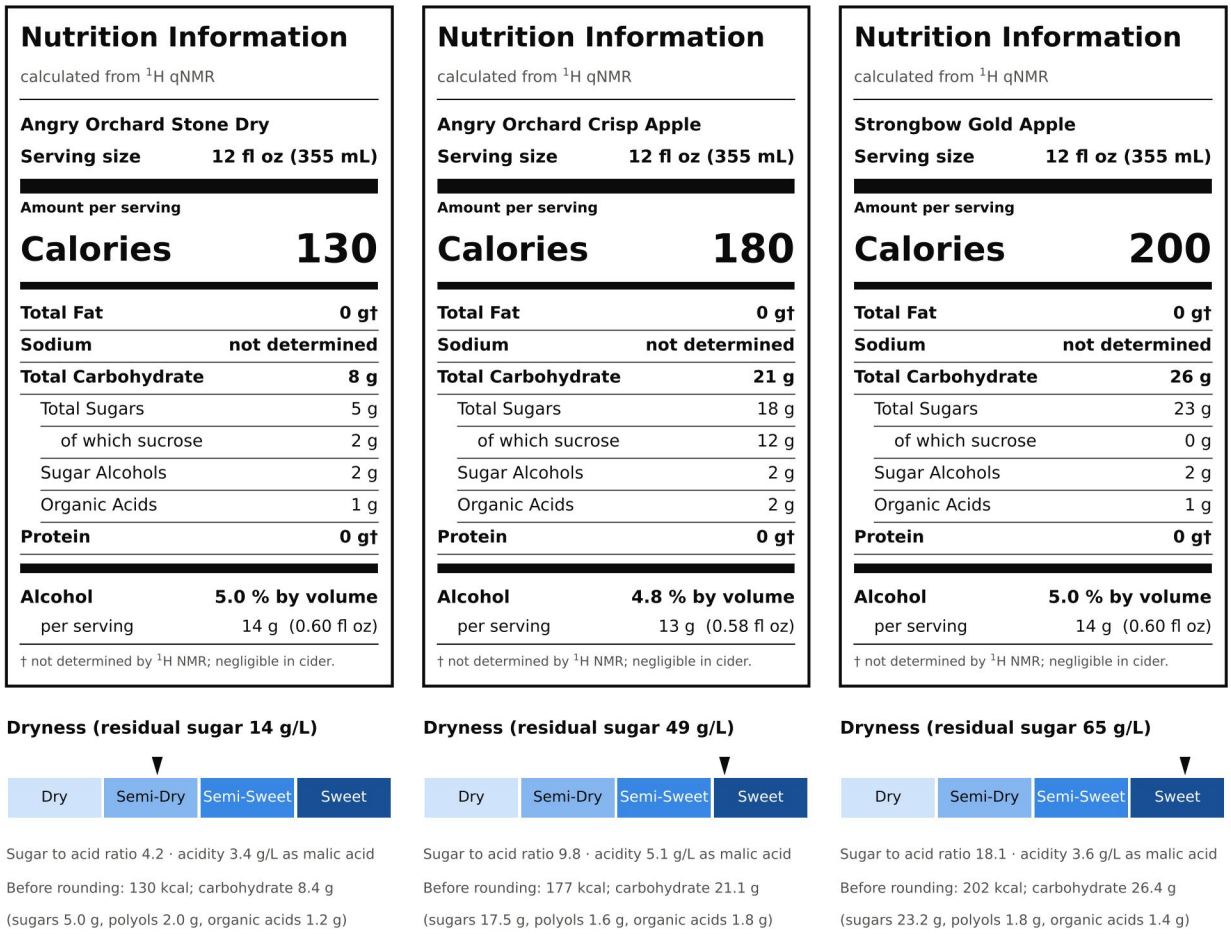

Figure 37. Nutrition panels calculated from the ¹H qNMR composition for three ciders (serving of 12 fl oz; values rounded according to the food labeling regulations), each with the placement of its residual sugar on the dryness scale of the American Cider Association. The panels are illustrations derived from the measurements of this study and are not the labels of the producers. † not determined by ¹H NMR.

11. Chemometrics: Choice of Pre-processing

11.1 Design of the comparison

A multivariate model sees only the table of numbers that is handed to it, and the way in which a spectrum is turned into that table usually matters more than the algorithm applied afterward.[191,214,236,237] The alternatives listed in Section 5.9 were therefore compared before any model was interpreted: five representations of the spectrum, three normalizations and four treatments of the variables after mean-centering, for the native and the freeze-dried tubes separately. The representations differ greatly in size. The freeze-dried spectra give 187 uniform bins of 0.04 ppm, 746 bins of 0.01 ppm, 187 optimized buckets, 27 targeted integrals and 17 fitted concentrations; the native spectra give 187, 746, 187, 30 and 15.

Three criteria were applied, each of which corresponds to a use of the data. The first asks whether similar samples are close to one another. The Stormalong cans and the cans of Downeast Original Blend provide groups of samples that are known to be the same product; for each of them the nearest neighbor in the pre-processed data was determined, and the fraction for which this neighbor is another lot of the same product was recorded as the lot retrieval rate (14 such samples with a freeze-dried tube, 18 with a native tube). The second criterion is the balanced accuracy of two classifiers, PLS discriminant analysis and a support vector machine, averaged over five labels (dryness, dominant acid, presence of sucrose, market segment and declared adjunct) in cross-validation grouped by producer. The third is the median Q^2 of PLS regression models for 14 (freeze-dried) or 13 (native) constituents. The results are assembled in Figure 38.

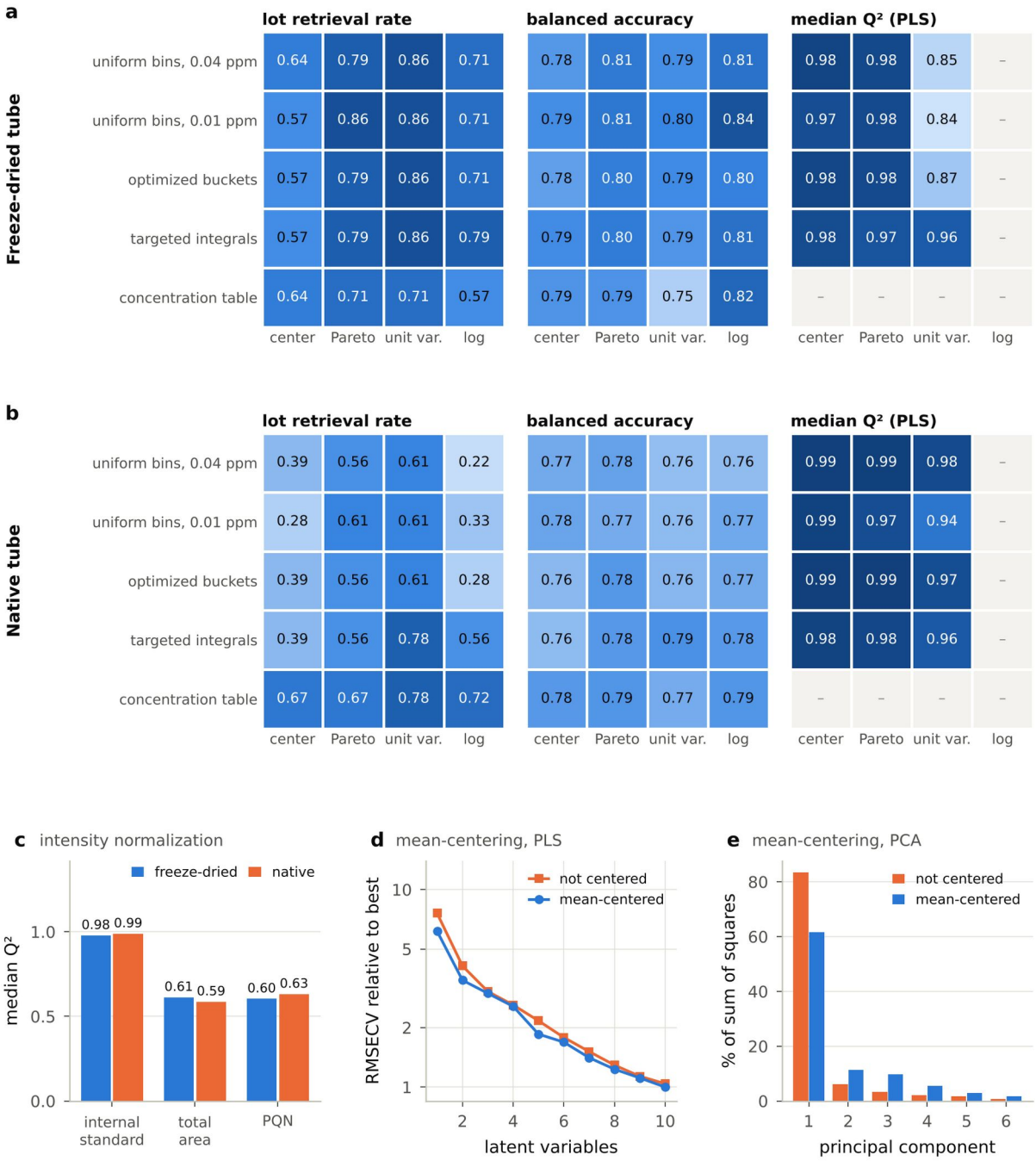

Figure 38. Comparison of pre-processing options. (a) Freeze-dried tube and (b) native tube, with intensities on the absolute scale of the internal standard: lot retrieval rate, mean balanced accuracy of two classifiers over five labels, and median Q² of PLS regression, for five representations of the spectrum (rows) and four treatments after mean-centering (columns). (c) Median Q² of PLS regression on optimized buckets (Pareto scaling) for three normalizations. (d) Effect of mean-centering on PLS regression: cross-validated error relative to the best model, median over ten constituents, against the number of latent variables. (e) Effect of mean-centering on PCA: share of the total sum of squares carried by each component.

11.2 Normalization

The decisive factor for quantitative work is the normalization. With intensities on the absolute scale provided by the internal standard, the median Q² of the PLS models on optimized buckets is 0.98 for the

freeze-dried and 0.99 for the native tube. Normalization to total area lowers it to 0.61 and 0.59, and probabilistic quotient normalization to 0.60 and 0.63 (Figure 38c). The reason is simple. Both normalizations were devised for samples of unknown dilution, such as urine, in which only relative composition has a meaning.[190,214] A cider is not diluted to an unknown degree. Its total signal is dominated by ethanol in the native tube and by sugar in the freeze-dried tube, and dividing by it removes exactly the information that a quantitative model needs, while it couples every other variable to the dominant constituent. For classification and for the retrieval of lots the normalization makes no consistent difference. Since the absolute scale is available at no cost, it was retained for all further work.

11.3 Mean-centering and scaling

Mean-centering is not a matter of choice for the methods used here, although its omission is not always visible. Without centering, the first principal component is the mean spectrum: it carries 84 % of the total sum of squares and is collinear with the mean (cosine 0.995), whereas after centering the first component carries 61 % of the variance and describes the differences between ciders (Figure 38e). A PLS model built on uncentered data must likewise spend its first latent variable on the mean, and its error is higher at every model size up to the point where enough latent variables are available to absorb the difference; with one latent variable the cross-validated error is 7.6 times that of the best model against 6.2 times for the centered model (Figure 38d). Methods based on distances are unaffected, since a common offset does not change a distance, and the usual software implementations of PCA and PLS center the data internally whether asked to or not.[215] All data were mean-centered, with the mean taken from the training samples only.

The scaling that follows centering determines how much weight a small signal receives. Without scaling, the buckets of the sugars and of ethanol account for nearly all of the variance, and the similarity of two ciders is judged by their sugar content alone: the lot retrieval rate is only 0.57 for the freeze-dried and 0.39 for the native tube. Scaling to unit variance gives every bucket the same weight and raises the retrieval rate to 0.86 and 0.61, but it also gives full weight to buckets that contain nothing but baseline, and the median Q^2 of the regression falls to 0.87 in the freeze-dried tube. Pareto scaling, in which each variable is divided by the square root of its standard deviation, lies between the two: retrieval rates of 0.79 and 0.56 with a median Q^2 of 0.98 and 0.99. The logarithmic transformation behaves erratically. It helps where the label is defined by a threshold on a minor constituent and it harms the retrieval of lots in the native tube (0.28), because it magnifies the noise of empty buckets. Pareto scaling was adopted as the only option that is among the best by every criterion.

11.4 Representation of the spectrum

The representation proved to matter least. Uniform bins of 0.04 ppm and optimized buckets of the same nominal width give retrieval rates, balanced accuracies and Q^2 values that agree within the second decimal place in both tubes, and bins of 0.01 ppm, with four times as many variables, bring no gain that could be distinguished from chance. The advantage claimed for optimized bucketing, that a peak is not divided between two variables when its position shifts, has little to act on in these data, because all spectra had been referenced individually and placed on a common axis beforehand. Optimized buckets were nevertheless adopted, on the principle that a variable should correspond to a signal, with the plain statement that the present data would have supported uniform bins equally well.

The two targeted representations deserve a separate comment. The table of 27 to 30 integrals performs as well as the full bucket table by all three criteria, and the table of fitted concentrations, with fewer than twenty variables, gives the highest lot retrieval rate in the native tube (0.78 after scaling to unit variance) because fitting has removed baseline and noise. For a matrix as well understood as cider, the targeted table contains essentially all of the information that the untargeted one contains. What the untargeted buckets add is independence from the assignment: they register a compound that nobody has looked for, and they do not

require the fitting procedure of Section 5.6. The following sections use the optimized buckets, on the absolute scale, mean-centered and Pareto-scaled, and report the concentration table alongside where the comparison is instructive.

The limits of this comparison should be stated. The differences among the three bucket tables, and most differences between scalings in the classification criterion, amount to one or two hundredths of balanced accuracy or to one lot in fourteen. With fewer than one hundred samples such differences are within the variation between cross-validation splits. Only three effects are large enough to be regarded as established: the loss of quantitative information on normalization to total area, the damage done to regression on bucket tables by scaling to unit variance, and the poor retrieval of similar samples from unscaled data.

12. Chemometrics: Unsupervised Analysis

12.1 Principal component analysis

The principal component analysis of the freeze-dried tubes is shown in Figure 39. The first three components carry 61.5, 11.3 and 9.8 % of the variance, and the loadings can be read like spectra. The first component is the sugar content. All of its large positive loadings belong to fructose and glucose, and the only negative loadings of any size belong to lactic acid, succinic acid and 1,3-propanediol, the constituents that increase as a cider becomes drier and more thoroughly fermented. The scores arrange the ciders accordingly from dry at the left to sweet at the right. The second component opposes lactic acid, acetic acid and 1,3-propanediol to sucrose and malic acid. It separates the lactate-dominant ciders, which lie above the axis whatever their sugar content, from the malate-dominant ones, and among the sweet ciders it distinguishes those sweetened with sucrose, at the bottom of the plot, from those that retain the sugars of the juice. The third component combines sucrose with lactic acid against fructose, glucose and malic acid and spreads the sucrose-sweetened group further. In other words, the three largest sources of variance in the spectra are the three decisions and events identified in Section 8: how much sugar remains, whether it is sucrose, and whether the malic acid has been converted.

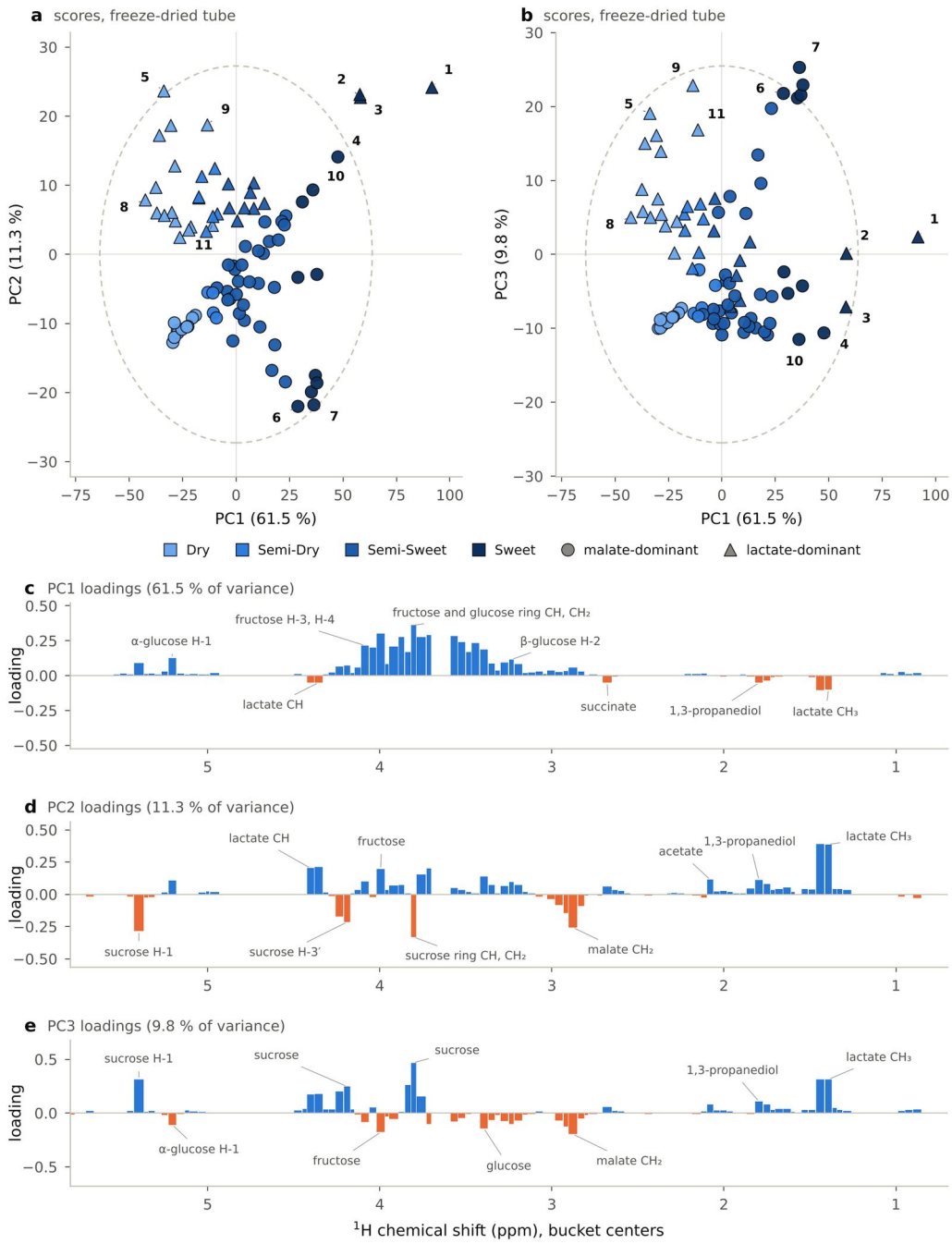

Figure 39. Principal component analysis of the freeze-dried tubes of 87 ciders (optimized buckets, absolute intensities, mean-centered, Pareto-scaled). (a), (b) Scores; color gives the dryness class of the American Cider Association, marker shape the dominant acid; the dashed ellipse is the 95 % limit of Hotelling's T². Numbered points: 1, Stew Leonard's fresh cider (held two months); 2, J.K.'s Scrumpy Farmhouse; 3, Awestruck Hibiscus and Ginger; 4, Strongbow Gold Apple; 5, Graft Wild Woods; 6, Angry Orchard Crisp Apple; 7, Downeast Guava Passion Fruit; 8, Tim's Cider #1; 9, Graft Book of Nomad Gose; 10, Woodchuck Amber; 11, E.Z. Orchards Poire. (c) to (e) Loadings of the first three components against the centers of the buckets, with assignments; buckets above 5.8 ppm have loadings below 0.02 in absolute value and are not shown.

Two diagnostic statistics identify ciders that do not conform to the common pattern. Hotelling's T², calculated for a model of four components, exceeds its 95 % limit (10.4, from the F distribution) for seven ciders. Four of them are the sweetest products of the study, the fresh juice above all; two are the fruit-

flavored ciders of Downeast that are sweetened with sucrose, which are extreme in the third component; and one is Graft Wild Woods, the cider richest in lactic acid. With 87 samples, four or five would be expected beyond such a limit by chance, so that the statistic does little more than name the extremes of the three axes just described. More informative is the residual, the part of a spectrum that the model does not describe. It is largest by a wide margin for the perry of E.Z. Orchards, whose sorbitol content has no counterpart among the apple ciders, and it is elevated for Graft Book of Nomad Gose (sorbitol and acetic acid), Downeast Cranberry Cider (citric and quinic acids) and Tim's Cider #1 (1,3-propanediol in place of glycerol). A large residual is thus a practical flag for a composition outside the experience of the model, which is the principle on which non-targeted screening of fruit juices is based.[126,128]

In the native tubes (Figure 40) ethanol takes the place of sugar as the dominant source of variance. The first component (56.3 %) is the alcohol content, with the two non-alcoholic ciders and the fresh juice at one end and the two strong ciders at the other; the second (19.8 %) is the sugar content, and the third (6.7 %) collects lactic acid, acetic acid and 1,3-propanediol. The loadings of this third component have a feature that deserves attention. They are of the same sign in nearly every bucket between 0.7 and 3.3 ppm, including buckets that contain no assigned signal. A pattern of this kind is not the spectrum of a compound; it is a broad background, either a baseline offset or the unresolved signal of macromolecules in an unfiltered cider, and the sample with by far the largest score, Tim's Cider #2, is a home-made cider. The native bucket table therefore contains a contribution that is not composition in the sense of Section 8, a point that recurs in Section 13.4.

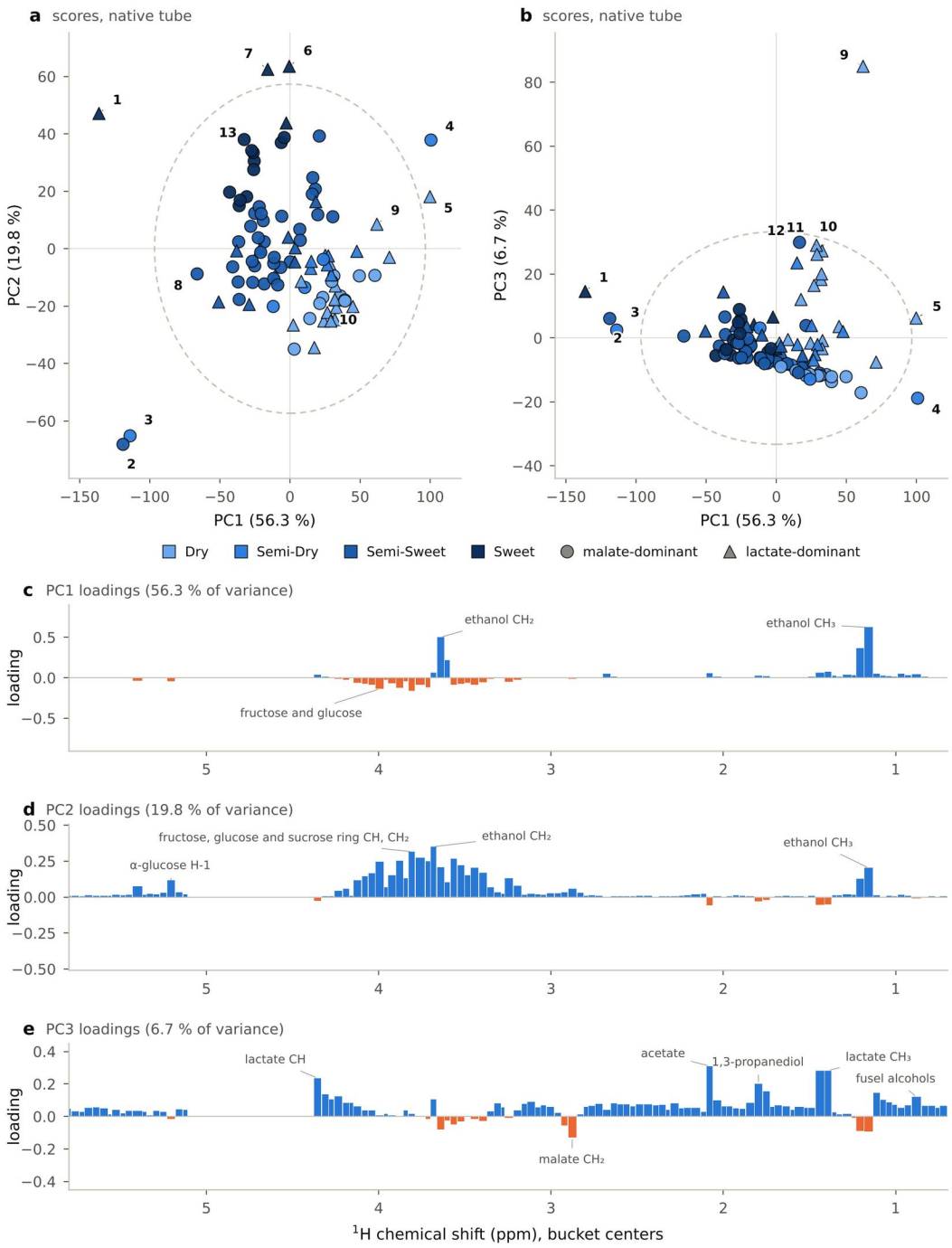

Figure 40. Principal component analysis of the native tubes of 96 ciders (optimized buckets, absolute intensities, mean-centered, Pareto-scaled). (a), (b) Scores, encoded as in the preceding figure. Numbered points: 1, Stew Leonard's fresh cider (held two months); 2, Original Sin White Widow (non-alcoholic); 3, 1911 Original Cider, alcohol removed; 4, Angry Orchard Strawman; 5, Etienne Dupont Cidre Triple; 6, Awestruck Hibiscus and Ginger; 7, J.K.'s Scrumpy Farmhouse; 8, Downeast Cranberry Blend; 9, Tim's Cider #2; 10, Tim's Cider #1; 11, Graft Book of Nomad Gose; 12, Citizen Cider Unified Press; 13, Angry Orchard Green Apple. (c) to (e) Loadings of the first three components with assignments.

The analysis of the concentration table leads to the same components with fewer variables. After Pareto scaling the first two components (42 and 22 %) are ethanol against the hexoses, and sucrose; after scaling to unit variance, which gives the minor constituents equal weight, the variance is spread over many

components (22 % in the first), which shows how little the minor constituents are correlated with one another.

12.2 Hierarchical clustering

Whether the ciders fall into natural groups was examined by hierarchical clustering with Ward's method on the principal component scores of the freeze-dried tubes (Figure 41). The answer is that they do not, or only weakly. The mean silhouette width lies between 0.26 and 0.32 for every number of clusters from two to ten, without a maximum that would indicate a preferred number, and the cophenetic correlation of the dendrogram is 0.52. Values of this size describe a continuous distribution on which any partition is to some extent arbitrary. The partitions obtained from different representations of the same ciders agree only moderately: the adjusted Rand index between the clusterings of the freeze-dried buckets, the native buckets and the concentration table lies between 0.30 and 0.50 for four clusters and falls to 0.22 to 0.40 for five or six.

With that reservation, a division into four clusters gives a useful summary (Table 17). The first cluster (34 ciders) contains the dry and semi-dry ciders, two thirds of them lactate-dominant, with the highest alcohol and acetic acid contents; all home-made ciders, the Basque cider and most ciders of Graft, Bad Seed and Stormalong's dry products are found here. The second and largest (41 ciders) holds the semi-sweet ciders in which the sugars of the juice predominate and malic acid is intact. The third (6 ciders) consists of the products sweetened with sucrose, at a mean of 35 g/L: four ciders of Angry Orchard and Woodchuck and the two fruit-flavored ciders of Downeast. The fourth (6 ciders) comprises the sweetest products of the study, rich in glucose and fructose, among them the fresh juice. No cluster coincides with a measurement campaign. The ciders of each campaign are distributed over the clusters in accordance with their composition, and the lots of one Stormalong product are neighbors or near neighbors in the dendrogram, with the two lots that underwent malolactic conversion placed, correctly, among the lactate-dominant ciders.

Table 17. Mean composition of the four clusters obtained by Ward's method from the freeze-dried tubes (g/L unless stated otherwise).

Cluster	n	Ethanol (% v/v)	Sugar	Sucrose	Malic	Lactic	Acetic	Glycerol	Sorbitol	kcal / 12 fl oz	Dry / Semi-Dry / Semi-Sweet / Sweet	Lactate-dominant
1	34	6.7	5.1	0.2	1.7	3.2	1.00	3.3	5.0	158	25 / 7 / 2 / 0	24 of 34
2	41	5.1	28.3	3.4	3.7	0.5	0.15	2.9	3.6	155	0 / 5 / 34 / 2	6 of 41
3	6	5.0	51.2	35.0	3.2	0.1	0.07	3.0	1.9	184	0 / 0 / 1 / 5	0 of 6
4	6	4.5	67.8	1.4	2.2	1.2	0.31	2.5	4.3	198	0 / 0 / 0 / 6	3 of 6

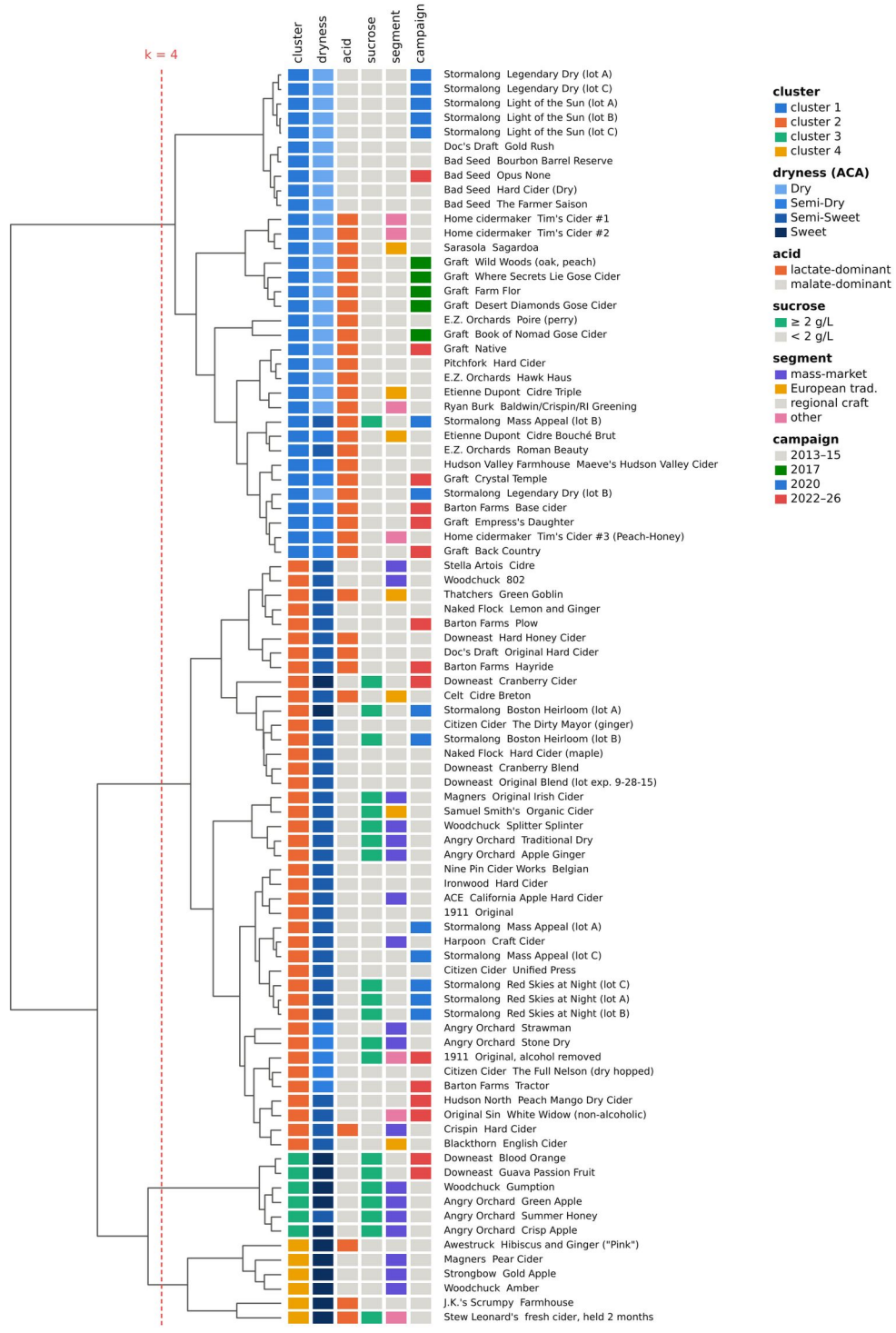

Figure 41. Hierarchical clustering of the freeze-dried tubes of 87 ciders (Ward's method on the principal component scores that carry 97 % of the variance). The strips give, for each cider, the cluster at the four-cluster level (dashed line), the dryness class, the dominant acid, the presence of sucrose at 2 g/L or more, the market segment and the measurement campaign.

12.3 t-SNE and UMAP

Nonlinear embeddings are now routinely shown beside score plots, and their appeal is that they draw separate islands where PCA shows a cloud. That appeal is also their danger.[238,239] Both methods were applied

to the same scores as the clustering (Figure 42), and two properties were measured. Trustworthiness, the extent to which the neighbors of a point in the map are neighbors in the data, is 0.98 for t-SNE and 0.96 for UMAP against 0.95 for the first two principal components, so that the embeddings do represent local neighborhoods somewhat more faithfully than a two-dimensional score plot. Their reproducibility is less satisfactory. When the calculation is repeated from different random starting configurations, the sets of ten nearest neighbors in the map overlap by 0.78 (t-SNE) and 0.64 (UMAP) on average, and for UMAP the overlap falls further when larger neighborhoods are used. A fifth to a third of the local structure of such a map is therefore a product of the particular run.

What is stable is the content already known from PCA and clustering. Both maps show a dry, lactate-dominant branch, a dry, malate-dominant group formed largely by Stormalong and Bad Seed ciders, a semi-sweet body, and the sucrose-sweetened and very sweet groups at the periphery; the lots of one product lie together, as do the five sour ciders measured in 2017, which come from one producer. The gaps between the islands, their relative distances and their sizes have no meaning and were not interpreted. In this data set the embeddings are an effective picture of results obtained by other means and add nothing to them.

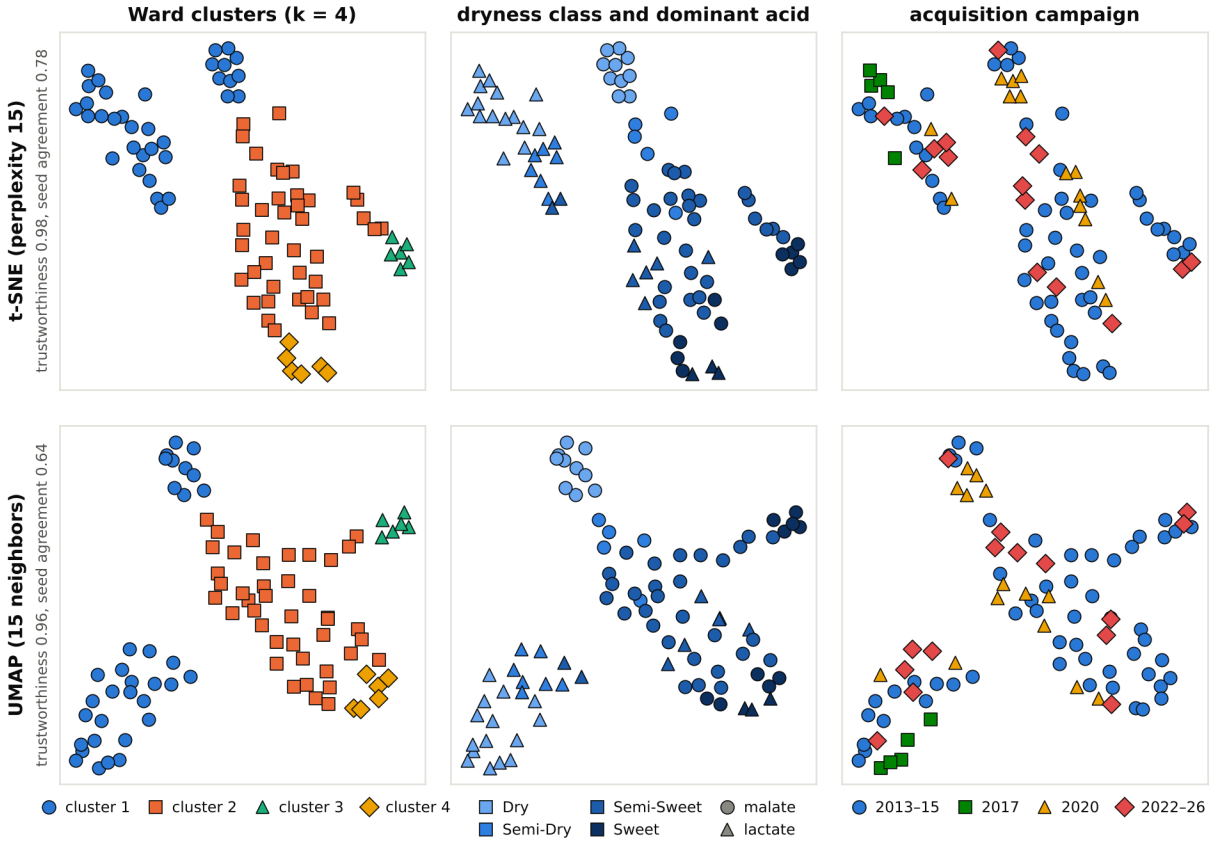

Figure 42. t-SNE (top) and UMAP (bottom) embeddings of the freeze-dried tubes of 87 ciders, calculated from the same principal component scores as the clustering and colored by cluster, by dryness class and dominant acid, and by measurement campaign. Trustworthiness and the agreement of nearest-neighbor sets between runs with different random starts are given at the left.

13. Chemometrics: Classification

13.1 Questions, labels and validation

Supervised classification answers a different question from the preceding section: not whether groups exist, but whether a property assigned to the ciders beforehand can be recognized from the spectrum. Nine properties were examined, and they are of three kinds. Four are defined by the composition itself: whether a cider is dry or not (below or above 18 g/L of sugar), its dryness class on the four-level scale, its dominant acid, and the presence of sucrose at 2 g/L or more. These labels were derived from the same spectra that serve as predictors, so that a successful classification proves only that the bucket table contains the information that the targeted analysis extracted from it. Three labels are external to the measurement: the market segment, as a contrast between mass-market and regional craft ciders and as a three-way division that adds the traditional European ciders; the producer, for the four producers represented by eight or more ciders; and the declaration of an adjunct such as fruit, spice, hops or barrel aging. The ninth label, the measurement campaign, is not a property of the cider at all and was included as a control.

The validation was designed to prevent the two errors that most often inflate reported accuracies.[204,220,223,240] Hyperparameters were chosen in an inner cross-validation, so that the performance estimate is not biased by their selection, and the folds were formed by producer, so that a model is never tested on ciders of a producer from which it has learned. For the producer label itself the folds were formed by product, which keeps the lots of one product together. Performance is expressed as balanced accuracy, the mean of the recalls of the classes, for which chance corresponds to 0.5 with two classes, 0.33 with three and 0.25 with four. Because the best of six classifiers is reported for each label, the significance was assessed by a permutation test that repeats the entire procedure, including the tuning of all six classifiers and the selection of the best, on data whose labels had been shuffled among the producers. When k of the 100 permutations reach the observed value, the p-value is (k + 1)/101, so that its smallest attainable value is 0.01.

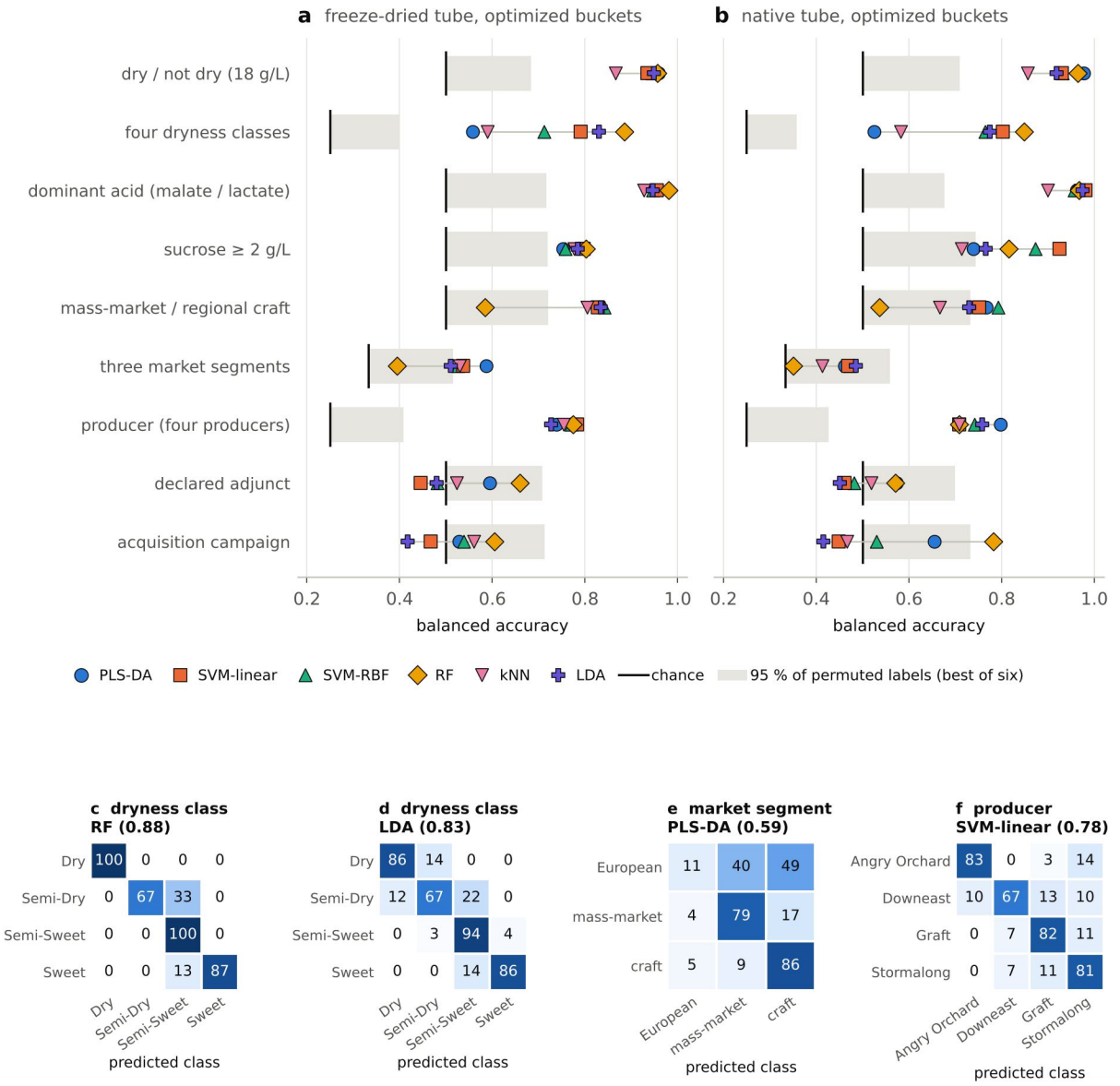

Figure 43. Supervised classification of the ciders from optimized buckets. (a) Freeze-dried tubes and (b) native tubes: balanced accuracy of six classifiers for nine labels in nested cross-validation grouped by producer (by product for the producer label). The vertical bar marks chance; the gray band extends from chance to the 95th percentile of the best-of-six balanced accuracy obtained with permuted labels. (c) to (f) Confusion matrices (percent of each true class, rows) for the dryness class with a random forest and with linear discriminant analysis, for the three market segments and for the four producers.

Table 18. Classification of the ciders from optimized buckets: best balanced accuracy among six classifiers in nested, grouped cross-validation (five repetitions), and the permutation test of the best-of-six statistic (one repetition; 100 permutations of the labels among the cross-validation groups).

Label	Tube	n	Classes	Chance	Best balanced accuracy (model)	Statistic tested	Permuted: mean	95th percentile	p
Dry / not dry (18 g/L)	freeze-dried	87	2	0.50	0.96 (PLS-DA)	0.96	0.57	0.68	0.01
	native	96	2	0.50	0.98 (PLS-DA)	0.97	0.56	0.71	0.01
Four dryness classes	freeze-dried	87	4	0.25	0.88 (RF)	0.88	0.29	0.40	0.01
	native	96	4	0.25	0.85 (RF)	0.81	0.28	0.36	0.01
Dominant acid	freeze-dried	87	2	0.50	0.98 (RF)	0.97	0.58	0.72	0.01
	native	96	2	0.50	0.98 (SVM-linear)	0.97	0.57	0.68	0.01
Sucrose ≥ 2 g/L	freeze-dried	87	2	0.50	0.80 (RF)	0.82	0.56	0.72	0.01
	native	96	2	0.50	0.93 (SVM-linear)	0.92	0.56	0.74	0.01
Mass-market / regional craft	freeze-dried	73	2	0.50	0.84 (SVM-RBF)	0.83	0.57	0.72	0.01
	native	82	2	0.50	0.79 (SVM-RBF)	0.82	0.59	0.73	0.01
Mass-market / regional craft, 2013–2015 only	freeze-dried	41	2	0.50	0.89 (SVM-linear)	0.89	0.57	0.74	0.01
	native	47	2	0.50	0.91 (SVM-RBF)	0.91	0.56	0.72	0.01
Three market segments	freeze-dried	80	3	0.33	0.59 (PLS-DA)	0.57	0.39	0.52	0.03
	native	89	3	0.33	0.49 (LDA)	0.53	0.40	0.56	0.08
Producer (four producers)	freeze-dried	36	4	0.25	0.78 (SVM-linear)	0.84	0.29	0.41	0.01
	native	40	4	0.25	0.80 (PLS-DA)	0.81	0.29	0.43	0.01
Declared adjunct	freeze-dried	87	2	0.50	0.66 (RF)	0.66	0.56	0.71	0.13
	native	96	2	0.50	0.57 (PLS-DA)	0.58	0.57	0.70	0.43
Measurement campaign (2013–2015 / later)	freeze-dried	87	2	0.50	0.60 (RF)	0.61	0.57	0.71	0.29
	native	96	2	0.50	0.78 (RF)	0.79	0.58	0.73	0.02

The statistic tested is the highest balanced accuracy among the six classifiers in a single repetition of the nested cross-validation, which is the quantity that the permutation null reproduces; it differs slightly from the five-repetition mean of the best model. p = 0.01 is the smallest value attainable with 100 permutations.

13.2 Labels defined by the composition

Before the individual results are considered, the null distributions of Table 18 deserve a remark, because they calibrate everything that follows. For a label with two classes, the best of six classifiers reaches a balanced accuracy of 0.56 to 0.59 on average when the labels carry no information at all, and 0.68 to 0.74 in one of twenty such trials. Chance is therefore not 0.5 for the quantity that is usually reported, the performance of the best model tried, and with samples of this number no two-class result below about 0.72 can be told apart from the outcome of choosing among models.[203] The gray bands of Figure 43 mark this limit.

The two-class labels that follow from the composition are recognized almost without error. Whether a cider contains less or more than 18 g/L of sugar is predicted with a balanced accuracy of 0.96 from the freeze-dried and 0.98 from the native buckets, and the dominant acid with 0.98 from either. Every classifier except the nearest-neighbor method exceeds 0.9 for both labels in both tubes; the nearest-neighbor method is distinctly weaker (0.86 to 0.93). The presence of sucrose is recognized less reliably from the freeze-dried buckets (0.80) than from the native ones (0.93). At least part of this difference has a simple cause. In five of

the 21 ciders that carry the positive label, sucrose had been partly inverted in the acidified freeze-dried sample and the reported value was taken from the native tube (Section 5.7); for three of them, the lots of Stormalong Red Skies at Night, the freeze-dried spectrum shows less than the limit of 2 g/L, so that the label no longer describes the spectrum from which it is to be predicted.

The four dryness classes are more instructive, because here the classifiers differ widely (Figure 43c, d). The random forest reaches 0.88 and 0.85, linear discriminant analysis 0.83 and 0.77, and PLS discriminant analysis only 0.56 and 0.53, little more than half of what it achieves for the two-class version of the same question. The explanation lies in the nature of the label and not in the data. The classes are consecutive intervals of one quantity, the sugar content. A decision tree divides its variables at thresholds and is the natural model for such a label, whereas PLS discriminant analysis fits one linear response to each class and assigns the class with the largest response;[220] the response of an interior class, which is low on both sides of an interval and high within it, cannot be approximated by a linear function, and the interior classes are masked by their neighbors. All errors of the random forest are confusions of adjacent classes, and the semi-dry class, with twelve ciders the smallest, is the least well recognized (67 %).

A better model than any of the six is available without further effort. When the limits of the classes are applied to the sugar content predicted by the PLS regression of Section 14, in cross-validation, one of the 96 ciders is assigned to a wrong class from the native buckets and one of 87 from the freeze-dried buckets, a balanced accuracy of 0.98 in both cases. An ordered class that is defined by limits on a measurable quantity should be predicted by regression of that quantity. The present data give no support for the common practice of applying discriminant analysis to labels of this kind.

These results are to be read for what they are. The four labels are functions of concentrations that were themselves determined from the spectra, and a model that reproduces them shows that the bucket table retains the information. The same accuracies are obtained from the table of fitted concentrations (dry or not dry 0.97 and 0.97, dominant acid 0.98 and 0.98), as they must be. The practical content of the result is this: a cider can be assigned to its dryness class and its acid type from a bucket table alone, without line fitting and without a library of reference spectra.

13.3 Labels external to the measurement

The labels of the second kind are the ones on which a classification can produce new information, and the results are correspondingly less uniform. Mass-market and regional craft ciders are distinguished with a balanced accuracy of 0.84 from the freeze-dried and 0.79 from the native buckets. The statistic tested by permutation, 0.83 and 0.82, lies above the 95th percentile of the permuted labels (0.72 and 0.73) and was not reached by any of the 100 permutations ($p = 0.01$). Four of the classifiers agree within a few hundredths; the random forest fails (0.58 and 0.54), because with 19 ciders against 63 it assigns nearly everything to the larger class. The distinction has an evident chemical basis (Section 8): the mass-market ciders of this study contain twice as much residual sugar in the median as the craft ciders, ten of the nineteen are sweetened mainly with sucrose, only one has undergone malolactic conversion, and their content of succinic acid is about half that of the craft ciders.

An objection must be met before this result is accepted. Eighteen of the nineteen mass-market ciders were measured in the first campaign, whereas the craft ciders are spread over all four, and a model could in principle have learned the campaign instead of the product. The classification was therefore repeated with the ciders of the first campaign alone, 18 mass-market against 23 (freeze-dried) or 29 (native) craft ciders. The balanced accuracy was no lower, 0.89 and 0.91, again with $p = 0.01$. The distinction is thus recognized within a single campaign and cannot be an effect of the measurement. Since the folds were formed by producer, each of the eight mass-market brands was judged by a model that had not seen it, which speaks for a compositional character common to these products. The statement is nevertheless one about nineteen

products of eight brands, nearly all of them purchased before 2016, and not about the category as it stands today.

The extension to three classes by the addition of the traditional European ciders was not successful. The balanced accuracy is 0.59 from the freeze-dried buckets ($p = 0.03$) and 0.49 from the native ones ($p = 0.08$), and the European ciders themselves are recognized in about one case in ten; they are assigned to the craft or the mass-market class in nearly equal proportions (Figure 43e). Seven ciders of six producers from three countries do not form a class in the dimensions that dominate the spectra. They range from the completely dry, lactate-dominant ciders of Etienne Dupont and Sarasola to the sweetened, malate-dominant English ciders, one of which contains 15 g/L of sucrose. The property that the targeted analysis found to be common to them, the high content of polyphenols (Section 8.6), occupies a few buckets of low intensity to which Pareto scaling gives little weight. The failure says more about the smallness and heterogeneity of the class than about the method.

The four producers represented by eight or more ciders, Angry Orchard, Downeast, Graft and Stormalong, are recognized with a balanced accuracy of 0.78 and 0.80 against a chance level of 0.25 ($p = 0.01$; Figure 43f). Each has a house style that is plain from the composition: sweetening with sucrose, unfiltered ciders with the sugars of the juice, sour ciders with lactic and acetic acids and 1,3-propanediol, and dry to semi-sweet, malate-dominant ciders of heirloom apples. It must be added that producer and campaign are almost completely confounded in this group. All fourteen Stormalong ciders were measured in 2020, the Graft ciders in 2017 and after 2022, and seven of the eight Angry Orchard ciders before 2016. For the native tubes, which carry a signature of the campaign (Section 13.4), it cannot be decided how much of the recognition is due to the cider; for the freeze-dried tubes, in which no such signature was detectable, the compositional reading is the more probable. In either case a model of four producers has no use beyond these four, since a cider of any other producer would be assigned to one of them.

The declaration of an adjunct is not recognized. The best balanced accuracies, 0.66 and 0.57, lie within the range obtained with permuted labels ($p = 0.13$ and 0.43). This label also shows how much the design of a permutation test matters. When the labels are permuted freely among the samples instead of among the producers, the result for the freeze-dried tubes appears significant ($p = 0.02$), because the null distribution becomes narrower (95th percentile 0.64 instead of 0.71). Adjuncts are a habit of producers: only 14 of the 36 producers are represented in both classes, and a free permutation destroys this association, which the grouped cross-validation respects. It tests against a hypothesis that is too easy to reject. The grouped permutation is the appropriate one, and by it nothing has been shown. The outcome is not surprising. The label covers fruit, hops, spices, honey and barrel aging, which have no spectral feature in common, and several of them contribute nothing that a ¹H spectrum at 300 MHz would register. Individual adjuncts are detected reliably by their own markers, citric and quinic acids for berries and citrus fruit and sorbitol for pear (Sections 8.4 and 8.5), and a targeted marker is the proper tool for this question.

13.4 The measurement campaign as a control

The campaign in which a cider was measured is not a property of the cider. A classifier that predicts it has found either a difference between the products that happened to be measured in different years or a difference in the measurement itself, and in the second case every other model built on the same data is open to doubt. The label was therefore included as a control, in the form of a contrast between the first campaign and the three later ones. From the freeze-dried buckets the campaign is not predicted: the best balanced accuracy is 0.60 ($p = 0.29$). From the native buckets it is. The random forest reaches 0.78, and the value of 0.79 obtained in the repetition used for the permutation test was reached by one of the 100 permutations ($p = 0.02$). PLS discriminant analysis reaches 0.66, and the remaining classifiers do not exceed 0.53.

The origin of this signal was located by removing and isolating parts of the spectrum (Figure 44). When the 24 buckets between the excluded water region and the excluded region of the standard, from 5.12 to 6.05

ppm, are removed from the native table, the random forest falls to 0.63 and PLS discriminant analysis to 0.52. Conversely, the fifteen buckets between 5.45 and 6.05 ppm alone, a range that excludes the anomeric signals of sucrose and glucose, predict the campaign with 0.79, 0.74 and 0.71 (random forest, support vector machine, PLS discriminant analysis), as well as the whole table and now with every classifier. Hardly any constituent of cider resonates in this range. The only assigned signal is the H-2 doublet of sorbic acid, and when the test is restricted to the 72 ciders in which no sorbic acid was found the result is unchanged (0.80, 0.73 and 0.71). What differs is the level of the region itself. Its median is 14 mmol of protons per liter per ppm in these ciders of the first campaign, with an interquartile range from 0.3 to 34, against 4 to 7 in the later campaigns; in the freeze-dried tubes, whose residual water signal is small, the quartiles lie between −0.3 and 1.9 in every campaign.

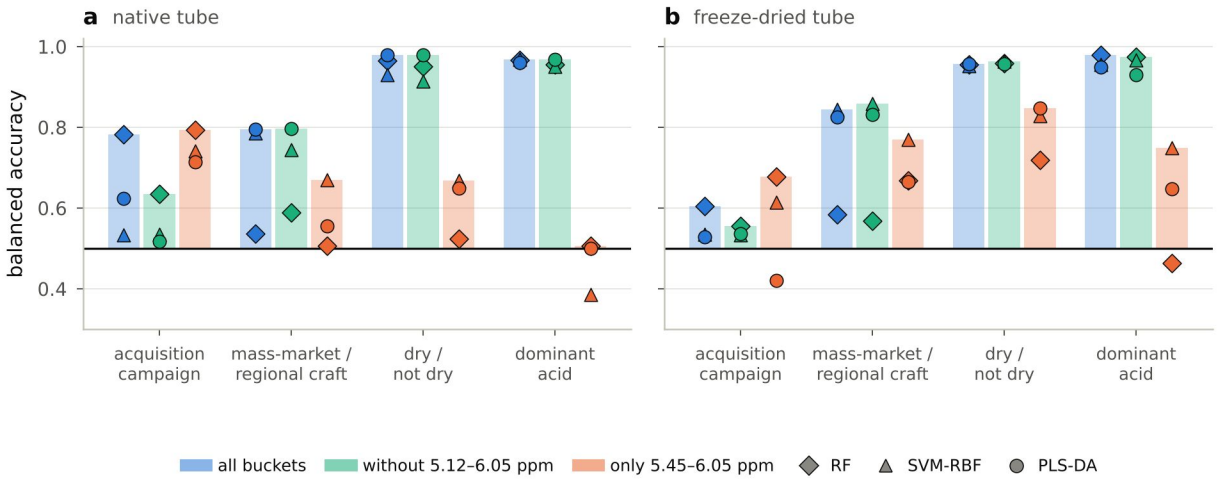

Figure 44. Location of the campaign signature. Balanced accuracy of three classifiers for the measurement campaign and for three properties of the ciders, from all optimized buckets, from the table without the buckets between 5.12 and 6.05 ppm, and from the fifteen buckets between 5.45 and 6.05 ppm alone; (a) native tubes, (b) freeze-dried tubes. Bars extend from chance to the best of the three classifiers.

The signature is therefore the flank of the residual water signal. Its height and shape depend on the efficiency of the water suppression, on the homogeneity of the field, and on the spectral width and filter settings, which differed between the campaigns, as in part did the spectrometer (Table 8). It is an instrumental feature, and it is the same broad contribution that appeared in the third principal component of the native tubes (Section 12.1). A random forest finds it readily because a single bucket with a characteristic level suffices for a tree, while methods that weigh all buckets together are dominated by the far larger variance of ethanol and sugar.

The consequences are limited. The chemical labels do not depend on the region: without it, dry and not dry are distinguished with 0.98, the dominant acid with 0.97 and the two market segments with 0.80, the same values as with it, and the contrast of the market segments has been shown above to hold within one campaign. In the freeze-dried tubes the same fifteen buckets do carry chemical information, since they predict dryness with 0.85, presumably through the wings of the anomeric signals of the concentrated sugars and through sorbic acid, which accompanies residual sugar; there the region is harmless and should be kept. For native tubes it is recommended that the range between the water signal and 6.05 ppm be excluded from bucket tables intended for classification, at the cost of the olefinic signals of sorbic acid, whose methyl doublet remains, and that the prediction of the campaign, or of whatever variable describes the history of the measurement, be made a routine control. A model built on native spectra is verified anew before it is applied to data from another instrument or another set of acquisition parameters.

14. Chemometrics: PLS Regression of Composition on the Spectra

14.1 Purpose and design

The targeted procedure of Section 5.6 yields the composition, but it depends on line fitting, on a library of reference spectra and on a set of rules that must be maintained. A regression model that predicts the same quantities directly from a table of buckets would dispense with all of that, and it could be transferred to routine use more easily. PLS regression[201] was therefore applied to the optimized buckets, with the pre-processing selected in Section 11, to predict the concentrations determined by qNMR. Models were built in three configurations: freeze-dried buckets against the non-volatile constituents measured in the same tubes; native buckets against ethanol, the volatile constituents, the sugars, the acids and the calculated energy; and native buckets against the polyols and minor acids, whose reference values come from the freeze-dried tube. One further model relates the native buckets to the alcohol content declared on the label. The number of latent variables, up to ten, was selected in an inner cross-validation, and the errors quoted are those of the outer, producer-grouped cross-validation (Table 19, Figure 45).

The meaning of these errors must be clear before the numbers are read. In the first two configurations the reference values were derived from the very spectra that serve as predictors. The model is then asked to reproduce a deterministic function of its input, and its error measures how well a linear combination of buckets imitates the targeted procedure, not how accurately the composition is known. Any bias of the reference values is inherited in full. Only the third configuration, in which predictor and reference come from different tubes and different acquisitions, and the model for the declared alcohol content are tests against independent information.

Table 19. PLS regression of the composition on optimized buckets (absolute intensities, mean-centered, Pareto-scaled): nested cross-validation grouped by producer. Concentrations in g/L unless stated otherwise.

Predicted quantity	n	Range	Latent variables	RMSECV	RMSECV (% of mean)	Q^2	RPD	RPD, targeted integrals
Native buckets; reference values from the native tube (non-volatile constituents bridged from the freeze-dried tube)								
Ethanol (% v/v)	96	0.1–10.4	10	0.026	0.5	1.000	59.4	42.7
Total sugar	96	0–102	8	1.2	5.1	0.996	16.6	7.5
Sucrose	96	0.0–42.9	10	0.617	15	0.996	15.4	5.6
Glucose	96	0.0–40.7	7	0.832	11	0.989	9.6	9.7
Fructose	96	0.0–73.5	9	0.905	7.5	0.994	13.2	8.8
Malic acid	96	0.0–7.3	10	0.247	8.8	0.986	8.4	10.7
Lactic acid	96	0.0–9.2	10	0.287	19	0.980	7.2	7.6
Succinic acid	96	0.0–1.5	10	0.071	15	0.916	3.5	3.5
Acetic acid	96	0.0–4.9	10	0.094	20	0.987	9.0	10.9
Total organic acids	96	1.9–13.9	10	0.584	10	0.927	3.7	2.8
Energy (kcal per 12 fl oz)	96	35–247	8	3.1	1.9	0.990	9.9	6.7
Methanol (mg/L)	96	0–233	10	25.9	49	0.798	2.2	3.8
Isoamyl alcohols (mg/L)	96	0–232	8	43.2	36	0.294	1.2	1.2
Sorbic acid (mg/L)	96	0–253	10	65.1	182	0.082	1.0	1.9
Freeze-dried buckets; reference values from the same tubes								
Total sugar	87	0–102	9	1.2	5.3	0.996	16.4	13.9
Sucrose	87	0.0–42.9	6	1.2	29	0.984	7.9	6.9
Glucose	87	0.0–40.7	5	1.1	14	0.983	7.7	8.0

Predicted quantity	n	Range	Latent variables	RMSECV	RMSECV (% of mean)	Q²	RPD	RPD, targeted integrals
Fructose	87	0.0–73.5	6	0.991	8.5	0.993	11.7	10.4
Malic acid	87	0.0–7.3	10	0.153	5.5	0.995	14.0	16.9
Lactic acid	87	0.0–9.2	10	0.103	6.6	0.998	20.7	40.3
Succinic acid	87	0.0–1.5	10	0.043	9.4	0.971	5.9	6.6
Citric acid	87	0.0–3.3	10	0.163	85	0.914	3.4	2.5
Quinic acid	87	0.0–3.2	10	0.155	50	0.821	2.4	2.0
Total organic acids	87	1.9–13.9	10	0.606	10	0.927	3.7	3.1
Glycerol	87	0.0–6.2	10	0.538	18	0.743	2.0	2.8
Sorbitol	87	0.0–22.6	10	0.312	7.6	0.987	8.8	2.2
1,3-Propanediol	87	0.0–2.5	9	0.116	63	0.953	4.7	5.0
Total polyols	87	1.8–25.8	10	0.682	9.3	0.950	4.5	3.0
Native buckets; reference values from the freeze-dried tube (independent preparation and acquisition)								
Total polyols	87	1.8–25.8	10	0.570	7.8	0.965	5.4	–
Sorbitol	87	0.0–22.6	10	0.569	14	0.956	4.8	–
Glycerol	87	0.0–6.2	9	0.630	21	0.648	1.7	–
1,3-Propanediol	87	0.0–2.5	10	0.176	96	0.892	3.1	–
Citric acid	87	0.0–3.3	10	0.328	171	0.653	1.7	–
Quinic acid	87	0.0–3.2	10	0.354	114	0.060	1.0	–
Native buckets; reference value from the label								
Declared ethanol (% v/v)	71	4.2–11.0	1	0.672	11	0.662	1.7	–

RMSECV, root mean square error of cross-validation; RPD, ratio of the standard deviation of the reference values to RMSECV. The last column gives the RPD of the corresponding model built on the table of targeted integrals.

14.2 Results

The principal constituents are reproduced with errors that are small against their range. Ethanol is predicted from the native buckets with a cross-validated error of 0.026 % v/v, total sugar with 1.2 g/L, sucrose with 0.62 g/L and the energy of a serving with 3.1 kcal; from the freeze-dried buckets, malic and lactic acids are predicted with 0.15 and 0.10 g/L and sorbitol with 0.31 g/L. The ratio of performance to deviation exceeds 7 for all of these, the range in which a model is conventionally regarded as fit for quantitative use.[205] A second group, comprising succinic and acetic acids, the sums of acids and of polyols, 1,3-propanediol and citric acid, is predicted with ratios between 3 and 9, adequate for screening. For constituents whose signals are weak or buried the models are of little or no quantitative use: glycerol (Q^2 0.74, ratio 2.0), quinic acid (0.82, 2.4) and methanol (0.80, 2.2) are predicted only roughly, and the isoamyl alcohols (Q^2 0.29) and sorbic acid (0.08) not at all. A bucket 0.04 ppm wide that contains the methyl signal of a compound at 100 mg/L also contains the flank of a signal a hundred times larger, and Pareto scaling does not give such a bucket enough weight.

The comparison with the table of targeted integrals is instructive in both directions (last column of Table 19). Where a constituent has a window of its own, the integral table is equal or better; this holds for lactic and malic acids, and markedly for methanol and sorbic acid, whose small signals are isolated by the window. Where the integral table has no variable specific to a constituent, the buckets are far superior: sorbitol, which is hidden under the sugars and has no window, is predicted from buckets with a ratio of 8.8 and from integrals with 2.2, and sucrose and total sugar in the native tube behave in the same way. The

untargeted representation thus recovers information from overlapped regions that fixed windows discard, while the targeted one protects small signals from large neighbors. Neither is uniformly better.

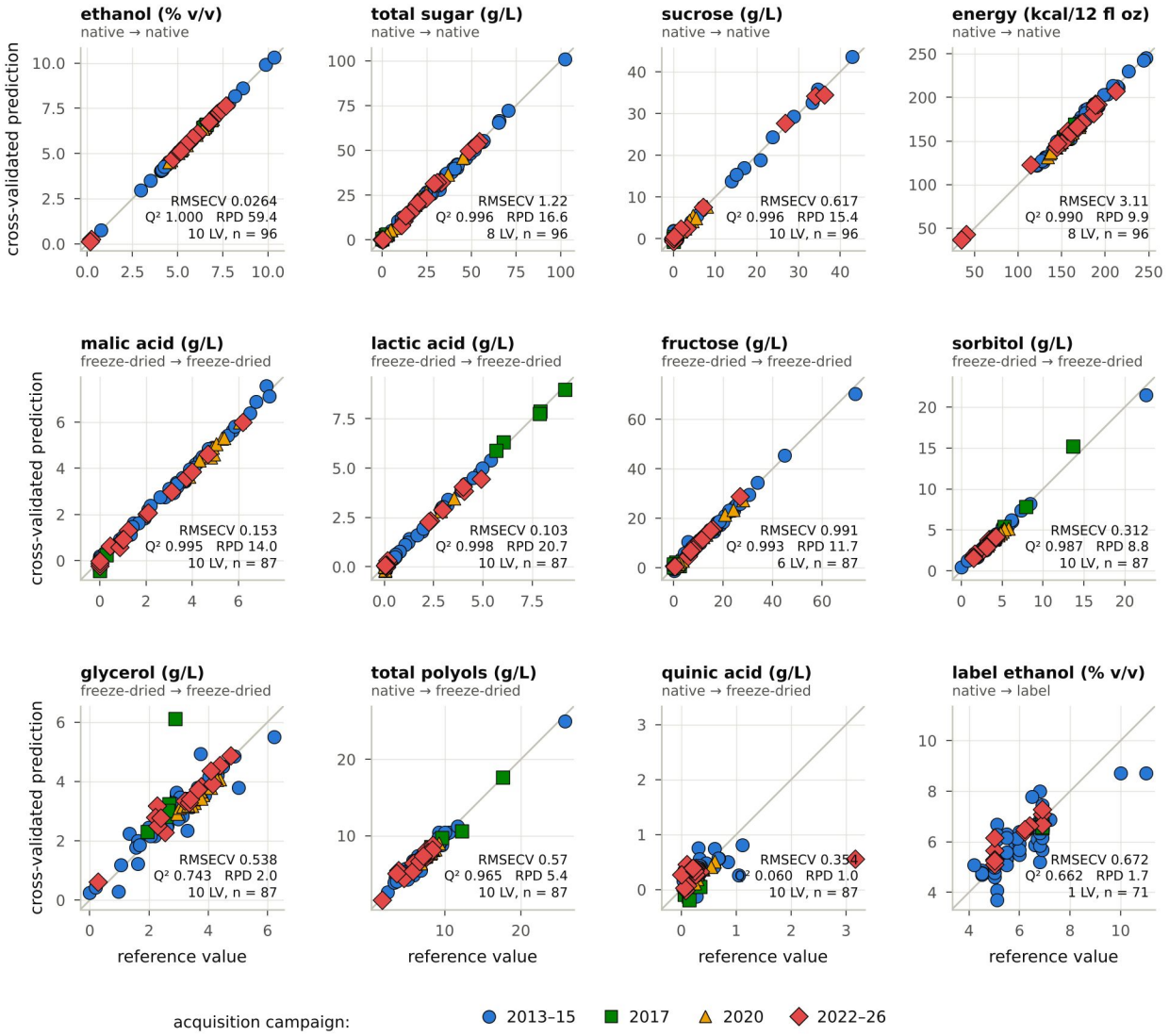

Figure 45. PLS regression of composition on optimized buckets: cross-validated predictions against the reference values for twelve models. The first line of each title gives the predicted quantity, the second the tube that supplies the buckets and the source of the reference values. Marker shape and color indicate the measurement campaign.

The cross-tube models are the genuine test, and they show what can be obtained from the native spectrum alone. The sum of the polyols is predicted with an error of 0.57 g/L (ratio 5.4), sorbitol with 0.57 g/L (4.8) and 1,3-propanediol with 0.18 g/L (3.1). Glycerol by itself is predicted poorly (0.63 g/L, ratio 1.7), citric acid no better, and quinic acid not at all (Q^2 0.06). The native spectrum therefore contains usable information on the total polyol content even though ethanol covers the polyol signals, presumably through the sorbitol signals between the ethanol multiplets and through the correlation of glycerol with ethanol. These models were applied to the nine ciders for which no freeze-dried tube exists (Table 20); the predicted polyol sums of 4 to 10 g/L would add 5 to 11 kcal per serving to the energy values reported for these ciders in Section 10.

Table 20. Polyols predicted from the native spectrum by PLS regression for the nine ciders without a freeze-dried tube (g/L), and the effect on the energy of a serving.

Sample	Cider	Ethanol (% v/v)	Sugar (g/L)	Total polyols, predicted	Glycerol, predicted	Sorbitol, predicted	1,3-Propanediol, predicted	Energy without polyols (kcal)	Energy with predicted polyols (kcal)
C026	Hudson Valley Farmhouse Cider Bourbon Barrel Aged	8.6	0.0	9.7	4.3	5.3	0.06	177	188
C032	Nine Pin Cider Works Signature NY Hard Cider	6.2	27.9	7.1	3.9	3.3	0.00	168	177
C038	Downeast Original Blend, overpressurized can	6.8	32.7	7.6	3.4	4.2	0.00	187	196
C039	Downeast Original Blend, sound can	6.0	51.8	7.6	3.2	4.4	0.00	198	207
C040	Downeast Original Blend, Batch Y	6.2	56.9	7.8	3.2	4.6	0.01	209	218
C043	Awestruck Eastern Dry	7.2	22.2	7.9	4.2	3.8	0.00	180	189
C083	Angry Orchard Peach Mango	4.7	33.3	4.4	2.4	1.9	0.02	145	150
C085	Locust Grove Honey Crisp Cider (4-3-24)	5.8	24.8	6.3	3.7	2.5	0.13	156	164
C086	Locust Grove Peach Cider (4-3-24)	5.5	0.0	5.1	3.4	1.8	0.00	115	121

Cross-validated errors of the models: total polyols 0.57 g/L, glycerol 0.63 g/L, sorbitol 0.57 g/L, 1,3-propanediol 0.18 g/L. Energy per 12 fl oz on the component basis; the last column adds glycerol and sorbitol as predicted individually.

The model for the declared alcohol content uses a single latent variable and has an error of 0.67 % v/v (Q^2 0.66). This is the same as the standard deviation of the differences between qNMR and label found in Section 7.2. No model can predict the declarations better than the declarations describe the contents, and the result confirms that the scatter resides in the labels.

14.3 What the models have learned

The regression coefficients show that the models are chemically sensible (Figure 46). In each case the large coefficients lie on the signals of the constituent concerned: the methylene multiplet for malic acid, the methyl doublet and the methine quartet for lactic acid, the bucket at 3.84 ppm for sorbitol, the high-field methylene signals for glycerol, the two multiplets for ethanol, and the anomeric doublet together with the fructosyl signals for sucrose. The model for malic acid carries, in addition, negative coefficients on the downfield half of the citric acid quartet. It has learned the correction for the overlap of citric acid that the targeted procedure applies by rule. Coefficients of opposite sign on neighboring sugar buckets in the models for sorbitol, glycerol and sucrose serve the same purpose, the subtraction of overlapping contributions.



Figure 46. Regression coefficients of six PLS models against the centers of the buckets (5.8 to 0.7 ppm), with the assignment of the largest coefficients. Blue, positive; orange, negative. Each panel is scaled to its largest coefficient.

Two features of the models are to be kept in mind when they are used. The number of latent variables selected reached the permitted maximum of ten in 24 of the 36 bucket models, and the error was still falling at that point (Figure 38d). A regression against independent reference values would normally reach its optimum with fewer components and would then deteriorate; continued improvement is what is expected when the reference is an exact function of the predictors, and the errors of the first two configurations are to be read accordingly. Secondly, the models were validated by cross-validation grouped by producer, which protects against the recognition of a producer. Their transfer to another instrument, to other acquisition parameters or to another sample preparation calls for a verification with samples of known composition, as

for any calibration. The predictions show no visible offset between the four campaigns (Figure 45), although these differ in acquisition parameters and, in part, in the spectrometer.

The practical conclusion is as follows. For the quantities that matter for dryness designation and nutrition labeling, namely alcohol, total sugar, the principal acids and energy, a PLS model on a bucket table is an adequate substitute for the targeted analysis once that analysis has provided the calibration. It is not a substitute for minor constituents, for additives, or for the recognition of anything unexpected, and it cannot improve on the accuracy of the values with which it was trained.

15. Conclusions

One ¹H NMR measurement with an internal standard, supplemented by a second measurement of the freeze-dried sample, determines the alcohol, the three sugars, the principal organic acids, glycerol, sorbitol and a series of minor constituents and markers of a cider. Applied uniformly to the spectra accumulated since 2013, with each cider included once, the procedure gives a consistent description of 96 ciders from 36 producers and of the juices of 56 apple varieties.

The juices of the old apple varieties differ from one another chiefly in the share of sucrose among their sugars, which extends from 5 to 52 %, in their malic acid, 3.4 to 10.4 g per 100 g of sugars, and in their sorbitol, which varies more than tenfold and is highest in the russets. Of the 53 lots, 48 would be classed as sharps by the cider maker, and the supermarket varieties measured with them lie outside the range of the old varieties in their ratio of sugar to acid. Citramalic acid and L-rhamnitol were identified as further markers of variety, and asparagine, the principal source of nitrogen for the yeast, is measured in the same spectrum. In the multivariate analysis the juices form a continuum governed by the sucrose share, by glucose and sorbitol, and by malic acid, without natural clusters. PLS regression on the bucketed spectra reproduces the sugars, sorbitol and malic acid with ratios of performance to deviation of 10 to 28, and the best pre-processing depends on the size of the signal: mean-centering alone for the principal constituents, Pareto or VAST scaling for the minor ones.

The ciders differ chiefly in two respects that are independent of each other. The first is residual sugar, from nil to 102 g/L, and its kind: the sugars of the juice in most craft ciders, added sucrose in most mass-market ciders. The second is the microbiological history recorded in the acids and polyols. Malolactic conversion has taken place in 36 of the ciders, almost never in mass-market products. In nine ciders glycerol has been converted to 1,3-propanediol by lactic acid bacteria, the signature of the natural ciders of the Basque Country, which was found in a Basque cider and in eight American ciders made by spontaneous or mixed fermentation; the molar balance of glycerol and the parallel rise of acetic acid agree with the known metabolism of these bacteria, and the quintet of 1,3-propanediol at 1.78 ppm reveals the conversion at an early stage. Two lots of products that are otherwise malate-dominant had undergone complete malolactic conversion, and one can had refermented. Sorbic acid, 5-hydroxymethylfurfural, citric and quinic acids and sorbitol indicate preservation, heat treatment, added fruit and pear respectively. The traditional European ciders are distinguished by their polyphenol signals, the property in which ciders made from dessert and heirloom apples differ most from them.

On the scale of the American Cider Association, 27 ciders are dry, 12 semi-dry, 42 semi-sweet and 15 sweet, and the word dry on a label does not predict the class. The same sugar content carries different designations in the United States, in Spain and in France, so that the measured value is the only unambiguous statement. The orchard-based scale, which weighs sugar against acid, separates the dry ciders usefully but has no resolving power for the sweeter commercial ciders. Seven further scales were derived from the same spectrum, for balance, acidity, acid profile, body, aromatic character, alcohol tier and energy; the balance index, a sweetness equivalent divided by the acidity, resolves the whole range of the market in five classes (12 bone dry, 10 dry, 11 off-dry, 27 medium and 27 sweet among the ciders for which the polyols were measured).

The energy of a 12 fl oz serving of the alcoholic ciders ranges from 115 to 247 kcal. Polyols and organic acids contribute about a tenth of the energy of a dry cider and about 5 g of carbohydrate by difference per serving, which is more than some labels declare in total. Alcohol, carbohydrate and energy were ranked for all 96 products by brand, and complete nutrition panels, including the entries of the proposed Alcohol Facts statement, were calculated from single spectra.

For multivariate analysis the spectra of ciders are best kept on the absolute scale that the internal standard provides, mean-centered and Pareto-scaled; whether uniform bins or optimized buckets are used is

immaterial for data referenced as these were. The ciders form a continuum without natural clusters. Dryness, dominant acid and the presence of sucrose are recognized from the buckets almost without error, which confirms the information content of the bucket table; the distinction between mass-market and craft ciders is recognized with a balanced accuracy above 0.8, also within a single campaign; and declared adjuncts are recognized by their individual markers and not by a common model. PLS regression on the buckets reproduces alcohol, sugar, the principal acids and energy well enough to replace the targeted analysis for those quantities, and it estimates the total polyols from the native spectrum alone with an error of about 0.6 g/L.

Taken together, the results show what a single ¹H NMR measurement, evaluated quantitatively and supported by multivariate models, offers the cider maker: the composition of the fruit and of the product, the record of the fermentation, the designation of dryness in the vocabulary of any market, and the nutrition values of the label.

Appendix A. Ciders Included in the Study

Table 21. The 96 ciders: identification, origin, market segment (Craft, regional craft; European, traditional European; Home, home or experimental cider maker; Juice, fresh unfermented juice), declared adjunct, measurement campaign, declared alcohol content (% v/v), tubes measured (N, native; FD, freeze-dried) and laboratory code of the spectra.

ID	Producer	Product	Origin	Segment	Declared adjunct	Campaign	Declared alcohol	Tubes	Laboratory code
C001	Home cidermaker (Tim)	Tim's Cider #1	Hudson Valley, NY	Home	–	2013–15	–	N+FD	Cider-001
C002	Home cidermaker (Tim)	Tim's Cider #2	Hudson Valley, NY	Home	–	2013–15	–	N+FD	Cider-002
C003	J.K.'s Scrumpy	Farmhouse Hard Cider (USDA Organic)	Michigan	Craft	–	2013–15	6.0	N+FD	Cider-003 (2013)
C004	Harpoon	Craft Cider	Massachusetts / Vermont	Mass-market	–	2013–15	4.8	N+FD	HarpoonCraftCider
C005	Stella Artois	Cidre	International brand	Mass-market	–	2013–15	4.5	N+FD	StellaArtois-001
C006	Home cidermaker (Tim)	Tim's Cider #3 (Peach-Honey)	Hudson Valley, NY	Home	peach, honey	2013–15	–	N+FD	Cider-003 (2015)
C007	Etienne Dupont	Cidre Triple	Normandy, France	European	–	2013–15	11.0	N+FD	Cider-008
C008	Crispin	Hard Apple Cider	California	Mass-market	–	2013–15	5.8	N+FD	Cider-009
C009	Strongbow	Gold Apple	International brand	Mass-market	–	2013–15	5.0	N+FD	Cider-010
C010	Woodchuck	Out on a Limb - Splitter Splinter (bourbon barrel aged)	Vermont	Mass-market	bourbon barrel	2013–15	5.5	N+FD	Cider-011
C011	Thatchers	Green Goblin Oak Aged English Cider	Somerset, England	European	–	2013–15	6.6	N+FD	Cider-012
C012	Magners	Pear Cider	Ireland	Mass-market	pear	2013–15	4.5	N+FD	Cider-013
C013	Etienne Dupont	Cidre Bouché Brut de Normandie	Normandy, France	European	–	2013–15	5.5	N+FD	Cider-014
C014	Naked Flock	Hard Cider (Belgian ale yeast, maple syrup)	Hudson Valley, NY	Craft	maple syrup	2013–15	6.8	N+FD	Cider-015
C015	Doc's Draft	Original Hard Apple Cider	Hudson Valley, NY	Craft	–	2013–15	5.0	N+FD	Cider-016
C016	Woodchuck	Amber	Vermont	Mass-market	–	2013–15	5.0	N+FD	Cider-017
C017	Blackthorn	English Cider	Somerset, England	European	–	2013–15	6.0	N+FD	Cider-018
C018	Woodchuck	802	Vermont	Mass-market	–	2013–15	5.0	N+FD	Cider-019
C019	ACE	California Apple Hard Cider	California	Mass-market	–	2013–15	5.0	N+FD	Cider-020
C020	Angry Orchard	Traditional Dry	New York / Ohio	Mass-market	–	2013–15	5.5	N+FD	Cider-021
C021	Woodchuck	Gumption	Vermont	Mass-market	–	2013–15	5.5	N+FD	Cider-022
C022	Doc's Draft	Gold Rush	Hudson Valley, NY	Craft	–	2013–15	7.2	N+FD	Cider-023
C023	Sarasola	Original Basque Cider (Sagardoa)	Basque Country, Spain	European	–	2013–15	6.0	N+FD	Cider-026
C024	Pitchfork	Hard Cider	Hudson Valley, NY	Craft	–	2013–15	6.9	N+FD	Cider-027

ID	Producer	Product	Origin	Segment	Declared adjunct	Campaign	Declared alcohol	Tubes	Laboratory code
C025	Celt	Cidre Breton Traditionnel	Brittany, France	European	–	2013–15	4.5	N+FD	Cider-028
C026	Hudson Valley Farmhouse Cider	Bourbon Barrel Aged	Hudson Valley, NY	Craft	bourbon barrel	2013–15	6.8	N	Cider-029
C027	Angry Orchard	Summer Honey	New York / Ohio	Mass-market	honey	2013–15	5.0	N+FD	Cider-030
C028	Angry Orchard	Green Apple	New York / Ohio	Mass-market	–	2013–15	5.0	N+FD	Cider-031
C029	Hudson Valley Farmhouse Cider	Maeve's Hudson Valley Cider	Hudson Valley, NY	Craft	–	2013–15	6.5	N+FD	Cider-032
C030	Nine Pin Cider Works	Belgian	Albany, NY	Craft	–	2013–15	6.3	N+FD	Cider-033
C031	Magners	Original Irish Cider	Ireland	Mass-market	–	2013–15	4.5	N+FD	Cider-034
C032	Nine Pin Cider Works	Signature NY Hard Cider	Albany, NY	Craft	–	2013–15	6.7	N	Cider-035
C033	Ironwood	Hard Cider	Niagara	Craft	–	2013–15	6.0	N+FD	Cider-036
C034	Samuel Smith's	Organic Cider	Yorkshire, England	European	–	2013–15	5.0	N+FD	Cider-037
C035	Bad Seed	Hard Apple Cider (Dry)	Hudson Valley, NY	Craft	–	2013–15	6.9	N+FD	Cider-038
C036	Angry Orchard	Crisp Apple	New York / Ohio	Mass-market	–	2013–15	5.0	N+FD	Cider-039
C037	Bad Seed	The Farmer Saison	Hudson Valley, NY	Craft	–	2013–15	6.9	N+FD	Cider-040
C038	Downeast	Original Blend - Batch X, overpressurized can	Massachusetts	Craft	–	2013–15	5.1	N	Cider-041 (overpressurized can)
C039	Downeast	Original Blend - Batch X, sound can	Massachusetts	Craft	–	2013–15	5.1	N	Cider-041 (good can)
C040	Downeast	Original Blend - Batch Y	Massachusetts	Craft	–	2013–15	5.1	N	Cider-042
C041	Downeast	Hard Honey Cider	Massachusetts	Craft	honey	2013–15	5.1	N+FD	Cider-043
C042	Downeast	Cranberry Blend	Massachusetts	Craft	cranberry	2013–15	5.1	N+FD	Cider-044
C043	Awestruck	Eastern Dry	New York	Craft	–	2013–15	6.8	N	Cider-045
C044	Stew Leonard's	Fresh apple cider, unpasteurized, held 2 months	Connecticut / New York	Juice	–	2013–15	–	N+FD	Cider-046
C045	Naked Flock	Lemon and Ginger	Hudson Valley, NY	Craft	lemon, ginger	2013–15	6.8	N+FD	Cider-047
C046	Citizen Cider	The Dirty Mayor (ginger)	Vermont	Craft	ginger	2013–15	6.9	N+FD	Cider-048
C047	Awestruck	Hibiscus and Ginger ("Pink")	New York	Craft	hibiscus, ginger	2013–15	6.8	N+FD	Cider-049
C048	Downeast	Original Blend (lot exp. 9-28-15)	Massachusetts	Craft	–	2013–15	5.1	N+FD	Cider-050
C049	Citizen Cider	The Full Nelson (dry hopped)	Vermont	Craft	hops	2013–15	6.9	N+FD	Cider-051
C050	Citizen Cider	Unified Press	Vermont	Craft	–	2013–15	6.8	N+FD	Cider-052
C051	Ryan Burk (cidermaker)	Baldwin - Crispin - Rhode Island Greening	New York	Home	–	2013–15	–	N+FD	Cider-053

ID	Producer	Product	Origin	Segment	Declared adjunct	Campaign	Declared alcohol	Tubes	Laboratory code
C052	Angry Orchard	Strawman	New York / Ohio	Mass-market	–	2013–15	10.0	N+FD	Cider-054
C053	Angry Orchard	Stone Dry	New York / Ohio	Mass-market	–	2013–15	5.5	N+FD	Cider-055
C054	E.Z. Orchards	Roman Beauty (heirloom)	Oregon	Craft	–	2013–15	4.2	N+FD	Cider-056
C055	E.Z. Orchards	Hawk Haus (Jonathan and Yarlington Mill)	Oregon	Craft	–	2013–15	6.6	N+FD	Cider-057
C056	E.Z. Orchards	Poire (Forelle / Comice / Bosc perry)	Oregon	Craft	pear (perry)	2013–15	5.9	N+FD	Cider-058
C057	Bad Seed	Bourbon Barrel Reserve	Hudson Valley, NY	Craft	bourbon barrel	2013–15	6.5	N+FD	Cider-059
C058	1911	Original	New York	Craft	–	2013–15	5.5	N+FD	Cider-060
C059	Angry Orchard	Apple Ginger	New York / Ohio	Mass-market	ginger	2013–15	5.0	N+FD	Cider-061
C060	Graft	Farm Flor (rustic, wild fermented)	Hudson Valley, NY	Craft	–	2017	6.9	N+FD	Graft-001
C061	Graft	Desert Diamonds Gose Cider	Hudson Valley, NY	Craft	tamarind, ginger, sea salt	2017	6.9	N+FD	Graft-002
C062	Graft	Book of Nomad Gose Cider	Hudson Valley, NY	Craft	gose adjuncts	2017	6.9	N+FD	Graft-003
C063	Graft	Where Secrets Lie Gose Cider	Hudson Valley, NY	Craft	gose adjuncts	2017	6.9	N+FD	Graft-004
C064	Graft	Wild Woods - Grim Wilderness Vol. 1 (oak aged peach)	Hudson Valley, NY	Craft	peach, oak	2017	6.9	N+FD	Graft-005
C065	Stormalong	Legendary Dry (lot A)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-001
C066	Stormalong	Legendary Dry (lot B)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-002
C067	Stormalong	Legendary Dry (lot C)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-003
C068	Stormalong	Mass Appeal (lot A)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-004
C069	Stormalong	Mass Appeal (lot B)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-005
C070	Stormalong	Mass Appeal (lot C)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-006
C071	Stormalong	Red Skies at Night (lot A)	Massachusetts	Craft	passion fruit, hibiscus	2020	–	N+FD	StormalongCider-007
C072	Stormalong	Red Skies at Night (lot B)	Massachusetts	Craft	passion fruit, hibiscus	2020	–	N+FD	StormalongCider-008
C073	Stormalong	Red Skies at Night (lot C)	Massachusetts	Craft	passion fruit, hibiscus	2020	–	N+FD	StormalongCider-009
C074	Stormalong	Light of the Sun (lot A)	Massachusetts	Craft	hops (Citra, Mosaic), citrus zest	2020	–	N+FD	StormalongCider-010
C075	Stormalong	Light of the Sun (lot B)	Massachusetts	Craft	hops (Citra, Mosaic), citrus zest	2020	–	N+FD	StormalongCider-011

ID	Producer	Product	Origin	Segment	Declared adjunct	Campaign	Declared alcohol	Tubes	Laboratory code
C076	Stormalong	Light of the Sun (lot C)	Massachusetts	Craft	hops (Citra, Mosaic), citrus zest	2020	–	N+FD	StormalongCider-012
C077	Stormalong	Boston Heirloom (lot A)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-013
C078	Stormalong	Boston Heirloom (lot B)	Massachusetts	Craft	–	2020	–	N+FD	StormalongCider-014
C079	Hudson North Cider Co.	Peach Mango Dry Cider	Hudson Valley, NY	Craft	peach, mango	2022–26	5.0	N+FD	HudsonNorthCiderCo-009
C080	Downeast	Cranberry Cider	Massachusetts	Craft	cranberry	2022–26	5.0	N+FD	DownEastCider-010
C081	Bad Seed	Opus None (aged on Cabernet skins, oak barrel)	Hudson Valley, NY	Craft	grape skins, oak	2022–26	–	N+FD	BadSeedCider-010
C082	Graft	Back Country (sour cider, citrus and bourbon)	Hudson Valley, NY	Craft	citrus, bourbon	2022–26	6.9	N+FD	GraftCider-009
C083	Angry Orchard	Peach Mango	New York / Ohio	Mass-market	peach, mango	2022–26	5.0	N	AngryOrchard-008
C084	Downeast	Blood Orange	Massachusetts	Craft	blood orange	2022–26	5.0	N+FD	DownEastCider-015
C085	Locust Grove	Honey Crisp Cider (4-3-24)	Hudson Valley, NY	Craft	–	2022–26	–	N	LocustGrove-028
C086	Locust Grove	Peach Cider (4-3-24)	Hudson Valley, NY	Craft	peach	2022–26	–	N	LocustGrove-029
C087	Downeast	Guava Passion Fruit	Massachusetts	Craft	guava, passion fruit	2022–26	5.0	N+FD	DownEastCider-016
C088	1911	Original Cider - Alcohol Removed	New York	Non-alcoholic	–	2022–26	0.5	N+FD	Beer-485
C089	Original Sin	White Widow non-alcoholic cider (blackberry, apple cider vinegar)	New York	Non-alcoholic	blackberry, cider vinegar	2022–26	0.5	N+FD	Beer-486
C090	Graft	Native (spontaneous natural cider)	Hudson Valley, NY	Craft	–	2022–26	6.9	N+FD	GraftCider-011
C091	Graft	Book of Nomad - Empress's Daughter (strawberry sake cider)	Hudson Valley, NY	Craft	strawberry, sake	2022–26	6.9	N+FD	Graft_Cider-009
C092	Graft	Book of Nomad - Crystal Temple (Cabernet Franc, black currant)	Hudson Valley, NY	Craft	Cabernet Franc, black currant	2022–26	6.9	N+FD	Graft_Cider-010
C093	Barton Farms	Base cider	New York	Craft	–	2022–26	–	N+FD	Barton-001
C094	Barton Farms	Plow (raspberry)	New York	Craft	raspberry	2022–26	6.2	N+FD	Barton-002
C095	Barton Farms	Tractor	New York	Craft	–	2022–26	6.4	N+FD	Barton-003
C096	Barton Farms	Hayride (black cherry)	New York	Craft	black cherry	2022–26	6.2	N+FD	Barton-004

Appendix B. Composition of the Ciders by ¹H qNMR

Table 22. Alcohol (% v/v), sugars, organic acids and polyols (g/L). nd, below the reporting limit; –, not determined (no freeze-dried tube).

ID	Cider	Ethanol	Fructose	Glucose	Sucrose	Sugar, total	Malic	Lactic	Succinic	Citric	Quinic	Acetic	Acids, total	Glycerol	Sorbitol	1,3-PDO
C001	Home cider maker Tim's Cider #1	6.74	0.18	nd	nd	0.2	nd	4.81	0.43	nd	0.05	1.61	6.9	nd	nd	2.49
C002	Home cider maker Tim's Cider #2	7.85	0.13	nd	nd	0.1	nd	5.40	0.47	nd	0.05	2.34	8.3	1.05	3.72	2.01
C003	J.K.'s Scrumpy Farmhouse Hard Cider (USDA Org…	5.81	44.8	20.3	0.39	65.5	3.25	2.35	0.58	nd	0.29	1.17	7.6	3.73	5.54	nd
C004	Harpoon Craft Cider	4.06	18.1	1.88	nd	20.0	4.90	0.09	0.24	nd	0.32	0.03	5.6	2.25	3.66	nd
C005	Stella Artois Cidre	4.17	19.0	15.7	0.66	35.4	3.68	0.16	0.08	0.32	0.15	0.07	4.5	1.33	2.17	nd
C006	Home cider maker Tim's Cider #3 (Peach-Honey)	6.10	4.51	4.50	nd	9.0	nd	3.90	0.33	nd	0.31	0.56	5.1	3.51	4.59	nd
C007	Etienne Dupont Cidre Triple	9.88	0.33	0.06	nd	0.4	0.74	3.15	1.01	nd	1.11	0.19	6.2	4.88	5.26	nd
C008	Crispin Hard Apple Cider	5.95	9.59	14.7	0.86	25.2	1.46	1.28	0.28	nd	0.04	0.65	3.7	3.13	2.80	nd
C009	Strongbow Gold Apple	5.04	24.2	40.7	0.34	65.3	2.83	0.57	0.13	nd	0.21	0.11	3.9	2.08	2.96	nd
C010	Woodchuck Out on a Limb - Splitter Splinter (…	5.03	10.6	6.90	13.9	31.3	3.37	0.09	0.25	nd	0.31	0.08	4.1	2.23	3.09	nd
C011	Thatchers Green Goblin Oak Aged English Cider	5.82	10.5	19.7	nd	30.2	0.87	1.20	0.36	nd	0.29	0.07	2.8	3.41	2.75	nd
C012	Magners Pear Cider	4.50	22.2	24.1	0.74	47.0	0.33	0.08	0.18	2.29	0.04	0.10	3.0	1.63	6.24	nd
C013	Etienne Dupont Cidre Bouché Brut de Normandie	6.22	12.9	0.52	nd	13.4	2.02	2.91	0.66	nd	0.84	0.31	6.7	3.31	7.30	nd
C014	Naked Flock Hard Cider (Belgian ale yeast, ma…	6.98	24.5	15.1	nd	39.7	5.75	1.06	0.50	nd	0.33	0.29	7.9	3.63	5.53	nd
C015	Doc's Draft Original Hard Apple Cider	4.84	20.7	10.3	nd	31.1	1.93	2.86	0.40	nd	0.24	0.14	5.6	3.06	4.14	nd
C016	Woodchuck Amber	4.39	22.7	33.5	nd	56.3	3.14	0.11	0.13	nd	0.13	0.10	3.6	2.62	2.50	nd
C017	Blackthorn English Cider	5.79	6.93	14.9	0.72	22.6	3.33	0.34	0.27	0.12	0.31	0.20	4.6	3.01	3.11	nd
C018	Woodchuck 802	4.37	18.7	23.3	nd	42.0	3.18	0.08	0.17	nd	0.13	0.10	3.7	1.98	1.54	nd
C019	ACE California Apple Hard Cider	4.91	20.0	12.2	nd	32.3	7.20	0.10	0.19	nd	0.38	0.23	8.1	2.99	4.28	nd
C020	Angry Orchard Traditional Dry	5.18	6.70	8.46	23.8	39.0	4.25	0.09	0.37	nd	0.20	0.15	5.1	2.82	2.45	nd
C021	Woodchuck Gumption	5.10	12.0	6.93	33.3	52.2	4.14	0.10	0.18	0.27	0.34	0.06	5.1	5.01	1.88	nd
C022	Doc's Draft Gold Rush	7.02	0.75	0.60	nd	1.3	6.52	0.47	0.79	0.11	0.25	0.21	8.4	3.71	2.74	nd

ID	Cider	Ethanol	Fructose	Glucose	Sucrose	Sugar, total	Malic	Lactic	Succinic	Citric	Quinic	Acetic	Acids, total	Glycerol	Sorbitol	1,3-PDO
C023	Sarasola Original Basque Cider (Sagardoa)	5.95	0.18	nd	nd	0.2	nd	4.11	0.41	nd	0.16	2.17	6.8	0.19	3.83	1.58
C024	Pitchfork Hard Cider	6.81	0.19	0.58	nd	0.8	nd	4.00	0.50	nd	0.34	0.84	5.7	3.89	5.74	nd
C025	Celt Cidre Breton Traditionnel	4.08	25.6	4.68	0.64	30.9	1.32	1.90	0.30	nd	0.61	0.30	4.4	1.63	8.41	nd
C026	Hudson Valley Farmhouse Bourbon Barrel Aged	8.62	nd	nd	nd	0.0	0.54	3.21	0.69	nd	0.24	0.97	5.7	–	–	–
C027	Angry Orchard Summer Honey	4.96	6.14	7.03	28.9	42.1	3.10	0.11	0.48	nd	0.14	nd	3.8	2.95	2.57	0.13
C028	Angry Orchard Green Apple	5.12	5.75	7.41	42.9	56.1	3.57	0.11	0.49	0.16	0.15	0.03	4.5	2.93	1.99	nd
C029	Hudson Valley Farmhouse Maeve's Hudson Valley…	6.70	10.1	5.38	nd	15.4	1.00	4.97	0.74	nd	0.09	1.01	7.8	4.48	5.15	nd
C030	Nine Pin Cider Works Belgian	6.79	17.4	7.56	nd	24.9	3.00	0.67	0.37	nd	0.28	0.08	4.4	3.47	3.65	nd
C031	Magners Original Irish Cider	4.14	5.81	6.99	16.9	29.7	1.67	0.78	0.26	nd	0.29	0.12	3.1	2.19	1.88	0.10
C032	Nine Pin Cider Works Signature NY Hard Cider	6.16	19.4	8.48	nd	27.9	5.08	0.38	0.52	nd	0.26	0.13	6.4	–	–	–
C033	Ironwood Hard Cider	5.56	17.2	8.94	0.66	26.8	3.68	0.43	0.27	nd	0.12	0.08	4.6	1.56	1.83	nd
C034	Samuel Smith's Organic Cider	4.93	4.11	5.26	15.1	24.4	3.45	0.79	0.24	nd	0.59	0.12	5.2	2.42	1.99	nd
C035	Bad Seed Hard Apple Cider (Dry)	7.21	nd	0.09	0.09	0.2	5.55	0.06	0.27	nd	0.30	0.18	6.4	2.62	3.24	nd
C036	Angry Orchard Crisp Apple	4.81	7.14	7.60	34.7	49.4	4.35	0.10	0.38	nd	0.20	0.13	5.2	2.93	1.59	nd
C037	Bad Seed The Farmer Saison	7.69	0.25	0.09	nd	0.3	4.86	0.09	0.40	nd	0.23	0.19	5.8	3.70	4.48	nd
C038	Downeast Original Blend - Batch X, overpressu…	6.82	25.9	6.23	0.54	32.7	4.47	0.36	0.30	nd	0.33	0.05	5.5	–	–	–
C039	Downeast Original Blend - Batch X, sound can	6.01	37.1	13.7	1.07	51.8	4.19	0.31	0.33	nd	0.37	nd	5.2	–	–	–
C040	Downeast Original Blend - Batch Y	6.20	36.6	16.1	4.15	56.9	1.56	2.31	0.29	nd	0.31	0.29	4.7	–	–	–
C041	Downeast Hard Honey Cider	3.48	16.3	14.3	0.52	31.1	nd	1.66	0.38	nd	0.04	0.06	2.1	2.96	0.74	nd
C042	Downeast Cranberry Blend	2.95	26.7	14.1	0.24	41.1	3.34	0.05	0.19	0.91	1.05	0.03	5.6	1.67	3.74	nd
C043	Awestruck Eastern Dry	7.20	7.52	8.03	6.66	22.2	3.51	0.40	0.58	nd	0.38	0.17	5.0	–	–	–
C044	Stew Leonard's Fresh apple cider, unpasteuriz…	0.75	73.5	21.9	6.96	102.3	1.41	1.99	0.03	0.10	0.13	0.17	3.8	0.96	5.05	nd
C045	Naked Flock Lemon and Ginger	4.95	17.3	18.9	nd	36.2	2.08	0.56	0.29	0.38	0.25	0.19	3.8	2.60	3.32	nd
C046	Citizen Cider The Dirty Mayor (ginger)	5.68	22.5	9.67	nd	32.1	4.44	0.06	0.40	0.19	0.20	0.11	5.4	3.42	4.20	nd
C047	Awestruck Hibiscus and Ginger ("Pink")	6.60	33.8	36.9	nd	70.7	2.19	2.32	0.52	nd	0.45	0.22	5.7	4.13	3.49	nd
C048	Downeast Original Blend (lot exp. 9-28-15)	5.12	30.4	11.8	0.15	42.4	3.35	0.30	0.27	nd	0.28	0.06	4.3	2.04	3.68	nd

ID	Cider	Ethanol	Fructose	Glucose	Sucrose	Sugar, total	Malic	Lactic	Succinic	Citric	Quinic	Acetic	Acids, total	Glycerol	Sorbitol	1,3-PDO
C049	Citizen Cider The Full Nelson (dry hopped)	6.04	7.16	3.59	nd	10.8	4.07	0.54	0.44	nd	0.23	0.12	5.4	3.51	4.27	nd
C050	Citizen Cider Unified Press	6.41	16.4	8.14	nd	24.6	4.69	0.09	0.46	nd	0.28	0.05	5.6	3.36	4.77	0.18
C051	Ryan Burk Baldwin - Crispin - Rhode Island Gr…	6.89	2.94	1.25	0.85	5.0	0.21	3.20	0.80	nd	0.35	0.41	5.0	4.30	4.99	nd
C052	Angry Orchard Strawman	10.35	5.89	3.89	1.10	10.9	6.77	0.10	0.77	nd	0.69	0.37	8.7	6.23	5.50	nd
C053	Angry Orchard Stone Dry	5.02	4.32	4.14	5.63	14.1	2.61	0.13	0.39	nd	0.25	0.11	3.5	3.30	2.42	nd
C054	E.Z. Orchards Roman Beauty (heirloom)	4.26	18.6	1.35	nd	19.9	nd	3.22	0.68	nd	0.68	0.53	5.1	3.17	4.40	nd
C055	E.Z. Orchards Hawk Haus (Jonathan and Yarling…	6.39	0.40	nd	nd	0.4	nd	4.51	0.51	nd	0.34	0.19	5.6	3.28	6.14	nd
C056	E.Z. Orchards Poire (Forelle / Comice / Bosc …	5.63	nd	0.15	nd	0.2	0.43	2.98	0.49	nd	0.12	0.34	4.4	3.05	22.6	0.15
C057	Bad Seed Bourbon Barrel Reserve	8.18	0.19	nd	nd	0.2	5.84	0.18	0.55	nd	0.32	0.10	7.0	4.02	2.00	nd
C058	1911 Original	6.09	14.7	6.42	nd	21.1	7.33	0.08	0.27	nd	0.32	0.12	8.1	2.90	4.67	nd
C059	Angry Orchard Apple Ginger	4.79	9.06	10.8	20.9	40.8	3.88	0.10	0.38	nd	0.19	0.13	4.7	2.81	1.59	nd
C060	Graft Farm Flor (rustic, wild fermented)	6.57	0.16	nd	nd	0.2	nd	7.92	0.49	nd	0.06	2.82	11.3	2.58	4.16	1.46
C061	Graft Desert Diamonds Gose Cider	6.39	0.85	1.19	nd	2.0	nd	7.87	0.49	nd	0.13	2.79	11.3	2.69	5.25	1.71
C062	Graft Book of Nomad Gose Cider	6.51	2.25	0.98	nd	3.2	nd	6.04	1.54	nd	0.11	4.86	12.5	2.89	13.6	1.15
C063	Graft Where Secrets Lie Gose Cider	6.59	nd	nd	nd	0.0	0.28	5.66	0.59	nd	0.15	2.81	9.5	2.63	7.89	1.72
C064	Graft Wild Woods - Grim Wilderness Vol. 1 (oa…	6.42	nd	2.94	nd	2.9	nd	9.16	0.58	nd	0.34	3.78	13.9	1.95	3.90	2.21
C065	Stormalong Legendary Dry (lot A)	6.92	2.56	2.78	nd	5.3	4.45	0.05	0.65	0.10	0.61	0.20	6.1	4.38	4.89	nd
C066	Stormalong Legendary Dry (lot B)	5.50	3.45	4.00	nd	7.4	nd	2.86	1.13	nd	0.29	0.23	4.5	4.32	3.16	nd
C067	Stormalong Legendary Dry (lot C)	6.87	2.45	2.50	nd	5.0	4.32	0.05	0.59	nd	0.54	0.16	5.7	3.83	4.69	nd
C068	Stormalong Mass Appeal (lot A)	4.53	20.5	1.94	0.13	22.6	4.71	0.04	0.42	nd	0.21	0.11	5.5	3.53	3.32	nd
C069	Stormalong Mass Appeal (lot B)	5.20	13.0	3.71	4.56	21.2	nd	3.54	0.37	nd	0.24	0.53	4.7	2.94	5.21	0.16
C070	Stormalong Mass Appeal (lot C)	5.17	15.5	1.45	1.69	18.7	4.74	0.04	0.55	nd	0.17	0.10	5.6	3.50	4.07	nd
C071	Stormalong Red Skies at Night (lot A)	5.19	11.9	3.31	4.59	19.8	4.80	0.03	0.55	0.18	0.24	0.11	5.9	3.46	5.06	nd
C072	Stormalong Red Skies at Night (lot B)	5.26	13.1	4.21	3.79	21.1	4.93	0.04	0.44	0.18	0.40	0.12	6.1	3.03	5.73	nd
C073	Stormalong Red Skies at Night (lot C)	4.64	15.4	6.61	2.77	24.8	3.86	0.05	0.47	0.14	0.29	0.06	4.9	3.16	5.26	nd
C074	Stormalong Light of the Sun (lot A)	6.03	1.41	0.67	nd	2.1	4.89	0.06	0.60	0.22	0.30	0.10	6.2	3.74	3.19	nd
C075	Stormalong Light of the Sun (lot B)	6.49	1.85	0.98	nd	2.8	5.29	0.07	0.72	0.24	0.27	0.14	6.7	4.28	4.05	nd

ID	Cider	Ethanol	Fructose	Glucose	Sucrose	Sugar, total	Malic	Lactic	Succinic	Citric	Quinic	Acetic	Acids, total	Glycerol	Sorbitol	1,3-PDO
C076	Stormalong Light of the Sun (lot C)	6.39	1.89	0.87	nd	2.8	5.05	0.05	0.60	0.26	0.26	0.15	6.4	4.07	4.03	nd
C077	Stormalong Boston Heirloom (lot A)	4.53	28.0	9.35	8.04	45.4	5.38	0.04	0.45	0.13	0.38	0.08	6.5	3.31	4.84	nd
C078	Stormalong Boston Heirloom (lot B)	5.04	24.2	7.32	5.43	36.9	6.07	0.03	0.49	0.10	0.25	0.14	7.1	3.75	4.49	nd
C079	Hudson North Cider Co. Peach Mango Dry Cider	5.36	12.0	8.97	0.25	21.2	1.23	0.12	0.18	0.10	0.08	0.15	1.9	2.54	2.13	nd
C080	Downeast Cranberry Cider	6.10	26.7	14.8	7.17	48.6	6.18	0.12	0.29	3.30	3.17	0.16	13.2	3.21	3.35	nd
C081	Bad Seed Opus None (aged on Cabernet skins, o…	7.16	0.13	0.16	nd	0.3	3.12	0.07	0.32	nd	0.11	0.04	3.7	3.35	2.59	nd
C082	Graft Back Country (sour cider, citrus and bo…	6.87	6.00	6.11	nd	12.1	nd	4.07	0.81	nd	0.35	0.55	5.8	3.77	3.77	nd
C083	Angry Orchard Peach Mango	4.74	2.90	3.61	26.8	33.3	2.71	0.29	0.25	nd	nd	0.08	3.3	–	–	–
C084	Downeast Blood Orange	5.11	10.3	10.0	34.1	54.4	2.05	0.19	0.23	1.85	0.27	0.07	4.7	2.27	1.63	nd
C085	Locust Grove Honey Crisp Cider (4-3-24)	5.84	12.4	12.5	nd	24.8	1.97	2.35	0.50	nd	nd	0.39	5.2	–	–	–
C086	Locust Grove Peach Cider (4-3-24)	5.48	nd	nd	nd	0.0	3.41	0.36	0.36	0.83	0.67	0.11	5.7	–	–	–
C087	Downeast Guava Passion Fruit	5.17	8.95	7.84	36.2	53.0	2.08	0.17	0.21	2.68	0.26	0.14	5.5	2.20	1.50	nd
C088	1911 Original Cider - Alcohol Removed	0.23	10.4	3.09	3.19	16.7	4.68	0.07	0.20	nd	0.19	0.12	5.3	2.29	3.56	nd
C089	Original Sin White Widow non-alcoholic cider …	0.14	11.7	6.21	1.77	19.7	0.45	0.10	0.01	0.87	0.11	0.54	2.1	0.27	1.49	nd
C090	Graft Native (spontaneous natural cider)	7.31	0.32	nd	nd	0.3	nd	3.99	0.52	nd	0.06	1.82	6.4	2.40	3.40	0.92
C091	Graft Book of Nomad - Empress's Daughter (str…	6.91	7.76	8.27	nd	16.0	nd	2.96	0.37	0.98	0.26	0.11	4.7	3.38	4.19	nd
C092	Graft Book of Nomad - Crystal Temple (Caberne…	7.67	5.76	5.14	nd	10.9	0.86	4.93	0.78	nd	nd	1.70	8.3	4.75	3.71	nd
C093	Barton Farms Base cider	6.77	6.86	6.27	0.06	13.2	0.99	2.33	0.92	nd	0.25	0.12	4.6	4.40	3.13	nd
C094	Barton Farms Plow (raspberry)	6.69	15.7	15.0	0.13	30.8	3.74	0.07	0.78	0.33	0.23	0.05	5.2	3.65	4.19	nd
C095	Barton Farms Tractor	6.73	6.77	6.27	0.08	13.1	3.99	0.05	0.82	nd	0.24	0.06	5.2	4.14	4.19	nd
C096	Barton Farms Hayride (black cherry)	6.74	14.8	14.2	0.09	29.1	1.01	2.24	0.86	0.18	0.22	0.12	4.6	4.09	3.10	nd

Table 23. Minor constituents (mg/L) and aromatic protons (mmol/L). 2,3-BD, 2,3-butanediol; HMF, 5-hydroxymethylfurfural; α-Xyl, doublet at 5.18 ppm calculated as α-xylose; α-GalA, α anomer of galacturonic acid; HCA, hydroxycinnamoyl esters calculated as chlorogenic acid. nd, below the reporting limit; –, not determined. Tentative assignments are marked †.

ID	Cider	Methanol	Isoamyl alcohols	Isobutanol†	1-Propanol†	2,3-BD	Sorbic acid	Formic acid	HMF	α-Xyl†	α-GalA	Alanine†	HCA	Aromatic H
C001	Home cider maker Tim's Cider #1	70.7	213	50.1	69.0	46.8	nd	25.2	nd	nd	nd	20.3	nd	1.9
C002	Home cider maker Tim's Cider #2	53.2	180	40.3	271	122	nd	nd	nd	nd	nd	nd	nd	3.8
C003	J.K.'s Scrumpy Farmhouse Hard Cider (USDA Org…	47.8	109	49.5	nd	156	nd	nd	nd	113	179	nd	nd	5.3
C004	Harpoon Craft Cider	nd	182	74.7	16.6	nd	nd	nd	nd	112	nd	45.4	46.9	3.7
C005	Stella Artois Cidre	nd	45.7	7.0	14.0	652	nd	nd	3.3	nd	109	6.9	32.4	2.0
C006	Home cider maker Tim's Cider #3 (Peach-Honey)	76.6	211	63.4	12.7	241	nd	nd	nd	58.2	nd	nd	nd	4.3
C007	Etienne Dupont Cidre Triple	69.2	153	34.1	16.8	40.7	nd	nd	nd	124	120	nd	513	12.0
C008	Crispin Hard Apple Cider	nd	86.8	35.2	5.8	144	nd	nd	nd	nd	172	14.8	199	3.4
C009	Strongbow Gold Apple	nd	47.2	7.7	20.8	57.2	nd	nd	nd	nd	99.6	nd	173	3.6
C010	Woodchuck Out on a Limb - Splitter Splinter (…	71.2	98.9	37.1	6.4	95.3	nd	nd	nd	65.3	451	nd	148	4.2
C011	Thatchers Green Goblin Oak Aged English Cider	44.6	92.1	15.9	17.6	179	nd	nd	nd	nd	403	nd	142	5.7
C012	Magners Pear Cider	41.7	63.3	nd	19.0	106	196	nd	nd	nd	169	nd	45.7	1.9
C013	Etienne Dupont Cidre Bouché Brut de Normandie	106	103	28.8	17.8	80.3	nd	nd	nd	136	172	nd	392	10.3
C014	Naked Flock Hard Cider (Belgian ale yeast, ma…	26.3	144	40.6	12.0	200	143	12.4	nd	167	120	nd	38.9	4.5
C015	Doc's Draft Original Hard Apple Cider	38.6	120	27.4	nd	40.4	247	nd	nd	96.6	209	nd	43.5	2.8
C016	Woodchuck Amber	37.3	34.7	12.0	12.8	172	nd	3.2	2.0	51.0	439	5.6	nd	2.0
C017	Blackthorn English Cider	nd	66.4	12.7	8.1	52.8	nd	nd	nd	64.5	nd	nd	235	5.6
C018	Woodchuck 802	21.0	33.1	7.5	26.8	55.3	nd	nd	nd	49.9	545	nd	56.6	0.9
C019	ACE California Apple Hard Cider	nd	39.9	11.1	16.9	291	152	nd	nd	148	73.3	nd	nd	1.6
C020	Angry Orchard Traditional Dry	nd	131	35.4	14.3	95.1	nd	nd	nd	77.8	161	49.8	67.9	2.6
C021	Woodchuck Gumption	22.4	61.0	32.6	14.0	189	nd	nd	nd	72.1	130	nd	209	4.2
C022	Doc's Draft Gold Rush	30.7	191	38.7	5.1	67.5	253	nd	nd	159	70.4	nd	39.8	5.5
C023	Sarasola Original Basque Cider (Sagardoa)	54.8	131	34.4	78.9	133	nd	nd	nd	nd	nd	nd	66.7	4.7
C024	Pitchfork Hard Cider	29.5	210	63.5	15.1	65.0	nd	nd	nd	76.5	nd	nd	90.9	2.6
C025	Celt Cidre Breton Traditionnel	31.8	48.7	17.3	6.5	183	nd	9.0	nd	81.6	nd	12.0	442	9.8
C026	Hudson Valley Farmhouse Bourbon Barrel Aged	85.5	167	52.1	8.3	–	nd	4.6	nd	nd	173	nd	83.2	3.2
C027	Angry Orchard Summer Honey	nd	139	24.4	24.0	42.8	nd	nd	nd	94.7	376	27.5	nd	2.3

ID	Cider	Methanol	Isoamyl alcohols	Isobutanol†	1-Propanol†	2,3-BD	Sorbic acid	Formic acid	HMF	α-Xyl†	α-GalA	Alanine†	HCA	Aromatic H
C028	Angry Orchard Green Apple	nd	144	29.3	16.2	234	nd	nd	nd	91.9	405	24.9	40.6	1.7
C029	Hudson Valley Farmhouse Maeve's Hudson Valley…	21.1	128	33.0	11.7	113	nd	5.5	nd	nd	111	nd	44.0	2.3
C030	Nine Pin Cider Works Belgian	70.4	163	20.0	12.0	209	111	4.1	nd	283	133	nd	34.2	3.9
C031	Magners Original Irish Cider	68.4	88.7	38.3	12.7	42.1	nd	4.1	3.5	55.9	300	7.4	81.3	4.2
C032	Nine Pin Cider Works Signature NY Hard Cider	39.1	170	53.5	34.4	–	102	nd	nd	nd	nd	nd	316	5.9
C033	Ironwood Hard Cider	13.8	90.3	26.6	5.7	92.4	50.0	nd	nd	81.7	86.5	nd	34.6	1.3
C034	Samuel Smith's Organic Cider	nd	80.2	35.5	8.5	193	nd	nd	nd	106	87.8	6.7	222	5.5
C035	Bad Seed Hard Apple Cider (Dry)	13.7	104	15.1	19.2	94.6	nd	nd	nd	81.4	30.5	nd	59.4	2.7
C036	Angry Orchard Crisp Apple	nd	127	28.7	13.8	168	nd	nd	nd	69.5	201	37.3	68.3	3.6
C037	Bad Seed The Farmer Saison	36.0	170	40.0	9.6	59.6	nd	nd	nd	149	70.0	23.6	42.6	3.3
C038	Downeast Original Blend - Batch X, overpressu…	nd	177	40.3	8.6	–	96.9	nd	nd	300	nd	65.7	61.9	1.9
C039	Downeast Original Blend - Batch X, sound can	nd	185	33.2	10.2	–	124	nd	nd	303	nd	nd	114	2.4
C040	Downeast Original Blend - Batch Y	nd	150	17.5	nd	–	245	nd	nd	262	nd	nd	159	1.9
C041	Downeast Hard Honey Cider	nd	123	44.3	9.5	146	181	7.8	nd	nd	nd	nd	nd	2.7
C042	Downeast Cranberry Blend	nd	65.0	9.9	nd	133	188	nd	nd	158	nd	18.1	nd	3.3
C043	Awestruck Eastern Dry	25.0	170	22.9	10.5	–	nd	nd	nd	nd	nd	44.9	83.7	4.2
C044	Stew Leonard's Fresh apple cider, unpasteuriz…	29.0	nd	9.0	7.1	nd	44.1	nd	nd	233	95.5	21.6	53.8	1.2
C045	Naked Flock Lemon and Ginger	12.8	86.3	24.6	20.9	27.4	137	nd	nd	60.6	50.8	nd	86.6	2.8
C046	Citizen Cider The Dirty Mayor (ginger)	27.6	137	16.8	8.7	29.0	106	5.0	nd	171	77.8	11.7	63.7	2.8
C047	Awestruck Hibiscus and Ginger ("Pink")	nd	178	37.5	19.9	172	nd	nd	nd	158	33.3	nd	nd	4.4
C048	Downeast Original Blend (lot exp. 9-28-15)	nd	95.2	12.9	5.5	51.4	134	nd	nd	196	nd	nd	nd	2.1
C049	Citizen Cider The Full Nelson (dry hopped)	33.0	133	17.1	19.0	390	121	nd	nd	256	63.6	nd	76.6	3.3
C050	Citizen Cider Unified Press	20.3	136	28.5	7.9	nd	nd	nd	nd	167	57.7	5.5	50.9	1.9
C051	Ryan Burk Baldwin - Crispin - Rhode Island Gr…	37.0	154	24.9	nd	83.8	nd	3.2	nd	117	34.5	nd	314	9.3
C052	Angry Orchard Strawman	20.3	232	61.0	21.3	206	nd	4.1	3.9	100	262	44.0	275	8.6
C053	Angry Orchard Stone Dry	nd	144	30.7	14.4	61.2	nd	nd	nd	73.4	207	41.3	74.7	4.5
C054	E.Z. Orchards Roman Beauty (heirloom)	104	78.5	28.5	nd	70.5	nd	nd	nd	256	233	nd	127	6.7
C055	E.Z. Orchards Hawk Haus (Jonathan and Yarling…	50.7	101	22.4	12.1	25.8	nd	nd	nd	136	100	nd	70.4	5.1
C056	E.Z. Orchards Poire (Forelle / Comice / Bosc …	202	103	32.7	19.9	86.4	nd	nd	nd	380	475	nd	nd	4.1
C057	Bad Seed Bourbon Barrel Reserve	22.7	158	56.7	9.3	58.5	nd	nd	nd	77.5	43.4	17.4	58.7	3.1

ID	Cider	Methanol	Isoamyl alcohols	Isobutanol†	1-Propanol†	2,3-BD	Sorbic acid	Formic acid	HMF	α-Xyl†	α-GalA	Alanine†	HCA	Aromatic H
C058	1911 Original	nd	124	20.6	nd	63.6	222	nd	nd	351	nd	10.1	64.6	2.7
C059	Angry Orchard Apple Ginger	nd	129	25.5	16.8	49.2	nd	nd	nd	54.9	174	24.1	63.5	4.5
C060	Graft Farm Flor (rustic, wild fermented)	49.1	170	62.0	235	162	nd	8.1	nd	nd	nd	nd	nd	3.6
C061	Graft Desert Diamonds Gose Cider	35.6	190	79.4	101	189	nd	6.8	nd	nd	nd	nd	nd	3.1
C062	Graft Book of Nomad Gose Cider	48.2	111	36.3	238	161	nd	1.9	nd	nd	nd	nd	nd	5.3
C063	Graft Where Secrets Lie Gose Cider	26.6	170	78.2	192	187	nd	3.1	nd	nd	nd	nd	60.1	4.9
C064	Graft Wild Woods - Grim Wilderness Vol. 1 (oa…	159	141	61.8	281	303	nd	5.8	nd	nd	nd	nd	nd	4.6
C065	Stormalong Legendary Dry (lot A)	152	146	27.9	15.9	98.8	nd	nd	nd	172	67.6	55.1	233	5.9
C066	Stormalong Legendary Dry (lot B)	110	179	46.8	19.2	149	nd	nd	nd	92.5	48.5	nd	111	5.6
C067	Stormalong Legendary Dry (lot C)	131	130	18.6	9.5	106	nd	nd	nd	144	39.5	59.3	161	5.0
C068	Stormalong Mass Appeal (lot A)	137	110	25.0	nd	44.7	nd	nd	nd	170	nd	67.6	60.1	3.2
C069	Stormalong Mass Appeal (lot B)	116	130	22.0	8.2	36.5	nd	nd	nd	149	30.6	nd	38.9	3.0
C070	Stormalong Mass Appeal (lot C)	137	108	19.7	nd	41.3	nd	nd	nd	190	41.8	52.4	60.2	2.0
C071	Stormalong Red Skies at Night (lot A)	142	142	22.7	nd	50.6	nd	nd	2.6	135	40.6	26.2	93.6	4.0
C072	Stormalong Red Skies at Night (lot B)	136	126	22.7	nd	30.9	nd	nd	nd	220	36.0	37.4	94.1	3.5
C073	Stormalong Red Skies at Night (lot C)	108	131	19.1	nd	27.9	nd	nd	nd	220	34.8	30.5	nd	3.8
C074	Stormalong Light of the Sun (lot A)	158	186	33.4	nd	44.3	nd	nd	nd	148	59.1	31.2	35.3	2.5
C075	Stormalong Light of the Sun (lot B)	207	183	29.9	5.1	37.3	nd	nd	nd	153	119	45.2	66.1	3.6
C076	Stormalong Light of the Sun (lot C)	180	170	26.4	nd	38.0	nd	nd	nd	167	85.2	47.8	79.2	3.4
C077	Stormalong Boston Heirloom (lot A)	132	115	19.0	nd	76.2	nd	nd	nd	256	nd	57.7	83.2	2.0
C078	Stormalong Boston Heirloom (lot B)	124	113	15.6	7.4	63.8	nd	nd	nd	187	34.9	37.5	82.6	2.0
C079	Hudson North Cider Co. Peach Mango Dry Cider	91.3	44.0	20.9	23.1	81.3	nd	nd	3.0	133	37.5	9.1	31.4	0.8
C080	Downeast Cranberry Cider	63.2	82.3	17.1	19.2	161	nd	nd	nd	51.9	111	7.8	nd	6.5
C081	Bad Seed Opus None (aged on Cabernet skins, o…	37.0	118	15.8	6.8	98.4	nd	nd	nd	153	76.2	nd	nd	3.1
C082	Graft Back Country (sour cider, citrus and bo…	114	53.2	23.5	105	243	nd	53.8	nd	140	nd	nd	nd	2.7
C083	Angry Orchard Peach Mango	nd	93.6	18.2	16.2	–	nd	nd	nd	nd	210	37.3	144	8.5
C084	Downeast Blood Orange	43.0	73.3	14.6	14.4	284	nd	nd	nd	59.8	43.1	nd	60.6	3.2
C085	Locust Grove Honey Crisp Cider (4-3-24)	nd	225	62.6	24.1	–	117	nd	nd	nd	nd	nd	nd	6.0
C086	Locust Grove Peach Cider (4-3-24)	185	97.3	33.8	25.0	–	nd	nd	nd	nd	339	79.4	nd	5.7
C087	Downeast Guava Passion Fruit	78.3	60.8	6.3	11.2	314	nd	4.2	nd	68.7	127	nd	nd	2.9

ID	Cider	Methanol	Isoamyl alcohols	Isobutanol†	1-Propanol†	2,3-BD	Sorbic acid	Formic acid	HMF	α-Xyl†	α-GalA	Alanine†	HCA	Aromatic H
C088	1911 Original Cider - Alcohol Removed	nd	nd	nd	nd	40.4	nd	nd	nd	103	nd	16.2	43.3	2.7
C089	Original Sin White Widow non-alcoholic cider …	nd	nd	nd	nd	256	nd	3.3	nd	nd	70.0	16.0	nd	3.2
C090	Graft Native (spontaneous natural cider)	34.7	58.0	24.8	87.6	102	nd	2.9	nd	36.6	nd	nd	nd	2.9
C091	Graft Book of Nomad - Empress's Daughter (str…	123	66.2	11.3	69.0	198	nd	nd	nd	94.3	88.5	nd	33.4	2.1
C092	Graft Book of Nomad - Crystal Temple (Caberne…	233	180	72.4	30.0	194	nd	nd	nd	nd	295	nd	nd	3.0
C093	Barton Farms Base cider	15.8	103	38.2	22.3	62.7	127	nd	2.2	82.7	nd	nd	90.1	3.0
C094	Barton Farms Plow (raspberry)	nd	85.1	13.5	15.9	66.5	107	nd	2.9	125	nd	7.0	103	3.5
C095	Barton Farms Tractor	nd	85.9	14.6	14.5	65.4	111	nd	1.4	150	nd	6.7	111	3.5
C096	Barton Farms Hayride (black cherry)	nd	102	26.5	24.1	57.1	120	nd	6.2	74.7	nd	nd	81.7	3.1

Appendix C. Dryness Classes and Nutrition Values

Table 24. Residual sugar (g/L), dryness classes, organic acids as malic acid (g/L), ratio of sugar to acid with the unadjusted class of the orchard-based scale, dominant acid, energy per 12 fl oz on the component and the TTB basis (kcal), and sugars, sugars plus polyols and estimated carbohydrate by difference per 12 fl oz (g).

ID	Cider	Sugar	ACA class	BJCP class	Acid as malic	Sugar / acid	Ratio class	Dominant acid	kcal	kcal, TTB	Sugars (g)	Sugars + polyols (g)	Carbohydrate (g)
C001	Home cider maker Tim's Cider #1	0.2	Dry	Dry	5.9	0.0	Dry	lactate	145	139	0.1	0.9	3.4
C002	Home cider maker Tim's Cider #2	0.1	Dry	Dry	7.2	0.0	Dry	lactate	172	165	0.0	2.5	5.4
C003	J.K.'s Scrumpy Farmhouse Hard Cider (USDA Org…	65.5	Sweet	Sweet	7.1	9.3	Sweet	lactate	227	220	23.2	26.5	29.3
C004	Harpoon Craft Cider	20.0	Semi-Sweet	Semi-sweet	5.4	3.7	Sweet	malate	122	117	7.1	9.2	11.2
C005	Stella Artois Cidre	35.4	Semi-Sweet	Semi-sweet	4.3	8.1	Sweet	malate	142	138	12.6	13.8	15.4
C006	Home cider maker Tim's Cider #3 (Peach-Honey)	9.0	Semi-Dry	Medium	4.0	2.2	Semi-Sweet	lactate	149	142	3.2	6.1	7.9
C007	Etienne Dupont Cidre Triple	0.4	Dry	Dry	4.8	0.1	Dry	lactate	215	206	0.1	3.7	6.0
C008	Crispin Hard Apple Cider	25.2	Semi-Sweet	Semi-sweet	3.5	7.2	Sweet	lactate	165	159	8.9	11.0	12.4
C009	Strongbow Gold Apple	65.3	Sweet	Sweet	3.6	18.1	Sweet	malate	202	198	23.2	24.9	26.4
C010	Woodchuck Out on a Limb - Splitter Splinter (…	31.3	Semi-Sweet	Semi-sweet	3.9	8.0	Sweet	malate	155	150	11.1	13.0	14.6
C011	Thatchers Green Goblin Oak Aged English Cider	30.2	Semi-Sweet	Semi-sweet	2.4	12.8	Sweet	lactate	169	163	10.7	12.9	14.0
C012	Magners Pear Cider	47.0	Sweet	Sweet	3.1	15.0	Sweet	malate	167	163	16.7	19.5	20.6
C013	Etienne Dupont Cidre Bouché Brut de Normandie	13.4	Semi-Dry	Medium	5.6	2.4	Semi-Sweet	lactate	161	154	4.8	8.5	11.0
C014	Naked Flock Hard Cider (Belgian ale yeast, ma…	39.7	Semi-Sweet	Semi-sweet	7.5	5.3	Sweet	malate	214	206	14.1	17.3	20.2
C015	Doc's Draft Original Hard Apple Cider	31.1	Semi-Sweet	Semi-sweet	4.8	6.5	Sweet	lactate	154	149	11.0	13.6	15.7
C016	Woodchuck Amber	56.3	Sweet	Sweet	3.5	16.0	Sweet	malate	177	173	20.0	21.8	23.2
C017	Blackthorn English Cider	22.6	Semi-Sweet	Semi-sweet	4.4	5.2	Sweet	malate	159	153	8.0	10.2	11.8
C018	Woodchuck 802	42.0	Semi-Sweet	Sweet	3.6	11.7	Sweet	malate	155	151	14.9	16.2	17.7
C019	ACE California Apple Hard Cider	32.3	Semi-Sweet	Semi-sweet	7.9	4.1	Sweet	malate	160	154	11.5	14.0	17.0
C020	Angry Orchard Traditional Dry	39.0	Semi-Sweet	Semi-sweet	5.0	7.8	Sweet	malate	170	164	13.8	15.7	17.6
C021	Woodchuck Gumption	52.2	Sweet	Sweet	4.9	10.7	Sweet	malate	190	183	18.5	21.0	22.8
C022	Doc's Draft Gold Rush	1.3	Dry	Dry	8.2	0.2	Dry	malate	158	150	0.5	2.8	5.8
C023	Sarasola Original Basque Cider (Sagardoa)	0.2	Dry	Dry	6.0	0.0	Dry	lactate	131	126	0.1	2.1	4.5
C024	Pitchfork Hard Cider	0.8	Dry	Dry	4.6	0.2	Dry	lactate	153	146	0.3	3.7	5.7
C025	Celt Cidre Breton Traditionnel	30.9	Semi-Sweet	Semi-sweet	3.6	8.5	Sweet	lactate	140	135	11.0	14.5	16.1
C026	Hudson Valley Farmhouse Bourbon Barrel Aged	0.0	Dry	Dry	4.9	0.0	Dry	lactate	177	172	0.0	–	–

ID	Cider	Sugar	ACA class	BJCP class	Acid as malic	Sugar / acid	Ratio class	Dominant acid	kcal	kcal, TTB	Sugars (g)	Sugars + polyols (g)	Carbohydrate (g)
C027	Angry Orchard Summer Honey	42.1	Semi-Sweet	Sweet	3.8	11.2	Sweet	malate	169	164	14.9	16.9	18.5
C028	Angry Orchard Green Apple	56.1	Sweet	Sweet	4.5	12.6	Sweet	malate	192	187	19.9	21.7	23.4
C029	Hudson Valley Farmhouse Maeve's Hudson Valley…	15.4	Semi-Dry	Medium	6.7	2.3	Semi-Sweet	lactate	174	166	5.5	8.9	11.7
C030	Nine Pin Cider Works Belgian	24.9	Semi-Sweet	Semi-sweet	4.1	6.1	Sweet	malate	183	177	8.8	11.4	13.1
C031	Magners Original Irish Cider	29.7	Semi-Sweet	Semi-sweet	2.8	10.7	Sweet	malate	133	129	10.5	12.0	13.3
C032	Nine Pin Cider Works Signature NY Hard Cider	27.9	Semi-Sweet	Semi-sweet	6.2	4.5	Sweet	malate	168	164	9.9	–	–
C033	Ironwood Hard Cider	26.8	Semi-Sweet	Semi-sweet	4.4	6.0	Sweet	malate	157	152	9.5	10.7	12.4
C034	Samuel Smith's Organic Cider	24.4	Semi-Sweet	Semi-sweet	4.7	5.2	Sweet	malate	143	138	8.7	10.2	12.2
C035	Bad Seed Hard Apple Cider (Dry)	0.2	Dry	Dry	6.2	0.0	Dry	malate	157	150	0.1	2.1	4.4
C036	Angry Orchard Crisp Apple	49.4	Sweet	Sweet	5.1	9.8	Sweet	malate	177	172	17.5	19.1	21.1
C037	Bad Seed The Farmer Saison	0.3	Dry	Dry	5.7	0.1	Dry	malate	169	161	0.1	3.0	5.1
C038	Downeast Original Blend - Batch X, overpressu…	32.7	Semi-Sweet	Semi-sweet	5.3	6.2	Sweet	malate	187	183	11.6	–	–
C039	Downeast Original Blend - Batch X, sound can	51.8	Sweet	Sweet	4.9	10.5	Sweet	malate	198	194	18.4	–	–
C040	Downeast Original Blend - Batch Y	56.9	Sweet	Sweet	4.0	14.1	Sweet	lactate	209	205	20.2	–	–
C041	Downeast Hard Honey Cider	31.1	Semi-Sweet	Semi-sweet	1.7	17.8	Sweet	lactate	121	116	11.0	12.4	13.1
C042	Downeast Cranberry Blend	41.1	Semi-Sweet	Sweet	5.0	8.3	Sweet	malate	129	125	14.6	16.5	18.5
C043	Awestruck Eastern Dry	22.2	Semi-Sweet	Semi-sweet	4.8	4.6	Sweet	malate	180	175	7.9	–	–
C044	Stew Leonard's Fresh apple cider, unpasteuriz…	102.3	Sweet	Sweet	3.3	31.4	Sweet	lactate	170	168	36.3	38.4	39.9
C045	Naked Flock Lemon and Ginger	36.2	Semi-Sweet	Semi-sweet	3.5	10.3	Sweet	malate	161	156	12.9	15.0	16.3
C046	Citizen Cider The Dirty Mayor (ginger)	32.1	Semi-Sweet	Semi-sweet	5.3	6.0	Sweet	malate	173	167	11.4	14.1	16.1
C047	Awestruck Hibiscus and Ginger ("Pink")	70.7	Sweet	Sweet	4.9	14.4	Sweet	lactate	247	239	25.1	27.8	29.9
C048	Downeast Original Blend (lot exp. 9-28-15)	42.4	Semi-Sweet	Sweet	4.0	10.5	Sweet	malate	173	168	15.1	17.1	18.7
C049	Citizen Cider The Full Nelson (dry hopped)	10.8	Semi-Dry	Medium	5.2	2.1	Semi-Sweet	malate	150	144	3.8	6.6	8.6
C050	Citizen Cider Unified Press	24.6	Semi-Sweet	Semi-sweet	5.4	4.5	Sweet	malate	177	171	8.7	11.7	13.7
C051	Ryan Burk Baldwin - Crispin - Rhode Island Gr…	5.0	Dry	Semi-dry	4.1	1.2	Semi-Dry	lactate	160	152	1.8	5.1	6.9
C052	Angry Orchard Strawman	10.9	Semi-Dry	Medium	8.4	1.3	Semi-Dry	malate	244	233	3.9	8.0	11.2
C053	Angry Orchard Stone Dry	14.1	Semi-Dry	Medium	3.4	4.2	Sweet	malate	130	125	5.0	7.0	8.4
C054	E.Z. Orchards Roman Beauty (heirloom)	19.9	Semi-Sweet	Medium	4.0	5.0	Sweet	lactate	127	122	7.1	9.8	11.7
C055	E.Z. Orchards Hawk Haus (Jonathan and Yarling…	0.4	Dry	Dry	4.3	0.1	Dry	lactate	144	137	0.1	3.5	5.5
C056	E.Z. Orchards Poire (Forelle / Comice / Bosc …	0.2	Dry	Dry	3.6	0.0	Dry	lactate	142	135	0.1	9.2	11.1

ID	Cider	Sugar	ACA class	BJCP class	Acid as malic	Sugar / acid	Ratio class	Dominant acid	kcal	kcal, TTB	Sugars (g)	Sugars + polyols (g)	Carbohydrate (g)
C057	Bad Seed Bourbon Barrel Reserve	0.2	Dry	Dry	6.8	0.0	Dry	malate	178	170	0.1	2.2	4.7
C058	1911 Original	21.1	Semi-Sweet	Semi-sweet	7.9	2.7	Semi-Sweet	malate	168	161	7.5	10.2	13.2
C059	Angry Orchard Apple Ginger	40.8	Semi-Sweet	Sweet	4.6	8.8	Sweet	malate	163	158	14.5	16.0	17.8
C060	Graft Farm Flor (rustic, wild fermented)	0.2	Dry	Dry	9.6	0.0	Dry	lactate	152	144	0.1	3.0	7.0
C061	Graft Desert Diamonds Gose Cider	2.0	Dry	Dry	9.6	0.2	Dry	lactate	153	144	0.7	4.1	8.2
C062	Graft Book of Nomad Gose Cider	3.2	Dry	Dry	11.7	0.3	Dry	lactate	165	156	1.1	7.4	11.9
C063	Graft Where Secrets Lie Gose Cider	0.0	Dry	Dry	8.4	0.0	Dry	lactate	154	146	0.0	4.3	7.7
C064	Graft Wild Woods - Grim Wilderness Vol. 1 (oa…	2.9	Dry	Dry	11.8	0.2	Dry	lactate	156	147	1.0	3.9	8.8
C065	Stormalong Legendary Dry (lot A)	5.3	Dry	Semi-dry	5.8	0.9	Dry	malate	162	155	1.9	5.2	7.4
C066	Stormalong Legendary Dry (lot B)	7.4	Dry	Semi-dry	3.8	2.0	Semi-Dry	lactate	134	127	2.6	5.3	6.9
C067	Stormalong Legendary Dry (lot C)	5.0	Dry	Semi-dry	5.4	0.9	Dry	malate	159	152	1.8	4.8	6.9
C068	Stormalong Mass Appeal (lot A)	22.6	Semi-Sweet	Semi-sweet	5.4	4.2	Sweet	malate	136	130	8.0	10.5	12.5
C069	Stormalong Mass Appeal (lot B)	21.2	Semi-Sweet	Semi-sweet	3.7	5.7	Sweet	lactate	148	142	7.5	10.5	12.2
C070	Stormalong Mass Appeal (lot C)	18.7	Semi-Sweet	Medium	5.6	3.4	Sweet	malate	144	138	6.6	9.3	11.4
C071	Stormalong Red Skies at Night (lot A)	19.8	Semi-Sweet	Medium	5.8	3.4	Sweet	malate	147	141	7.0	10.1	12.2
C072	Stormalong Red Skies at Night (lot B)	21.1	Semi-Sweet	Semi-sweet	5.9	3.6	Sweet	malate	151	145	7.5	10.6	12.9
C073	Stormalong Red Skies at Night (lot C)	24.8	Semi-Sweet	Semi-sweet	4.7	5.2	Sweet	malate	142	136	8.8	11.8	13.6
C074	Stormalong Light of the Sun (lot A)	2.1	Dry	Dry	6.1	0.3	Dry	malate	138	131	0.7	3.2	5.5
C075	Stormalong Light of the Sun (lot B)	2.8	Dry	Dry	6.7	0.4	Dry	malate	150	143	1.0	4.0	6.4
C076	Stormalong Light of the Sun (lot C)	2.8	Dry	Dry	6.3	0.4	Dry	malate	147	140	1.0	3.9	6.2
C077	Stormalong Boston Heirloom (lot A)	45.4	Sweet	Sweet	6.3	7.2	Sweet	malate	171	165	16.1	19.0	21.4
C078	Stormalong Boston Heirloom (lot B)	36.9	Semi-Sweet	Semi-sweet	7.0	5.3	Sweet	malate	170	163	13.1	16.0	18.6
C079	Hudson North Cider Co. Peach Mango Dry Cider	21.2	Semi-Sweet	Semi-sweet	1.8	11.6	Sweet	malate	144	140	7.5	9.2	9.9
C080	Downeast Cranberry Cider	48.6	Sweet	Sweet	11.3	4.3	Sweet	malate	212	204	17.2	19.6	24.3
C081	Bad Seed Opus None (aged on Cabernet skins, o…	0.3	Dry	Dry	3.6	0.1	Dry	malate	154	147	0.1	2.2	3.6
C082	Graft Back Country (sour cider, citrus and bo…	12.1	Semi-Dry	Medium	4.7	2.6	Semi-Sweet	lactate	169	161	4.3	7.0	9.1
C083	Angry Orchard Peach Mango	33.3	Semi-Sweet	Semi-sweet	3.3	10.1	Sweet	malate	145	142	11.8	–	–
C084	Downeast Blood Orange	54.4	Sweet	Sweet	4.6	11.9	Sweet	malate	188	183	19.3	20.7	22.4
C085	Locust Grove Honey Crisp Cider (4-3-24)	24.8	Semi-Sweet	Semi-sweet	4.7	5.3	Sweet	lactate	156	153	8.8	–	–
C086	Locust Grove Peach Cider (4-3-24)	0.0	Dry	Dry	5.3	0.0	Dry	malate	115	111	0.0	–	–

ID	Cider	Sugar	ACA class	BJCP class	Acid as malic	Sugar / acid	Ratio class	Dominant acid	kcal	kcal, TTB	Sugars (g)	Sugars + polyols (g)	Carbohydrate (g)
C087	Downeast Guava Passion Fruit	53.0	Sweet	Sweet	5.5	9.6	Sweet	malate	188	183	18.8	20.1	22.2
C088	1911 Original Cider - Alcohol Removed	16.7	Semi-Dry	Medium	5.2	3.2	Sweet	malate	41	38	5.9	8.0	9.9
C089	Original Sin White Widow non-alcoholic cider …	19.7	Semi-Sweet	Medium	2.1	9.4	Sweet	malate	35	34	7.0	7.6	8.4
C090	Graft Native (spontaneous natural cider)	0.3	Dry	Dry	5.6	0.1	Dry	lactate	160	153	0.1	2.5	4.8
C091	Graft Book of Nomad - Empress's Daughter (str…	16.0	Semi-Dry	Medium	3.9	4.1	Sweet	lactate	174	167	5.7	8.4	10.1
C092	Graft Book of Nomad - Crystal Temple (Caberne…	10.9	Semi-Dry	Medium	7.3	1.5	Semi-Dry	lactate	187	178	3.9	6.9	9.9
C093	Barton Farms Base cider	13.2	Semi-Dry	Medium	4.0	3.3	Sweet	lactate	167	160	4.7	7.4	9.0
C094	Barton Farms Plow (raspberry)	30.8	Semi-Sweet	Semi-sweet	5.2	6.0	Sweet	malate	191	184	10.9	13.7	15.6
C095	Barton Farms Tractor	13.1	Semi-Dry	Medium	5.1	2.6	Semi-Sweet	malate	168	160	4.7	7.6	9.5
C096	Barton Farms Hayride (black cherry)	29.1	Semi-Sweet	Semi-sweet	4.0	7.2	Sweet	lactate	189	182	10.3	12.9	14.6

Appendix D. Composition of the Juices of 56 Apple Varieties

Table 25. Sugars, sorbitol and malic acid in the 56 apple juices. Left: concentrations per liter of the juice as measured (g/L). Right: shares of the sugars (% by mass of fructose + glucose + sucrose), ratio of fructose to glucose, malic acid and sorbitol per 100 g of sugars, and malic acid in the juice standardized to 110 g/L of sugars.

No.	Variety	Sucrose	Glucose	Fructose	Sugars	Sorbitol	Malic acid	Sucrose (%)	Glucose (%)	Fructose (%)	Fru / Glc	Malic acid (g/100 g)	Sorbitol (g/100 g)	Malic acid, std. (g/L)
A01	Calville Blanc d'Hiver	122.5	33.5	138.8	295	7.1	19.6	41.6	11.4	47.1	4.1	6.6	2.4	7.3
A02	Fall Pippin	62.2	19.3	132.3	214	2.8	9.6	29.1	9.0	61.9	6.9	4.5	1.3	4.9
A03	Dr. Matthews	33.2	16.2	88.5	138	1.8	6.0	24.1	11.8	64.1	5.4	4.4	1.3	4.8
A04	Chieftain	57.6	18.5	90.7	167	1.5	8.3	34.5	11.1	54.4	4.9	5.0	0.9	5.5
A05	Pink Pearl	55.4	18.8	90.0	164	5.4	14.0	33.8	11.5	54.8	4.8	8.5	3.3	9.4
A06	Oaken Pin	51.0	42.4	132.8	226	4.1	23.1	22.5	18.7	58.7	3.1	10.2	1.8	11.2
A07	King David	47.4	19.0	108.0	174	1.9	10.0	27.2	10.9	61.9	5.7	5.7	1.1	6.3
A08	Benoni †	50.0	20.7	112.4	183	2.9	9.0	27.3	11.3	61.4	5.4	4.9	1.6	5.4
A09	Green Newtown Pippin	83.5	24.5	99.6	208	6.0	17.6	40.2	11.8	48.0	4.1	8.5	2.9	9.3
A10	Ashmead's Kernel	78.6	28.4	121.1	228	7.8	13.4	34.4	12.5	53.1	4.3	5.9	3.4	6.5
A11	Golden Harvey	101.5	26.9	138.8	267	16.9	20.6	38.0	10.1	51.9	5.2	7.7	6.3	8.5
A12	Yates	34.5	24.0	79.4	138	2.3	7.8	25.0	17.4	57.6	3.3	5.6	1.7	6.2
A13	Esopus Spitzenburg	60.2	25.6	130.8	217	2.9	9.0	27.8	11.8	60.4	5.1	4.2	1.3	4.6
A14	Tompkins King	58.6	33.6	150.1	242	4.6	13.6	24.2	13.9	62.0	4.5	5.6	1.9	6.2
A15	Golden Russet	66.8	42.7	184.2	294	15.2	12.2	22.7	14.5	62.7	4.3	4.2	5.2	4.6
A16	Decio	78.6	17.0	67.9	164	7.6	7.6	48.1	10.4	41.5	4.0	4.6	4.7	5.1
A17	Edward VII	48.0	24.4	145.2	218	5.8	9.8	22.1	11.2	66.7	5.9	4.5	2.6	5.0
A18	Rhode Island Greening	47.0	24.6	88.3	160	3.6	8.5	29.4	15.4	55.2	3.6	5.3	2.3	5.9
A19	Baldwin	104.2	13.2	83.8	201	3.3	16.5	51.8	6.6	41.7	6.3	8.2	1.6	9.0
A20	Starr †	8.3	15.8	75.1	99	2.1	8.7	8.4	15.9	75.7	4.8	8.8	2.1	9.7
A21	Razor Russet	66.1	27.3	138.2	232	2.9	9.2	28.5	11.8	59.7	5.1	4.0	1.3	4.4
A22	GoldRush	62.0	12.1	65.0	139	0.8	10.1	44.6	8.7	46.7	5.4	7.2	0.6	7.9
A23	Jewett Red (Nodhead)	44.5	29.5	146.3	220	3.3	8.4	20.2	13.4	66.4	5.0	3.8	1.5	4.2
A24	American Beauty (Sterling)	75.4	19.2	88.2	183	6.4	12.6	41.3	10.5	48.3	4.6	6.9	3.5	7.6
A25	Kidd's Orange Red	42.2	22.1	117.9	182	2.1	8.2	23.2	12.1	64.7	5.3	4.5	1.2	4.9
A26	Opalescent	41.9	22.2	95.0	159	3.2	7.4	26.3	14.0	59.7	4.3	4.6	2.0	5.1

No.	Variety	Sucrose	Glucose	Fructose	Sugars	Sorbitol	Malic acid	Sucrose (%)	Glucose (%)	Fructose (%)	Fru / Glc	Malic acid (g/100 g)	Sorbitol (g/100 g)	Malic acid, std. (g/L)
A27	Ben Davis	30.6	15.2	74.7	121	1.7	6.5	25.4	12.6	61.9	4.9	5.4	1.4	6.0
A28	Chestnut Crab	7.0	18.9	101.5	127	2.6	8.9	5.5	14.8	79.7	5.4	7.0	2.0	7.7
A29	Scarlet Crofton †	61.6	22.9	115.9	200	7.2	9.1	30.8	11.4	57.8	5.1	4.5	3.6	5.0
A30	Quinte	38.3	15.6	77.1	131	3.8	9.4	29.2	11.9	58.8	4.9	7.2	2.9	7.9
A31	Barnack Beauty	69.8	36.0	123.6	229	3.7	9.4	30.4	15.7	53.9	3.4	4.1	1.6	4.5
A32	Sweet Winter Pennock	31.2	18.2	88.1	138	1.2	5.7	22.7	13.3	64.0	4.8	4.1	0.9	4.6
A33	Bancroft	55.3	11.6	125.4	192	5.3	18.0	28.7	6.0	65.2	10.8	9.4	2.8	10.3
A34	Titus Pippin	33.8	18.3	105.9	158	2.9	8.4	21.4	11.6	67.1	5.8	5.3	1.9	5.9
A35	Northern Spy	91.2	35.1	172.5	299	9.7	13.1	30.5	11.7	57.7	4.9	4.4	3.3	4.8
A36	Pear Apple	89.5	18.6	155.1	263	11.1	19.7	34.0	7.1	58.9	8.3	7.5	4.2	8.2
A37	Golden Noble	66.8	31.1	127.1	225	7.9	23.4	29.7	13.8	56.5	4.1	10.4	3.5	11.4
A38	Devonshire Quarrenden	22.5	25.7	119.8	168	6.9	8.3	13.4	15.3	71.3	4.7	5.0	4.1	5.5
A39	Virginia Beauty	66.1	40.9	119.4	226	5.6	12.6	29.2	18.1	52.7	2.9	5.6	2.5	6.1
A40	Winter Banana	23.4	41.7	119.3	184	3.1	8.8	12.7	22.6	64.7	2.9	4.8	1.7	5.3
A41	Marie-Joseph d'Othée	39.6	16.7	75.8	132	2.5	13.1	30.0	12.6	57.4	4.5	9.9	1.9	10.9
A42	Mother	69.0	22.1	123.0	214	6.1	9.8	32.2	10.3	57.5	5.6	4.6	2.9	5.0
A43	Pitmaston Pineapple	50.2	66.7	212.8	330	18.3	14.2	15.2	20.2	64.5	3.2	4.3	5.6	4.7
A44	Roxbury Russet	77.4	41.2	148.9	268	28.3	13.0	28.9	15.4	55.7	3.6	4.9	10.6	5.4
A45	Shiawassee Beauty	33.8	13.9	143.5	191	4.8	12.8	17.7	7.3	75.1	10.3	6.7	2.5	7.4
A46	Wagener	75.6	23.6	127.1	226	2.8	12.8	33.4	10.4	56.2	5.4	5.7	1.2	6.2
A47	Sutton Beauty	47.0	22.8	103.5	173	2.3	9.5	27.1	13.2	59.7	4.5	5.5	1.3	6.0
A48	Yellow Newtown Pippin	73.4	42.6	154.1	270	18.1	9.6	27.2	15.8	57.0	3.6	3.6	6.7	3.9
A49	Collamer	29.2	12.3	99.9	141	2.4	10.7	20.7	8.7	70.7	8.2	7.6	1.7	8.4
A51	Montreal Peach	69.5	23.9	177.4	271	3.9	9.1	25.7	8.8	65.5	7.4	3.4	1.4	3.7
A52	Goof	17.1	39.7	108.5	165	3.8	8.3	10.4	24.0	65.6	2.7	5.0	2.3	5.5
A53	Ellison's Orange	89.2	22.8	97.6	210	9.6	10.0	42.5	10.9	46.6	4.3	4.8	4.6	5.3
A54	Granny Smith ‡	52.3	20.7	66.3	139	2.4	15.1	37.5	14.9	47.6	3.2	10.8	1.7	11.9
A55	Kingston Black †	45.8	46.0	128.8	221	7.7	9.9	20.7	20.9	58.4	2.8	4.5	3.5	4.9
G15	Gala (2015) ‡	44.9	60.8	221.5	327	11.1	6.6	13.7	18.6	67.7	3.6	2.0	3.4	2.2
F15	Fuji (2015) ‡	8.0	50.9	110.7	170	5.0	2.4	4.7	30.0	65.3	2.2	1.4	3.0	1.5

† juices in which fermentation had begun (ethanol detected); ‡ supermarket fruit. Sucrose values are lower limits because of hydrolysis.

Table 26. Minor constituents of the 56 apple juices per liter of the juice as measured (mg/L; aromatic protons in mmol/L), residual ethanol of the freeze-dried sample, and the potential alcohol content of the juice standardized to 110 g/L of sugars.

No.	Variety	Quinic acid	Citramalic acid	L-Rhamnitol	Asparagine	Aspartic acid	Alanine	Xylose (α)	Hydroxycinnamates as CGA	Aromatic H (mmol/L)	Ethanol, residual	Potential alcohol (% v/v)
A01	Calville Blanc d'Hiver	847	64.8	193	nd	nd	nd	359	621	8.8	nd	6.7
A02	Fall Pippin	222	34.7	74.2	1061	nd	10.2	345	217	4.0	nd	6.7
A03	Dr. Matthews	220	nd	121	nd	nd	nd	227	133	3.4	nd	6.6
A04	Chieftain	638	19.7	86.3	nd	nd	13.3	290	75.7	0.7	nd	6.7
A05	Pink Pearl	291	48.1	44.2	nd	nd	nd	358	120	3.3	nd	6.7
A06	Oaken Pin	833	15.4	91.0	1625	468	10.2	121	159	6.0	nd	6.6
A07	King David	326	27.8	75.0	187	nd	nd	480	180	3.0	nd	6.6
A08	Benoni †	423	59.3	114	nd	nd	nd	937	339	9.4	65.5	6.6
A09	Green Newtown Pippin	637	119	153	2181	nd	67.2	217	255	6.4	nd	6.7
A10	Ashmead's Kernel	431	172	175	559	419	14.1	147	278	4.1	nd	6.7
A11	Golden Harvey	390	39.7	141	515	772	45.8	139	243	7.3	nd	6.7
A12	Yates	693	33.9	82.3	nd	nd	nd	104	132	2.6	nd	6.6
A13	Esopus Spitzenburg	690	54.4	95.5	143	183	nd	432	nd	5.2	nd	6.6
A14	Tompkins King	696	45.6	92.0	853	nd	nd	258	233	4.8	nd	6.6
A15	Golden Russet	98.1	nd	190	nd	nd	nd	284	257	2.2	nd	6.6
A16	Decio	276	29.1	92.7	nd	nd	nd	172	291	0.8	nd	6.7
A17	Edward VII	460	12.6	136	427	nd	5.1	301	271	6.1	nd	6.6
A18	Rhode Island Greening	475	56.8	84.2	257	132	nd	107	166	2.9	nd	6.7
A19	Baldwin	552	nd	168	1392	434	16.7	92.2	203	4.8	nd	6.7
A20	Starr †	466	nd	43.3	190	nd	11.7	215	238	3.8	452	6.6
A21	Razor Russet	206	nd	81.8	365	nd	7.0	314	107	3.0	nd	6.7
A22	GoldRush	239	25.3	59.5	575	nd	6.0	132	68.4	2.3	nd	6.7
A23	Jewett Red (Nodhead)	460	117	42.3	450	429	23.3	394	213	0.0	nd	6.6
A24	American Beauty (Sterling)	326	17.5	242	290	nd	8.1	237	212	5.4	nd	6.7
A25	Kidd's Orange Red	613	24.1	51.3	nd	124	nd	456	81.6	2.3	nd	6.6
A26	Opalescent	453	10.3	54.5	304	nd	5.1	383	145	0.0	nd	6.6
A27	Ben Davis	385	12.5	74.6	357	nd	nd	283	251	3.9	nd	6.6

No.	Variety	Quinic acid	Citramalic acid	L-Rhamnitol	Asparagine	Aspartic acid	Alanine	Xylose (α)	Hydroxycinnamates as CGA	Aromatic H (mmol/L)	Ethanol, residual	Potential alcohol (% v/v)
A28	Chestnut Crab	306	26.7	136	nd	nd	nd	425	244	5.2	nd	6.6
A29	Scarlet Crofton †	136	nd	97.2	455	nd	17.4	641	268	4.0	518	6.7
A30	Quinte	260	49.3	nd	388	nd	nd	534	266	2.6	nd	6.7
A31	Barnack Beauty	130	38.0	197	nd	nd	nd	336	460	6.6	nd	6.7
A32	Sweet Winter Pennock	301	nd	56.6	nd	nd	nd	193	180	3.9	nd	6.6
A33	Bancroft	289	41.4	69.6	nd	382	nd	444	217	3.6	nd	6.7
A34	Titus Pippin	919	56.0	183	708	227	nd	591	243	5.8	nd	6.6
A35	Northern Spy	352	484	222	nd	nd	15.2	430	511	8.7	nd	6.7
A36	Pear Apple	936	nd	166	nd	nd	14.9	342	292	6.0	nd	6.7
A37	Golden Noble	2150	96.2	222	nd	nd	6.3	745	401	8.4	nd	6.7
A38	Devonshire Quarrenden	592	10.8	60.8	287	nd	nd	630	349	6.5	nd	6.6
A39	Virginia Beauty	1151	42.4	118	708	190	35.9	362	141	5.0	nd	6.7
A40	Winter Banana	1173	54.3	63.8	nd	nd	nd	324	630	9.4	nd	6.6
A41	Marie-Joseph d'Othée	584	148	175	187	283	nd	169	113	1.4	nd	6.7
A42	Mother	422	123	193	290	272	7.2	563	177	7.0	nd	6.7
A43	Pitmaston Pineapple	345	11.4	626	394	671	6.9	347	474	9.4	nd	6.6
A44	Roxbury Russet	125	nd	393	nd	311	nd	285	471	7.9	nd	6.7
A45	Shiawassee Beauty	417	24.8	157	nd	nd	nd	419	176	2.4	nd	6.6
A46	Wagener	457	44.2	97.1	520	296	8.0	129	142	3.0	nd	6.7
A47	Sutton Beauty	429	nd	74.5	288	nd	7.0	162	328	5.8	nd	6.6
A48	Yellow Newtown Pippin	237	nd	107	590	1357	10.9	427	268	6.3	nd	6.6
A49	Collamer	508	18.8	111	nd	nd	nd	448	133	3.0	nd	6.6
A51	Montreal Peach	374	337	72.0	nd	nd	nd	266	361	6.8	nd	6.6
A52	Goof	473	12.7	101	205	nd	10.8	315	109	2.7	nd	6.6
A53	Ellison's Orange	256	nd	51.2	454	nd	23.3	423	170	4.5	nd	6.7
A54	Granny Smith ‡	387	10.7	93.2	4306	nd	34.0	109	125	5.3	nd	6.7
A55	Kingston Black †	1203	58.7	135	nd	nd	nd	260	178	6.6	244	6.6
G15	Gala (2015) ‡	473	nd	66.2	841	519	94.1	208	182	3.9	nd	6.6
F15	Fuji (2015) ‡	169	119	58.5	841	242	12.2	210	71.3	0.2	nd	6.6

nd, below the reporting limit. CGA, chlorogenic acid. Potential alcohol: hexose equivalents × 0.511 × 0.92 / 7.893.

References

1. Bharti, S. K.; Roy, R. Quantitative ¹H NMR spectroscopy. *TrAC Trends in Analytical Chemistry* **2012**, *35*, 5–26. https://doi.org/10.1016/j.trac.2012.02.007
2. Pauli, G. F.; Gödecke, T.; Jaki, B. U.; Lankin, D. C. Quantitative ¹H NMR. Development and Potential of an Analytical Method: An Update. *Journal of Natural Products* **2012**, *75* (4), 834–851. https://doi.org/10.1021/np200993k
3. Holzgrabe, U. Quantitative NMR spectroscopy in pharmaceutical applications. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2010**, *57* (2), 229–240. https://doi.org/10.1016/j.pnmrs.2010.05.001
4. Malz, F.; Jancke, H. Validation of quantitative NMR. *Journal of Pharmaceutical and Biomedical Analysis* **2005**, *38* (5), 813–823. https://doi.org/10.1016/j.jpba.2005.01.043
5. Edwards, J. C. *¹H qNMR – Quantitative Chemical Fingerprint of Cider*; Technical Report; Process NMR Associates, LLC: Poughkeepsie, NY, 2016.
6. Edwards, J. C. *Quantitative ¹H NMR Compositional Survey of Fifty-Five Heirloom and Heritage Apple Varieties*; White Paper PNA-WP-HeirloomApple-NMR-002; Process NMR Associates, LLC: Poughkeepsie, NY, 2026.
7. Edwards, J. C. *Quantitative ¹H NMR Spectroscopic Fingerprinting of Alcoholic Cider: Chemical Component Distribution Across 56 Commercial Products and an Updated Survey of Apple, Apple Juice, and Cider NMR Methodology*; White Paper Series; Process NMR Associates, LLC: Poughkeepsie, NY, May 2026.
8. Juniper, B. E.; Mabberley, D. J. *The Story of the Apple*; Timber Press: Portland, OR, 2006.
9. Cornille, A.; Giraud, T.; Smulders, M. J.; Roldán-Ruiz, I.; Gladieux, P. The domestication and evolutionary ecology of apples. *Trends in Genetics* **2014**, *30* (2), 57–65. https://doi.org/10.1016/j.tig.2013.10.002
10. Cornille, A.; Gladieux, P.; Smulders, M. J. M.; Roldán-Ruiz, I.; Laurens, F.; Le Cam, B.; Nersesyan, A.; Clavel, J.; Olonova, M.; Feugey, L.; et al. New Insight into the History of Domesticated Apple: Secondary Contribution of the European Wild Apple to the Genome of Cultivated Varieties. *PLoS Genetics* **2012**, *8* (5), e1002703. https://doi.org/10.1371/journal.pgen.1002703
11. Velasco, R.; Zharkikh, A.; Affourtit, J.; Dhingra, A.; Cestaro, A.; Kalyanaraman, A.; Fontana, P.; Bhatnagar, S. K.; Troggio, M.; Pruss, D.; et al. The genome of the domesticated apple (*Malus* × *domestica* Borkh.). *Nature Genetics* **2010**, *42* (10), 833–839. https://doi.org/10.1038/ng.654
12. Duan, N.; Bai, Y.; Sun, H.; Wang, N.; Ma, Y.; Li, M.; Wang, X.; Jiao, C.; Legall, N.; Mao, L.; et al. Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nature Communications* **2017**, *8* (1), 249. https://doi.org/10.1038/s41467-017-00336-7
13. Beach, S. A.; Booth, N. O.; Taylor, O. M. *The Apples of New York*; New York Agricultural Experiment Station; J. B. Lyon: Albany, NY, 1905; Vols. 1 and 2.
14. Bussey, D. J.; Whealy, K., Ed. *The Illustrated History of Apples in the United States and Canada*; JAK KAW Press: Mount Horeb, WI, 2016; Vols. 1–7.
15. Volk, G. M.; Henk, A. D. Historic American Apple Cultivars: Identification and Availability. *Journal of the American Society for Horticultural Science* **2016**, *141* (3), 292–301. https://doi.org/10.21273/jashs.141.3.292
16. Routson, K. J.; Reilley, A. A.; Henk, A. D.; Volk, G. M. Identification of Historic Apple Trees in the Southwestern United States and Implications for Conservation. *HortScience* **2009**, *44* (3), 589–594. https://doi.org/10.21273/hortsci.44.3.589
17. Volk, G. M.; Chao, C. T.; Norelli, J.; Brown, S. K.; Fazio, G.; Peace, C.; McFerson, J.; Zhong, G. Y.; Bretting, P. The vulnerability of US apple (*Malus*) genetic resources. *Genetic Resources and Crop Evolution* **2015**, *62* (5), 765–794. https://doi.org/10.1007/s10722-014-0194-2
18. Burford, T. *Apples of North America: Exceptional Varieties for Gardeners, Growers, and Cooks*; Timber Press: Portland, OR, 2013.
19. Morgan, J.; Richards, A. *The New Book of Apples*, rev. ed.; Ebury Press: London, 2002.
20. Watson, B. *Cider, Hard and Sweet: History, Traditions, and Making Your Own*, 3rd ed.; The Countryman Press: Woodstock, VT, 2013.

21. Kerrigan, W. *Johnny Appleseed and the American Orchard: A Cultural History*; Johns Hopkins University Press: Baltimore, MD, 2012.
22. Hogg, R. *The Fruit Manual: A Guide to the Fruits and Fruit Trees of Great Britain*, 5th ed.; Journal of Horticulture Office: London, 1884.
23. Coxe, W. *A View of the Cultivation of Fruit Trees, and the Management of Orchards and Cider*; M. Carey and Son: Philadelphia, 1817. https://doi.org/10.5962/bhl.title.18915
24. Lea, A. G. H.; Drilleau, J. F. Cidermaking. In *Fermented Beverage Production*, 2nd ed.; Lea, A. G. H., Piggott, J. R., Eds.; Springer: Boston, MA, 2003; pp 59–87. https://doi.org/10.1007/978-1-4615-0187-9_4
25. Merwin, I. A.; Valois, S.; Padilla-Zakour, O. I. Cider Apples and Cider-Making Techniques in Europe and North America. In *Horticultural Reviews*; Janick, J., Ed.; Wiley: Hoboken, NJ, 2008; Vol. 34, pp 365–415. https://doi.org/10.1002/9780470380147.ch6
26. Lea, A. *Craft Cider Making*, 3rd ed.; The Crowood Press: Ramsbury, U.K., 2015.
27. Jolicoeur, C. *The New Cider Maker's Handbook: A Comprehensive Guide for Craft Producers*; Chelsea Green Publishing: White River Junction, VT, 2013.
28. Sanoner, P.; Guyot, S.; Marnet, N.; Molle, D.; Drilleau, J. F. Polyphenol Profiles of French Cider Apple Varieties (*Malus domestica* sp.). *Journal of Agricultural and Food Chemistry* **1999**, *47* (12), 4847–4853. https://doi.org/10.1021/jf990563y
29. Guyot, S.; Marnet, N.; Sanoner, P.; Drilleau, J. F. Variability of the Polyphenolic Composition of Cider Apple (*Malus domestica*) Fruits and Juices. *Journal of Agricultural and Food Chemistry* **2003**, *51* (21), 6240–6247. https://doi.org/10.1021/jf0301798
30. Thompson-Witrick, K. A.; Goodrich, K. M.; Neilson, A. P.; Hurley, E. K.; Peck, G. M.; Stewart, A. C. Characterization of the Polyphenol Composition of 20 Cultivars of Cider, Processing, and Dessert Apples (*Malus* × *domestica* Borkh.) Grown in Virginia. *Journal of Agricultural and Food Chemistry* **2014**, *62* (41), 10181–10191. https://doi.org/10.1021/jf503379t
31. Valois, S.; Merwin, I. A.; Padilla-Zakour, O. I. Characterization of Fermented Cider Apple Cultivars Grown in Upstate New York. *Journal of the American Pomological Society* **2006**, *60* (3), 113–128. https://doi.org/10.71318/apom.2006.60.3.113
32. Peck, G.; Miles, C. Assessing the Production Scale and Research and Extension Needs of U.S. Hard Cider Producers. *Journal of Extension* **2015**, *53* (5), Article 10. https://doi.org/10.34068/joe.53.05.18
33. Miles, C. A.; Alexander, T. R.; Peck, G.; Galinato, S. P.; Gottschalk, C.; van Nocker, S. Growing Apples for Hard Cider Production in the United States—Trends and Research Opportunities. *HortTechnology* **2020**, *30* (2), 148–155. https://doi.org/10.21273/horttech04488-19
34. Lee, C. Y.; Mattick, L. R. Composition and Nutritive Value of Apple Products. In *Processed Apple Products*; Downing, D. L., Ed.; Springer: New York, 1989; pp 303–322. https://doi.org/10.1007/978-1-4684-8225-6_14
35. Mattick, L. R.; Moyer, J. C. Composition of Apple Juice. *Journal of AOAC International* **1983**, *66* (5), 1251–1255. https://doi.org/10.1093/jaoac/66.5.1251
36. Fuleki, T.; Pelayo, E.; Palabay, R. B. Sugar Composition of Varietal Juices Produced from Fresh and Stored Apples. *Journal of Agricultural and Food Chemistry* **1994**, *42* (6), 1266–1275. https://doi.org/10.1021/jf00042a003
37. Karadeniz, F.; Ekşi, A. Sugar composition of apple juices. *European Food Research and Technology* **2002**, *215* (2), 145–148. https://doi.org/10.1007/s00217-002-0505-2
38. Eisele, T. A.; Drake, S. R. The partial compositional characteristics of apple juice from 175 apple varieties. *Journal of Food Composition and Analysis* **2005**, *18* (2–3), 213–221. https://doi.org/10.1016/j.jfca.2004.01.002
39. Loescher, W. H. Physiology and metabolism of sugar alcohols in higher plants. *Physiologia Plantarum* **1987**, *70* (3), 553–557. https://doi.org/10.1111/j.1399-3054.1987.tb02857.x
40. Hudina, M.; Štampar, F. Sugars and organic acids contents of European *Pyrus communis* L. and Asian *Pyrus serotina* Rehd. pear cultivars. *Acta Alimentaria* **2000**, *29* (3), 217–230. https://doi.org/10.1556/aalim.29.2000.3.2
41. Fuleki, T.; Pelayo, E.; Palabay, R. B. Carboxylic Acid Composition of Varietal Juices Produced from Fresh and Stored Apples. *Journal of Agricultural and Food Chemistry* **1995**, *43* (3), 598–607. https://doi.org/10.1021/jf00051a009

42. Wu, J.; Gao, H.; Zhao, L.; Liao, X.; Chen, F.; Wang, Z.; Hu, X. Chemical compositional characterization of some apple cultivars. *Food Chemistry* **2007**, *103* (1), 88–93. https://doi.org/10.1016/j.foodchem.2006.07.030

43. Ma, B.; Chen, J.; Zheng, H.; Fang, T.; Ogutu, C.; Li, S.; Han, Y.; Wu, B. Comparative assessment of sugar and malic acid composition in cultivated and wild apples. *Food Chemistry* **2015**, *172*, 86–91. https://doi.org/10.1016/j.foodchem.2014.09.032

44. Zhang, Y.; Li, P.; Cheng, L. Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh. *Food Chemistry* **2010**, *123* (4), 1013–1018. https://doi.org/10.1016/j.foodchem.2010.05.053

45. Hecke, K.; Herbinger, K.; Veberič, R.; Trobec, M.; Toplak, H.; Štampar, F.; Keppel, H.; Grill, D. Sugar-, acid- and phenol contents in apple cultivars from organic and integrated fruit cultivation. *European Journal of Clinical Nutrition* **2006**, *60* (9), 1136–1140. https://doi.org/10.1038/sj.ejcn.1602430

46. Hyson, D. A. A Comprehensive Review of Apples and Apple Components and Their Relationship to Human Health. *Advances in Nutrition* **2011**, *2* (5), 408–420. https://doi.org/10.3945/an.111.000513

47. Belton, P. S.; Delgadillo, I.; Gil, A. M.; Roma, P.; Casuscelli, F.; Colquhoun, I. J.; Dennis, M. J.; Spraul, M. High-field proton NMR studies of apple juices. *Magnetic Resonance in Chemistry* **1997**, *35* (13), S52–S60. https://doi.org/10.1002/(sici)1097-458x(199712)35:13<s52::aid-omr212>3.0.co;2-d

48. Belton, P. S.; Colquhoun, I. J.; Kemsley, E.; Delgadillo, I.; Roma, P.; Dennis, M.; Sharman, M.; Holmes, E.; Nicholson, J. K.; Spraul, M. Application of chemometrics to the ¹H NMR spectra of apple juices: discrimination between apple varieties. *Food Chemistry* **1998**, *61* (1–2), 207–213. https://doi.org/10.1016/s0308-8146(97)00103-9

49. Vogels, J. T. W. E.; Terwel, L.; Tas, A. C.; van den Berg, F.; Dukel, F.; van der Greef, J. Detection of Adulteration in Orange Juices by a New Screening Method Using Proton NMR Spectroscopy in Combination with Pattern Recognition Techniques. *Journal of Agricultural and Food Chemistry* **1996**, *44* (1), 175–180. https://doi.org/10.1021/jf950062m

50. French, R. K. *The History and Virtues of Cyder*; Robert Hale: London, 1982.

51. Jarvis, B. Cider (Cyder; Hard Cider). In *Encyclopedia of Food Microbiology*, 2nd ed.; Batt, C. A., Tortorello, M. L., Eds.; Academic Press: Oxford, 2014; pp 437–443. https://doi.org/10.1016/b978-0-12-384730-0.00066-5

52. Argia. Zagardoa: markesak pikaroa lau hankatan jarri zuenekoa (report on the documentation centre of Sagardoetxea, the Basque Cider Museum, Astigarraga). *Argia* No. 2809. https://www.argia.eus/en/argia-astekaria/2809/zagardoa-markesak-pikaroa-lau-hankatan-jarri-zuenekoa (accessed 2026-11-13).

53. National Geographic. Basque Country's Cider Houses Keep an Ancient History Alive. https://www.nationalgeographic.com/travel/article/basque-cider-houses-keep-cultural-and-culinary-history-alive (accessed 2026-11-13).

54. UNESCO Intergovernmental Committee for the Safeguarding of the Intangible Cultural Heritage. Decision 19.COM 7.B.7: Asturian Cider Culture, Inscription on the Representative List of the Intangible Cultural Heritage of Humanity; Asunción, December 2024. https://ich.unesco.org/en/decisions/19.COM/7.B.7 (accessed 2026-11-13).

55. Le Paulmier, J. *De Vino et Pomaceo Libri Duo*; G. Auvray: Paris, 1588 (French translation by J. de Cahaignes: *Traité du Vin et du Sidre*; Caen, 1589).

56. Le Quéré, J.-M.; Husson, F.; Renard, C. M. G. C.; Primault, J. French cider characterization by sensory, technological and chemical evaluations. *LWT – Food Science and Technology* **2006**, *39* (9), 1033–1044. https://doi.org/10.1016/j.lwt.2006.02.018

57. Evelyn, J. *Pomona, or an Appendix Concerning Fruit-Trees, in Relation to Cider; the Making and Several Ways of Ordering It*; published with *Sylva, or a Discourse of Forest-Trees*; J. Martyn and J. Allestry: London, 1664.

58. Worlidge, J. *Vinetum Britannicum: or, a Treatise of Cider, and Such Other Wines and Drinks That Are Extracted from All Manner of Fruits Growing in This Kingdom*; T. Dring and T. Burrel: London, 1676.

59. Baker, G. *An Essay Concerning the Cause of the Endemial Colic of Devonshire*; J. Hughs: London, 1767.

60. Picinelli, A.; Suárez, B.; Moreno, J.; Rodríguez, R.; Caso-García, L. M.; Mangas, J. J. Chemical Characterization of Asturian Cider. *Journal of Agricultural and Food Chemistry* **2000**, *48* (9), 3997–4002. https://doi.org/10.1021/jf991284d

61. del Campo, G.; Santos, J.; Berregi, I.; Velasco, S.; Ibarburu, I.; Dueñas, M. T.; Irastorza, A. Ciders Produced by Two Types of Presses and Fermented in Stainless Steel and Wooden Vats. *Journal of the Institute of Brewing* **2003**, *109* (4), 342–348. https://doi.org/10.1002/j.2050-0416.2003.tb00608.x

62. Alonso-Salces, R. M.; Guyot, S.; Herrero, C.; Berrueta, L. A.; Drilleau, J. F.; Gallo, B.; Vicente, F. Chemometric characterisation of Basque and French ciders according to their polyphenolic profiles. *Analytical and Bioanalytical Chemistry* **2004**, *379* (3), 464–475. https://doi.org/10.1007/s00216-004-2625-y

63. Picinelli Lobo, A.; Antón-Díaz, M. J.; Mangas Alonso, J. J.; Suárez Valles, B. Characterization of Spanish ciders by means of chemical and olfactometric profiles and chemometrics. *Food Chemistry* **2016**, *213*, 505–513. https://doi.org/10.1016/j.foodchem.2016.06.063

64. Coton, E.; Coton, M.; Guichard, H. Cider (Cyder; Hard Cider): The Product and Its Manufacture. In *Encyclopedia of Food and Health*; Caballero, B., Finglas, P. M., Toldrá, F., Eds.; Academic Press: Oxford, 2016; pp 119–128. https://doi.org/10.1016/b978-0-12-384947-2.00163-x

65. Rorabaugh, W. J. *The Alcoholic Republic: An American Tradition*; Oxford University Press: New York, 1979.

66. Proulx, A.; Nichols, L. *Cider: Making, Using & Enjoying Sweet & Hard Cider*, 3rd ed.; Storey Publishing: North Adams, MA, 2003.

67. Alcohol and Tobacco Tax and Trade Bureau. Cider FAQs (CID1-CID15). U.S. Department of the Treasury. https://www.ttb.gov/faqs/cider (accessed 2026-11-13).

68. Good Fruit Grower. Hard Cider Sales Show Regional Resiliency (interview with M. McGrath, American Cider Association). https://goodfruit.com/hard-cider-sales-show-regional-resiliency (accessed 2026-11-13).

69. Brewbound. CiderCon Data Talk: Cider to Reach Natural Equilibrium as National Brands Bow Out (regional cidery share of off-premise cider sales, 2019–2022). https://www.brewbound.com/news/cidercon-data-talk-cider-to-reach-natural-equilibrium-as-national-brands-bow-out-opportunity-for-cider-with-hard-seltzer-slowdown (accessed 2026-11-13).

70. Becot, F. A.; Bradshaw, T. L.; Conner, D. S. Apple Market Expansion through Value-added Hard Cider Production: Current Production and Prospects in Vermont. *HortTechnology* **2016**, *26* (2), 220–229. https://doi.org/10.21273/horttech.26.2.220

71. Tozer, P. R.; Galinato, S. P.; Ross, C. F.; Miles, C. A.; McCluskey, J. J. Sensory Analysis and Willingness to Pay for Craft Cider. *Journal of Wine Economics* **2015**, *10* (3), 314–328. https://doi.org/10.1017/jwe.2015.30

72. Calvert, M. D.; Neill, C. L.; Stewart, A. C.; Chang, E. A. B.; Whitehead, S. R.; Lahne, J. Appeal of the Apple: Exploring Consumer Perceptions of Hard Cider in the Northeast and Mid-Atlantic United States. *Journal of the American Society of Brewing Chemists* **2024**, *82* (3), 294–309. https://doi.org/10.1080/03610470.2023.2253707

73. Phetxumphou, K.; Cox, A. N.; Lahne, J. Development and Characterization of a Check-All-That-Apply (CATA) Lexicon for Virginia Hard (Alcoholic) Ciders. *Journal of the American Society of Brewing Chemists* **2020**, *78* (4), 299–307. https://doi.org/10.1080/03610470.2020.1768784

74. Kessinger, J.; Earnhart, G.; Hamilton, K. P.; Phetxumphou, K.; Neill, C.; Stewart, A. C.; Lahne, J. Exploring Perceptions and Categorization of Virginia Hard Ciders through the Application of Sorting Tasks. *Journal of the American Society of Brewing Chemists* **2021**, *79* (2), 187–200. https://doi.org/10.1080/03610470.2020.1843927

75. Littleson, B.; Chang, E.; Neill, C.; Phetxumphou, K.; Sandbrook, A.; Stewart, A.; Lahne, J. Sensory and Chemical Properties of Virginia Hard Cider: Effects of Apple Cultivar Selection and Fermentation Strategy. *Journal of the American Society of Brewing Chemists* **2023**, *81* (1), 141–154. https://doi.org/10.1080/03610470.2022.2057780

76. Suárez Valles, B.; Pando Bedriñana, R.; Fernández Tascón, N.; Querol Simón, A.; Rodríguez Madrera, R. Yeast species associated with the spontaneous fermentation of cider. *Food Microbiology* **2007**, *24* (1), 25–31. https://doi.org/10.1016/j.fm.2006.04.001

77. Morrissey, W.; Davenport, B.; Querol, A.; Dobson, A. The role of indigenous yeasts in traditional Irish cider fermentations. *Journal of Applied Microbiology* **2004**, *97* (3), 647–655. https://doi.org/10.1111/j.1365-2672.2004.02354.x

78. Coton, E.; Coton, M.; Levert, D.; Casaregola, S.; Sohier, D. Yeast ecology in French cider and black olive natural fermentations. *International Journal of Food Microbiology* **2006**, *108* (1), 130–135. https://doi.org/10.1016/j.ijfoodmicro.2005.10.016

79. Cousin, F.; Le Guellec, R.; Schlusselhuber, M.; Dalmasso, M.; Laplace, J. M.; Cretenet, M. Microorganisms in Fermented Apple Beverages: Current Knowledge and Future Directions. *Microorganisms* **2017**, *5* (3), 39. https://doi.org/10.3390/microorganisms5030039

80. Scanes, K.; Hohmann, S.; Prior, B. Glycerol Production by the Yeast *Saccharomyces cerevisiae* and its Relevance to Wine: A Review. *South African Journal of Enology and Viticulture* **1998**, *19* (1), 17–24. https://doi.org/10.21548/19-1-2239

81. Whiting, G. C. Organic Acid Metabolism of Yeasts during Fermentation of Alcoholic Beverages – A Review. *Journal of the Institute of Brewing* **1976**, *82* (2), 84–92. https://doi.org/10.1002/j.2050-0416.1976.tb03731.x

82. Romano, P.; Suzzi, G. Origin and Production of Acetoin during Wine Yeast Fermentation. *Applied and Environmental Microbiology* **1996**, *62* (2), 309–315. https://doi.org/10.1128/aem.62.2.309-315.1996

83. Antón, M. J.; Suárez Valles, B.; García Hevia, A.; Picinelli Lobo, A. Aromatic Profile of Ciders by Chemical Quantitative, Gas Chromatography-Olfactometry, and Sensory Analysis. *Journal of Food Science* **2014**, *79* (1), S92–S99. https://doi.org/10.1111/1750-3841.12323

84. Hang, Y. D.; Woodams, E. E. Influence of apple cultivar and juice pasteurization on hard cider and eau-de-vie methanol content. *Bioresource Technology* **2010**, *101* (4), 1396–1398. https://doi.org/10.1016/j.biortech.2009.09.069

85. Sánchez, A.; Rodríguez, R.; Coton, M.; Coton, E.; Herrero, M.; García, L. A.; Díaz, M. Population dynamics of lactic acid bacteria during spontaneous malolactic fermentation in industrial cider. *Food Research International* **2010**, *43* (8), 2101–2107. https://doi.org/10.1016/j.foodres.2010.07.010

86. Rosend, J.; Kaleda, A.; Kuldjärv, R.; Arju, G.; Nisamedtinov, I. The Effect of Apple Juice Concentration on Cider Fermentation and Properties of the Final Product. *Foods* **2020**, *9* (10), 1401. https://doi.org/10.3390/foods9101401

87. Calugar, P. C.; Coldea, T. E.; Salanţă, L. C.; Pop, C. R.; Pasqualone, A.; Burja-Udrea, C.; Zhao, H.; Mudura, E. An Overview of the Factors Influencing Apple Cider Sensory and Microbial Quality from Raw Materials to Emerging Processing Technologies. *Processes* **2021**, *9* (3), 502. https://doi.org/10.3390/pr9030502

88. Torres, A. P.; Oliveira, F. A. R.; Silva, C. L. M.; Fortuna, S. P. The Influence of pH on the Kinetics of Acid Hydrolysis of Sucrose. *Journal of Food Process Engineering* **1994**, *17* (2), 191–208. https://doi.org/10.1111/j.1745-4530.1994.tb00335.x

89. Ohtsuki, T.; Sato, K.; Sugimoto, N.; Akiyama, H.; Kawamura, Y. Absolute quantitative analysis for sorbic acid in processed foods using proton nuclear magnetic resonance spectroscopy. *Analytica Chimica Acta* **2012**, *734*, 54–61. https://doi.org/10.1016/j.aca.2012.04.033

90. Ye, M.; Yue, T.; Yuan, Y. Evolution of polyphenols and organic acids during the fermentation of apple cider. *Journal of the Science of Food and Agriculture* **2014**, *94* (14), 2951–2957. https://doi.org/10.1002/jsfa.6639

91. Alberti, A.; Machado dos Santos, T. P.; Ferreira Zielinski, A. A.; Eleutério dos Santos, C. M.; Braga, C. M.; Demiate, I. M.; Nogueira, A. Impact on chemical profile in apple juice and cider made from unripe, ripe and senescent dessert varieties. *LWT* **2016**, *65*, 436–443. https://doi.org/10.1016/j.lwt.2015.08.045

92. Laaksonen, O.; Kuldjärv, R.; Paalme, T.; Virkki, M.; Yang, B. Impact of apple cultivar, ripening stage, fermentation type and yeast strain on phenolic composition of apple ciders. *Food Chemistry* **2017**, *233*, 29–37. https://doi.org/10.1016/j.foodchem.2017.04.067

93. Real Decreto 72/2017, de 10 de febrero, por el que se aprueba la norma de calidad de las diferentes categorías de la sidra natural y de la sidra. *Boletín Oficial del Estado* **2017**, No. 44 (21 February), BOE-A-2017-1749. https://www.boe.es/buscar/pdf/2017/BOE-A-2017-1749-consolidado.pdf (accessed 2026-11-13).

94. Gobierno Vasco, Dirección de Calidad e Industrias Alimentarias. Pliego de Condiciones de la Denominación de Origen Protegida «Euskal Sagardoa / Sidra Natural del País Vasco» (product specification); application published in *Official Journal of the European Union* **2023**, *C 49*. https://www.euskadi.eus/informacion/calidad_diferenciada_dopigp/web01-a2elikal/es/ (accessed 2026-11-13).

95. Dueñas, M.; Irastorza, A.; Fernández, K.; Bilbao, A.; Huerta, A. Microbial Populations and Malolactic Fermentation of Apple Cider using Traditional and Modified Methods. *Journal of Food Science* **1994**, *59* (5), 1060–1064. https://doi.org/10.1111/j.1365-2621.1994.tb08190.x

96. Rodríguez Madrera, R.; Picinelli Lobo, A.; Suárez Valles, B. Phenolic Profile of Asturian (Spain) Natural Cider. *Journal of Agricultural and Food Chemistry* **2006**, *54* (1), 120–124. https://doi.org/10.1021/jf051717e

97. Zheng, J.; Wittouck, S.; Salvetti, E.; Franz, C. M. A. P.; Harris, H. M. B.; Mattarelli, P.; O'Toole, P. W.; Pot, B.; Vandamme, P.; Walter, J.; et al. A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *International Journal of Systematic and Evolutionary Microbiology* **2020**, *70* (4), 2782–2858. https://doi.org/10.1099/ijsem.0.004107

98. Sauvageot, N.; Gouffi, K.; Laplace, J. M.; Auffray, Y. Glycerol metabolism in *Lactobacillus collinoides*: production of 3-hydroxypropionaldehyde, a precursor of acrolein. *International Journal of Food Microbiology* **2000**, *55* (1–3), 167–170. https://doi.org/10.1016/s0168-1605(00)00191-4

99. Claisse, O.; Lonvaud-Funel, A. Assimilation of glycerol by a strain of *Lactobacillus collinoides* isolated from cider. *Food Microbiology* **2000**, *17* (5), 513–519. https://doi.org/10.1006/fmic.2000.0342

100. Garai-Ibabe, G.; Ibarburu, I.; Berregi, I.; Claisse, O.; Lonvaud-Funel, A.; Irastorza, A.; Dueñas, M. Glycerol metabolism and bitterness producing lactic acid bacteria in cidermaking. *International Journal of Food Microbiology* **2008**, *121* (3), 253–261. https://doi.org/10.1016/j.ijfoodmicro.2007.11.004

101. Vollenweider, S.; Lacroix, C. 3-Hydroxypropionaldehyde: applications and perspectives of biotechnological production. *Applied Microbiology and Biotechnology* **2004**, *64* (1), 16–27. https://doi.org/10.1007/s00253-003-1497-y

102. Claisse, O.; Lonvaud-Funel, A. Primers and a Specific DNA Probe for Detecting Lactic Acid Bacteria Producing 3-Hydroxypropionaldehyde from Glycerol in Spoiled Ciders. *Journal of Food Protection* **2001**, *64* (6), 833–837. https://doi.org/10.4315/0362-028x-64.6.833

103. de las Heras, E.; Zuriarrain-Ocio, A.; Zuriarrain, J.; Bordagaray, A.; Dueñas, M. T.; Berregi, I. Quantitative Determination of Acrolein in Cider by ¹H NMR Spectrometry. *Foods* **2020**, *9* (12), 1820. https://doi.org/10.3390/foods9121820

104. American Cider Association. What is "DRY"? American Cider Association Board Endorses Dryness Language. https://ciderassociation.org/?p=1478 (accessed 2026-11-13).

105. Strong, G.; Wilcox, M.; England, K. *BJCP Cider Style Guidelines (Including Perry), 2025 Edition*; Beer Judge Certification Program: St. Louis Park, MN, 2025. https://www.bjcp.org/beer-styles/introduction-to-cider-guidelines/ (accessed 2026-11-13).

106. Symoneaux, R.; Baron, A.; Marnet, N.; Bauduin, R.; Chollet, S. Impact of apple procyanidins on sensory perception in model cider (part 1): Polymerisation degree and concentration. *LWT – Food Science and Technology* **2014**, *57* (1), 22–27. https://doi.org/10.1016/j.lwt.2013.11.016

107. Symoneaux, R.; Chollet, S.; Bauduin, R.; Le Quéré, J.; Baron, A. Impact of apple procyanidins on sensory perception in model cider (part 2): Degree of polymerization and interactions with the matrix components. *LWT – Food Science and Technology* **2014**, *57* (1), 28–34. https://doi.org/10.1016/j.lwt.2014.01.007

108. Symoneaux, R.; Guichard, H.; Le Quéré, J. M.; Baron, A.; Chollet, S. Could cider aroma modify cider mouthfeel properties? *Food Quality and Preference* **2015**, *45*, 11–17. https://doi.org/10.1016/j.foodqual.2015.04.004

109. New York Cider Association. Orchard-Based Cider Dryness Scale; presentation at CiderCon 2018, Baltimore, MD, February 2018. https://ciderguide.com/wp-content/uploads/2018/02/NYCADrynessCiderCon-2018.pdf (accessed 2026-11-13).

110. Calvert, M.; Chang, E.; Lahne, J.; Stewart, A.; Neill, C. L.; Hamilton, A. M. *Effective Communication of Cider Sensory Quality: A White Paper*; Publication FST-455NP; Virginia Cooperative Extension: Blacksburg, VA, 2023. https://ext.vt.edu/content/dam/pubs_ext_vt_edu/FST/fst-455/FST-455.pdf (accessed 2026-11-13).

111. Calvert, M. D.; Cole, E.; Stewart, A. C.; Neill, C. L.; Lahne, J. Can Cider Chemistry Predict Sensory Dryness? Benchmarking the Merlyn Dryness Scale. *Journal of the American Society of Brewing Chemists* **2023**, *81* (4), 514–519. https://doi.org/10.1080/03610470.2022.2121562

112. Picchi, M.; Domizio, P.; Wilson, M.; Santos, J.; Orrin, F.; Zanoni, B.; Canuti, V. Chemical Characterization, Sensory Definition and Prediction Model of the Cider Dryness from New York State Apples. *Foods* **2023**, *12* (11), 2191. https://doi.org/10.3390/foods12112191

113. Décret n° 2025-162 du 20 février 2025 portant application de l'article L. 412-1 du code de la consommation en ce qui concerne les cidres et poirés. *Journal Officiel de la République Française*, 21 February 2025 (replaces Décret n° 53-978 du 30 septembre 1953).

114. UFC-Que Choisir. Cidre: doux, demi-sec, brut… comment choisir un cidre (buying guide; sugar contents of the designations under the decree of 1953). https://www.quechoisir.org/guide-d-achat-cidre-n63155/ (accessed 2026-11-13).
115. Commission Delegated Regulation (EU) 2019/33 of 17 October 2018 Supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council; Annex III, Part B (Terms Indicating the Sugar Content of Wine). *Official Journal of the European Union* **2019**, *L 9*, 2–45.
116. U.S. Food and Drug Administration. Nutrition Labeling of Food. *Code of Federal Regulations*, Title 21, Section 101.9; paragraph (c)(1)(i), caloric content and general factors, including the sugar alcohol factors of paragraph (c)(1)(i)(F). https://www.ecfr.gov/current/title-21/section-101.9 (accessed 2026-11-13).
117. Alcohol and Tobacco Tax and Trade Bureau. Industry Circular 2017-2: Changes to the Definition of Wine Eligible for the "Hard Cider" Tax Rate and New Labeling Requirements; see also *Code of Federal Regulations*, Title 27, Section 24.331. https://www.ttb.gov/industry-circulars/ttb-industry-circulars-17-2 (accessed 2026-11-13).
118. Alcohol and Tobacco Tax and Trade Bureau. TTB Procedure 2020-1: Testing of Calorie, Fat, Carbohydrate, and Protein Content of Alcohol Beverages; Acceptable Tolerance Levels; 28 September 2020 (supersedes TTB Procedure 2004-1; see also TTB Ruling 2013-2, Voluntary Nutrient Content Statements). https://www.ttb.gov/node/2472 (accessed 2026-11-13).
119. Alcohol and Tobacco Tax and Trade Bureau, Scientific Services Division. Official Method SSD:TM:400, Calories in Flavored Malt Beverages (and the companion methods SSD:TM:401 to 407). https://www.ttb.gov/images/pdfs/ssd/ssdtm400calsfmb.pdf (accessed 2026-11-13).
120. Alcohol and Tobacco Tax and Trade Bureau. Alcohol Facts Statements in the Labeling of Wines, Distilled Spirits, and Malt Beverages; Notice of Proposed Rulemaking (Notice No. 237). *Federal Register* **2025**, *90* (11), 6654; comment period extended to 15 August 2025 by Notice No. 239, *Federal Register* **2025**, *90* (65). https://www.govinfo.gov/content/pkg/FR-2025-01-17/html/2025-00957.htm (accessed 2026-11-13).
121. Merrill, A. L.; Watt, B. K. *Energy Value of Foods: Basis and Derivation*; Agriculture Handbook No. 74 (slightly revised); U.S. Department of Agriculture, Agricultural Research Service: Washington, DC, 1973.
122. Livesey, G. Health potential of polyols as sugar replacers, with emphasis on low glycaemic properties. *Nutrition Research Reviews* **2003**, *16* (2), 163–191. https://doi.org/10.1079/nrr200371
123. International Food Information Council. What Is Glycerin? 28 April 2020. https://ific.org/?p=17531 (accessed 2026-11-13).
124. Food and Agriculture Organization of the United Nations. *Food Energy – Methods of Analysis and Conversion Factors*; FAO Food and Nutrition Paper 77; FAO: Rome, 2003.
125. Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the Provision of Food Information to Consumers, Annex XIV (Conversion Factors). *Official Journal of the European Union* **2011**, *L 304*, 18–63.
126. Spraul, M.; Schütz, B.; Rinke, P.; Koswig, S.; Humpfer, E.; Schäfer, H.; Mörtter, M.; Fang, F.; Marx, U. C.; Minoja, A. NMR-Based Multi Parametric Quality Control of Fruit Juices: SGF Profiling. *Nutrients* **2009**, *1* (2), 148–155. https://doi.org/10.3390/nu1020148
127. Spraul, M.; Schütz, B.; Humpfer, E.; Mörtter, M.; Schäfer, H.; Koswig, S.; Rinke, P. Mixture analysis by NMR as applied to fruit juice quality control. *Magnetic Resonance in Chemistry* **2009**, *47* (S1), S130–S137. https://doi.org/10.1002/mrc.2528
128. Monakhova, Y. B.; Schütz, B.; Schäfer, H.; Spraul, M.; Kuballa, T.; Hahn, H.; Lachenmeier, D. W. Validation studies for multicomponent quantitative NMR analysis: the example of apple fruit juice. *Accreditation and Quality Assurance* **2014**, *19* (1), 17–29. https://doi.org/10.1007/s00769-013-1026-3
129. Ogrinc, N.; Košir, I. J.; Spangenberg, J. E.; Kidrič, J. The application of NMR and MS methods for detection of adulteration of wine, fruit juices, and olive oil. A review. *Analytical and Bioanalytical Chemistry* **2003**, *376* (4), 424–430. https://doi.org/10.1007/s00216-003-1804-6
130. Mannina, L.; Sobolev, A. P.; Viel, S. Liquid state ¹H high field NMR in food analysis. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2012**, *66*, 1–39. https://doi.org/10.1016/j.pnmrs.2012.02.001

131. Hatzakis, E. Nuclear Magnetic Resonance (NMR) Spectroscopy in Food Science: A Comprehensive Review. *Comprehensive Reviews in Food Science and Food Safety* **2019**, *18* (1), 189–220. https://doi.org/10.1111/1541-4337.12408

132. Sobolev, A. P.; Thomas, F.; Donarski, J.; Ingallina, C.; Circi, S.; Cesare Marincola, F.; Capitani, D.; Mannina, L. Use of NMR applications to tackle future food fraud issues. *Trends in Food Science & Technology* **2019**, *91*, 347–353. https://doi.org/10.1016/j.tifs.2019.07.035

133. Cuny, M.; Vigneau, E.; Le Gall, G.; Colquhoun, I.; Lees, M.; Rutledge, D. N. Fruit juice authentication by ¹H NMR spectroscopy in combination with different chemometrics tools. *Analytical and Bioanalytical Chemistry* **2008**, *390* (1), 419–427. https://doi.org/10.1007/s00216-007-1708-y

134. Marchetti, L.; Pellati, F.; Benvenuti, S.; Bertelli, D. Use of ¹H NMR to Detect the Percentage of Pure Fruit Juices in Blends. *Molecules* **2019**, *24* (14), 2592. https://doi.org/10.3390/molecules24142592

135. Le Gall, G.; Puaud, M.; Colquhoun, I. J. Discrimination between Orange Juice and Pulp Wash by ¹H Nuclear Magnetic Resonance Spectroscopy: Identification of Marker Compounds. *Journal of Agricultural and Food Chemistry* **2001**, *49* (2), 580–588. https://doi.org/10.1021/jf001046e

136. Sobolev, A.; Mannina, L.; Proietti, N.; Carradori, S.; Daglia, M.; Giusti, A.; Antiochia, R.; Capitani, D. Untargeted NMR-Based Methodology in the Study of Fruit Metabolites. *Molecules* **2015**, *20* (3), 4088–4108. https://doi.org/10.3390/molecules20034088

137. Tang, F.; Hatzakis, E. NMR-Based Analysis of Pomegranate Juice Using Untargeted Metabolomics Coupled with Nested and Quantitative Approaches. *Analytical Chemistry* **2020**, *92* (16), 11177–11185. https://doi.org/10.1021/acs.analchem.0c01553

138. Vermathen, M.; Marzorati, M.; Baumgartner, D.; Good, C.; Vermathen, P. Investigation of Different Apple Cultivars by High Resolution Magic Angle Spinning NMR. A Feasibility Study. *Journal of Agricultural and Food Chemistry* **2011**, *59* (24), 12784–12793. https://doi.org/10.1021/jf203733u

139. Vermathen, M.; Marzorati, M.; Diserens, G.; Baumgartner, D.; Good, C.; Gasser, F.; Vermathen, P. Metabolic profiling of apples from different production systems before and after controlled atmosphere (CA) storage studied by ¹H high resolution-magic angle spinning (HR-MAS) NMR. *Food Chemistry* **2017**, *233*, 391–400. https://doi.org/10.1016/j.foodchem.2017.04.089

140. Tomita, S.; Nemoto, T.; Matsuo, Y.; Shoji, T.; Tanaka, F.; Nakagawa, H.; Ono, H.; Kikuchi, J.; Ohnishi-Kameyama, M.; Sekiyama, Y. A NMR-based, non-targeted multistep metabolic profiling revealed L-rhamnitol as a metabolite that characterised apples from different geographic origins. *Food Chemistry* **2015**, *174*, 163–172. https://doi.org/10.1016/j.foodchem.2014.11.028

141. Eisenmann, P.; Ehlers, M.; Weinert, C.; Tzvetkova, P.; Silber, M.; Rist, M.; Luy, B.; Muhle-Goll, C. Untargeted NMR Spectroscopic Analysis of the Metabolic Variety of New Apple Cultivars. *Metabolites* **2016**, *6* (3), 29. https://doi.org/10.3390/metabo6030029

142. Iaccarino, N.; Varming, C.; Agerlin Petersen, M.; Viereck, N.; Schütz, B.; Toldam-Andersen, T. B.; Randazzo, A.; Balling Engelsen, S. Ancient Danish Apple Cultivars—A Comprehensive Metabolite and Sensory Profiling of Apple Juices. *Metabolites* **2019**, *9* (7), 139. https://doi.org/10.3390/metabo9070139

143. Francini, A.; Romeo, S.; Cifelli, M.; Gori, D.; Domenici, V.; Sebastiani, L. ¹H NMR and PCA-based analysis revealed variety dependent changes in phenolic contents of apple fruit after drying. *Food Chemistry* **2017**, *221*, 1206–1213. https://doi.org/10.1016/j.foodchem.2016.11.038

144. Berregi, I.; Santos, J. I.; del Campo, G.; Miranda, J. I.; Aizpurua, J. M. Quantitation determination of chlorogenic acid in cider apple juices by ¹H NMR spectrometry. *Analytica Chimica Acta* **2003**, *486* (2), 269–274. https://doi.org/10.1016/s0003-2670(03)00496-3

145. Berregi, I.; del Campo, G.; Caracena, R.; Miranda, J. I. Quantitative determination of formic acid in apple juices by ¹H NMR spectrometry. *Talanta* **2007**, *72* (3), 1049–1053. https://doi.org/10.1016/j.talanta.2006.12.031

146. del Campo, G.; Berregi, I.; Caracena, R.; Santos, J. I. Quantitative analysis of malic and citric acids in fruit juices using proton nuclear magnetic resonance spectroscopy. *Analytica Chimica Acta* **2006**, *556* (2), 462–468. https://doi.org/10.1016/j.aca.2005.09.039

147. del Campo, G.; Berregi, I.; Iturriza, N.; Santos, J. I. Ripening and Changes in Chemical Composition of Seven Cider Apple Varieties. *Food Science and Technology International* **2006**, *12* (6), 477–487. https://doi.org/10.1177/1082013206073009

148. Del Campo, G.; Santos, J. I.; Iturriza, N.; Berregi, I.; Munduate, A. Use of the ¹H Nuclear Magnetic Resonance Spectra Signals from Polyphenols and Acids for Chemometric Characterization of Cider Apple Juices. *Journal of Agricultural and Food Chemistry* **2006**, *54* (8), 3095–3100. https://doi.org/10.1021/jf051818c

149. Zuriarrain, A.; Zuriarrain, J.; Villar, M.; Berregi, I. Quantitative determination of ethanol in cider by ¹H NMR spectrometry. *Food Control* **2015**, *50*, 758–762. https://doi.org/10.1016/j.foodcont.2014.10.024

150. Zuriarrain, A.; Zuriarrain, J.; Puertas, A. I.; Dueñas, M. T.; Berregi, I. Quantitative determination of lactic and acetic acids in cider by ¹H NMR spectrometry. *Food Control* **2015**, *52*, 49–53. https://doi.org/10.1016/j.foodcont.2014.12.028

151. Cusano, E.; Simonato, B.; Consonni, R. Fermentation process of apple juice investigated by NMR spectroscopy. *LWT* **2018**, *96*, 147–151. https://doi.org/10.1016/j.lwt.2018.05.021

152. Cusano, E.; Cagliani, L. R.; Consonni, R.; Simonato, B.; Zapparoli, G. NMR-based metabolic profiling of different yeast fermented apple juices. *LWT* **2020**, *118*, 108771. https://doi.org/10.1016/j.lwt.2019.108771

153. Guyot, S.; Le Guernevé, C.; Marnet, N.; Drilleau, J. F. Methods for Determining the Degree of Polymerization of Condensed Tannins: A New ¹H-NMR Procedure Applied to Cider Apple Procyanidins. In *Plant Polyphenols 2*; Gross, G. G., Hemingway, R. W., Yoshida, T., Eds.; Springer: Boston, MA, 1999; pp 211–222. https://doi.org/10.1007/978-1-4615-4139-4_11

154. Carter, J. F.; Yates, H. S. A.; Tinggi, U. Stable Isotope and Chemical Compositions of European and Australasian Ciders as a Guide to Authenticity. *Journal of Agricultural and Food Chemistry* **2015**, *63* (3), 975–982. https://doi.org/10.1021/jf5030054

155. Godelmann, R.; Fang, F.; Humpfer, E.; Schütz, B.; Bansbach, M.; Schäfer, H.; Spraul, M. Targeted and Nontargeted Wine Analysis by ¹H NMR Spectroscopy Combined with Multivariate Statistical Analysis. Differentiation of Important Parameters: Grape Variety, Geographical Origin, Year of Vintage. *Journal of Agricultural and Food Chemistry* **2013**, *61* (23), 5610–5619. https://doi.org/10.1021/jf400800d

156. Godelmann, R.; Kost, C.; Patz, C. D.; Ristow, R.; Wachter, H. Quantitation of Compounds in Wine Using ¹H NMR Spectroscopy: Description of the Method and Collaborative Study. *Journal of AOAC International* **2016**, *99* (5), 1295–1304. https://doi.org/10.5740/jaoacint.15-0318

157. López-Rituerto, E.; Cabredo, S.; López, M.; Avenoza, A.; Busto, J. H.; Peregrina, J. M. A Thorough Study on the Use of Quantitative ¹H NMR in Rioja Red Wine Fermentation Processes. *Journal of Agricultural and Food Chemistry* **2009**, *57* (6), 2112–2118. https://doi.org/10.1021/jf803245r

158. Avenoza, A.; Busto, J. H.; Canal, N.; Peregrina, J. M. Time Course of the Evolution of Malic and Lactic Acids in the Alcoholic and Malolactic Fermentation of Grape Must by Quantitative ¹H NMR (qHNMR) Spectroscopy. *Journal of Agricultural and Food Chemistry* **2006**, *54* (13), 4715–4720. https://doi.org/10.1021/jf060778p

159. Viggiani, L.; Morelli, M. A. C. Characterization of Wines by Nuclear Magnetic Resonance: A Work Study on Wines from the Basilicata Region in Italy. *Journal of Agricultural and Food Chemistry* **2008**, *56* (18), 8273–8279. https://doi.org/10.1021/jf801513u

160. Duarte, I.; Barros, A.; Belton, P. S.; Righelato, R.; Spraul, M.; Humpfer, E.; Gil, A. M. High-Resolution Nuclear Magnetic Resonance Spectroscopy and Multivariate Analysis for the Characterization of Beer. *Journal of Agricultural and Food Chemistry* **2002**, *50* (9), 2475–2481. https://doi.org/10.1021/jf011345j

161. Nord, L. I.; Vaag, P.; Duus, J. Ø. Quantification of Organic and Amino Acids in Beer by ¹H NMR Spectroscopy. *Analytical Chemistry* **2004**, *76* (16), 4790–4798. https://doi.org/10.1021/ac0496852

162. Lachenmeier, D. W.; Frank, W.; Humpfer, E.; Schäfer, H.; Keller, S.; Mörtter, M.; Spraul, M. Quality control of beer using high-resolution nuclear magnetic resonance spectroscopy and multivariate analysis. *European Food Research and Technology* **2005**, *220* (2), 215–221. https://doi.org/10.1007/s00217-004-1070-7

163. Rodrigues, J. E.; Gil, A. M. NMR methods for beer characterization and quality control. *Magnetic Resonance in Chemistry* **2011**, *49* (S1), S37–S45. https://doi.org/10.1002/mrc.2844

164. Monakhova, Y. B.; Schäfer, H.; Humpfer, E.; Spraul, M.; Kuballa, T.; Lachenmeier, D. W. Application of automated eightfold suppression of water and ethanol signals in ¹H NMR to provide sensitivity for analyzing alcoholic beverages. *Magnetic Resonance in Chemistry* **2011**, *49* (11), 734–739. https://doi.org/10.1002/mrc.2823

165. Isaac-Lam, M. F. Determination of Alcohol Content in Alcoholic Beverages Using 45 MHz Benchtop NMR Spectrometer. *International Journal of Spectroscopy* **2016**, *2016*, 1–8. https://doi.org/10.1155/2016/2526946

166. Matviychuk, Y.; Haycock, S.; Rutan, T.; Holland, D. J. Quantitative analysis of wine and other fermented beverages with benchtop NMR. *Analytica Chimica Acta* **2021**, *1182*, 338944. https://doi.org/10.1016/j.aca.2021.338944

167. Monakhova, Y. B.; Kuballa, T.; Lachenmeier, D. W. Chemometric methods in NMR spectroscopic analysis of food products. *Journal of Analytical Chemistry* **2013**, *68* (9), 755–766. https://doi.org/10.1134/s1061934813090098

168. Smolinska, A.; Blanchet, L.; Buydens, L. M.; Wijmenga, S. S. NMR and pattern recognition methods in metabolomics: From data acquisition to biomarker discovery: A review. *Analytica Chimica Acta* **2012**, *750*, 82–97. https://doi.org/10.1016/j.aca.2012.05.049

169. Kemsley, E. K.; Defernez, M.; Marini, F. Multivariate statistics: Considerations and confidences in food authenticity problems. *Food Control* **2019**, *105*, 102–112. https://doi.org/10.1016/j.foodcont.2019.05.021

170. Riedl, J.; Esslinger, S.; Fauhl-Hassek, C. Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Analytica Chimica Acta* **2015**, *885*, 17–32. https://doi.org/10.1016/j.aca.2015.06.003

171. Franklin County Chamber of Commerce. 22nd Annual Franklin County CiderDays, 4 to 6 November 2016, Franklin County, MA (event announcement). https://chamber.franklincc.org/events/details/franklin-county-ciderdays-6220 (accessed 2026-11-13).

172. Orange Pippin Ltd. Orchard Directory: Second Chance Orchards, 7319 Kirkville Road, East Syracuse, NY 13057. https://www.orangepippin.com/orchards/united-states/new-york/second-chance-orchards (accessed 2026-11-13).

173. National Fruit Collection, Brogdale (Department for Environment, Food and Rural Affairs; University of Reading). Variety records of the apple collection. https://www.nationalfruitcollection.org.uk (accessed 2026-11-13).

174. Calhoun, C. L., Jr. *Old Southern Apples: A Comprehensive History and Description of Varieties for Collectors, Growers, and Fruit Enthusiasts*, revised and expanded ed.; Chelsea Green Publishing: White River Junction, VT, 2010.

175. Denisen, E. L.; Cott, A. E.; Pickett, B. S. Chieftain, a New Apple from Iowa. *Fruit Varieties and Horticultural Digest* **1966**, *20* (4), 62. https://doi.org/10.71318/apom.1966.20.4.62

176. Crosby, J. A.; Janick, J.; Pecknold, P. C.; Goffreda, J. C.; Korban, S. S. 'GoldRush' Apple. *HortScience* **1994**, *29* (7), 827–828. https://doi.org/10.21273/hortsci.29.7.827

177. Bernstein, H. J.; Pople, J. A.; Schneider, W. G. The Analysis of Nuclear Magnetic Resonance Spectra: I. Systems of Two and Three Nuclei. *Canadian Journal of Chemistry* **1957**, *35* (1), 67–83. https://doi.org/10.1139/v57-011

178. Hulme, A. C. The Isolation of L-Citramalic Acid from the Peel of the Apple Fruit. *Biochimica et Biophysica Acta* **1954**, *14*, 36–43. https://doi.org/10.1016/0006-3002(54)90127-4

179. Sugimoto, N.; Engelgau, P.; Jones, A. D.; Song, J.; Beaudry, R. Citramalate synthase yields a biosynthetic pathway for isoleucine and straight- and branched-chain ester formation in ripening apple fruit. *Proceedings of the National Academy of Sciences* **2021**, *118* (3), e2009988118. https://doi.org/10.1073/pnas.2009988118

180. Umino, M.; Onozato, M.; Sakamoto, T.; Koishi, M.; Fukushima, T. Analyzing Citramalic Acid Enantiomers in Apples and Commercial Fruit Juice by Liquid Chromatography–Tandem Mass Spectrometry with Pre-Column Derivatization. *Molecules* **2023**, *28* (4), 1556. https://doi.org/10.3390/molecules28041556

181. Di Matteo, G.; Spano, M.; Esposito, C.; Santarcangelo, C.; Baldi, A.; Daglia, M.; Mannina, L.; Ingallina, C.; Sobolev, A. P. NMR Characterization of Ten Apple Cultivars from the Piedmont Region. *Foods* **2021**, *10* (2), 289. https://doi.org/10.3390/foods10020289

182. Aprea, E.; Charles, M.; Endrizzi, I.; Corollaro, M. L.; Betta, E.; Biasioli, F.; Gasperi, F. Sweet taste in apple: the role of sorbitol, individual sugars, organic acids and volatile compounds. *Scientific Reports* **2017**, *7* (1), 44950. https://doi.org/10.1038/srep44950

183. Burroughs, L. F. The Amino-Acids of Apple Juices and Ciders. *Journal of the Science of Food and Agriculture* **1957**, *8* (3), 122–131. https://doi.org/10.1002/jsfa.2740080304

184. Wojdyło, A.; Oszmiański, J.; Laskowski, P. Polyphenolic Compounds and Antioxidant Activity of New and Old Apple Varieties. *Journal of Agricultural and Food Chemistry* **2008**, *56* (15), 6520–6530. https://doi.org/10.1021/jf800510j

185. Kschonsek, J.; Wolfram, T.; Stöckl, A.; Böhm, V. Polyphenolic Compounds Analysis of Old and New Apple Cultivars and Contribution of Polyphenolic Profile to the In Vitro Antioxidant Capacity. *Antioxidants* **2018**, *7* (1), 20. https://doi.org/10.3390/antiox7010020

186. Sousa, S.; Magalhães, A.; Ferreira, M. M. C. Optimized bucketing for NMR spectra: Three case studies. *Chemometrics and Intelligent Laboratory Systems* **2013**, *122*, 93–102. https://doi.org/10.1016/j.chemolab.2013.01.006

187. De Meyer, T.; Sinnaeve, D.; Van Gasse, B.; Tsiporkova, E.; Rietzschel, E. R.; De Buyzere, M. L.; Gillebert, T. C.; Bekaert, S.; Martins, J. C.; Van Criekinge, W. NMR-Based Characterization of Metabolic Alterations in Hypertension Using an Adaptive, Intelligent Binning Algorithm. *Analytical Chemistry* **2008**, *80* (10), 3783–3790. https://doi.org/10.1021/ac7025964

188. Bruce, S. D.; Higinbotham, J.; Marshall, I.; Beswick, P. H. An Analytical Derivation of a Popular Approximation of the Voigt Function for Quantification of NMR Spectra. *Journal of Magnetic Resonance* **2000**, *142* (1), 57–63. https://doi.org/10.1006/jmre.1999.1911

189. Weljie, A. M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C. M. Targeted Profiling: Quantitative Analysis of ¹H NMR Metabolomics Data. *Analytical Chemistry* **2006**, *78* (13), 4430–4442. https://doi.org/10.1021/ac060209g

190. Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in ¹H NMR Metabonomics. *Analytical Chemistry* **2006**, *78* (13), 4281–4290. https://doi.org/10.1021/ac051632c

191. van den Berg, R. A.; Hoefsloot, H. C.; Westerhuis, J. A.; Smilde, A. K.; van der Werf, M. J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **2006**, *7* (1), 142. https://doi.org/10.1186/1471-2164-7-142

192. Keun, H. C.; Ebbels, T. M.; Antti, H.; Bollard, M. E.; Beckonert, O.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Improved analysis of multivariate data by variable stability scaling: application to NMR-based metabolic profiling. *Analytica Chimica Acta* **2003**, *490* (1–2), 265–276. https://doi.org/10.1016/s0003-2670(03)00094-1

193. Mantel, N. The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research* **1967**, *27* (2), 209–220.

194. Ward, J. H. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* **1963**, *58* (301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

195. Sokal, R. R.; Rohlf, F. J. The Comparison of Dendrograms by Objective Methods. *Taxon* **1962**, *11* (2), 33–40. https://doi.org/10.2307/1217208

196. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **1987**, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

197. Hubert, L.; Arabie, P. Comparing partitions. *Journal of Classification* **1985**, *2* (1), 193–218. https://doi.org/10.1007/bf01908075

198. van der Maaten, L.; Hinton, G. Visualizing Data Using t-SNE. *Journal of Machine Learning Research* **2008**, *9*, 2579–2605.

199. McInnes, L.; Healy, J.; Saul, N.; Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* **2018**, *3* (29), 861. https://doi.org/10.21105/joss.00861

200. Venna, J.; Kaski, S. Neighborhood Preservation in Nonlinear Projection Methods: An Experimental Study. In *Artificial Neural Networks – ICANN 2001*; Dorffner, G., Bischof, H., Hornik, K., Eds.; Lecture Notes in Computer Science, Vol. 2130; Springer: Berlin, 2001; pp 485–491. https://doi.org/10.1007/3-540-44668-0_68

201. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **2001**, *58* (2), 109–130. https://doi.org/10.1016/s0169-7439(01)00155-1

202. Geladi, P.; Kowalski, B. R. Partial least-squares regression: a tutorial. *Analytica Chimica Acta* **1986**, *185*, 1–17. https://doi.org/10.1016/0003-2670(86)80028-9

203. Varma, S.; Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* **2006**, *7* (1), 91. https://doi.org/10.1186/1471-2105-7-91

204. Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E. J. J.; van Duijnhoven, J. P. M.; van Dorsten, F. A. Assessment of PLSDA cross validation. *Metabolomics* **2008**, *4* (1), 81–89. https://doi.org/10.1007/s11306-007-0099-6

205. Williams, P. C.; Sobering, D. Comparison of Commercial near Infrared Transmittance and Reflectance Instruments for Analysis of Whole Grains and Seeds. *Journal of Near Infrared Spectroscopy* **1993**, *1* (1), 25–32. https://doi.org/10.1255/jnirs.3

206. Rundlöf, T.; Mathiasson, M.; Bekiroglu, S.; Hakkarainen, B.; Bowden, T.; Arvidsson, T. Survey and qualification of internal standards for quantification by ¹H NMR spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis* **2010**, *52* (5), 645–651. https://doi.org/10.1016/j.jpba.2010.02.007

207. Weber, M.; Hellriegel, C.; Rueck, A.; Wuethrich, J.; Jenks, P. Using high-performance ¹H NMR (HP-qNMR) for the certification of organic reference materials under accreditation guidelines—Describing the overall process with focus on homogeneity and stability assessment. *Journal of Pharmaceutical and Biomedical Analysis* **2014**, *93*, 102–110. https://doi.org/10.1016/j.jpba.2013.09.007

208. Schoenberger, T. Determination of standard sample purity using the high-precision ¹H-NMR process. *Analytical and Bioanalytical Chemistry* **2012**, *403* (1), 247–254. https://doi.org/10.1007/s00216-012-5777-1

209. Saito, T.; Nakaie, S.; Kinoshita, M.; Ihara, T.; Kinugasa, S.; Nomura, A.; Maeda, T. Practical guide for accurate quantitative solution state NMR analysis. *Metrologia* **2004**, *41* (3), 213–218. https://doi.org/10.1088/0026-1394/41/3/015

210. Lawson, C. L.; Hanson, R. J. *Solving Least Squares Problems*; Society for Industrial and Applied Mathematics: Philadelphia, PA, 1995. https://doi.org/10.1137/1.9781611971217

211. Angyal, S. J. The Composition of Reducing Sugars in Solution. *Advances in Carbohydrate Chemistry and Biochemistry* **1984**, *42*, 15–68. https://doi.org/10.1016/s0065-2318(08)60122-5

212. Barclay, T.; Ginic-Markovic, M.; Johnston, M. R.; Cooper, P.; Petrovsky, N. Observation of the keto tautomer of D-fructose in D_2O using ¹H NMR spectroscopy. *Carbohydrate Research* **2012**, *347* (1), 136–141. https://doi.org/10.1016/j.carres.2011.11.003

213. Anderson, P. E.; Reo, N. V.; DelRaso, N. J.; Doom, T. E.; Raymer, M. L. Gaussian binning: a new kernel-based method for processing NMR spectroscopic data for metabolomics. *Metabolomics* **2008**, *4* (3), 261–272. https://doi.org/10.1007/s11306-008-0117-3

214. Craig, A.; Cloarec, O.; Holmes, E.; Nicholson, J. K.; Lindon, J. C. Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets. *Analytical Chemistry* **2006**, *78* (7), 2262–2267. https://doi.org/10.1021/ac0519312

215. Bro, R.; Smilde, A. K. Centering and scaling in component analysis. *Journal of Chemometrics* **2003**, *17* (1), 16–33. https://doi.org/10.1002/cem.773

216. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* **1987**, *2* (1–3), 37–52. https://doi.org/10.1016/0169-7439(87)80084-9

217. Bro, R.; Smilde, A. K. Principal component analysis. *Analytical Methods* **2014**, *6* (9), 2812–2831. https://doi.org/10.1039/c3ay41907j

218. Jolliffe, I. T.; Cadima, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2016**, *374* (2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

219. Barker, M.; Rayens, W. Partial least squares for discrimination. *Journal of Chemometrics* **2003**, *17* (3), 166–173. https://doi.org/10.1002/cem.785

220. Brereton, R. G.; Lloyd, G. R. Partial least squares discriminant analysis: taking the magic away. *Journal of Chemometrics* **2014**, *28* (4), 213–225. https://doi.org/10.1002/cem.2609

221. Cortes, C.; Vapnik, V. Support-vector networks. *Machine Learning* **1995**, *20* (3), 273–297. https://doi.org/10.1007/bf00994018

222. Breiman, L. Random Forests. *Machine Learning* **2001**, *45* (1), 5–32. https://doi.org/10.1023/a:1010933404324

223. Szymańska, E.; Saccenti, E.; Smilde, A. K.; Westerhuis, J. A. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* **2012**, *8* (S1), 3–16. https://doi.org/10.1007/s11306-011-0330-3

224. Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **2020**, *17* (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

225. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.

226. Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010**, *29* (9), 2176–2179. https://doi.org/10.1021/om100106e

227. Wishart, D. S.; Guo, A.; Oler, E.; Wang, F.; Anjum, A.; Peters, H.; Dizon, R.; Sayeeda, Z.; Tian, S.; Lee, B. L.; et al. HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Research* **2022**, *50* (D1), D622–D631. https://doi.org/10.1093/nar/gkab1062

228. Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; et al. BioMagResBank. *Nucleic Acids Research* **2008**, *36* (Database), D402–D408. https://doi.org/10.1093/nar/gkm957

229. Jensen, H. D.; Krogfelt, K. A.; Cornett, C.; Hansen, S. H.; Christensen, S. B. Hydrophilic Carboxylic Acids and Iridoid Glycosides in the Juice of American and European Cranberries (*Vaccinium macrocarpon* and *V. oxycoccos*), Lingonberries (*V. vitis-idaea*), and Blueberries (*V. myrtillus*). *Journal of Agricultural and Food Chemistry* **2002**, *50* (23), 6871–6874. https://doi.org/10.1021/jf0205110

230. Alcohol and Tobacco Tax and Trade Bureau. Standards of Identity for Wine: Class 5, Fruit Wine. *Code of Federal Regulations*, Title 27, Section 4.21(e); volatile acidity in paragraph (e)(3). https://www.ecfr.gov/current/title-27/section-4.21 (accessed 2026-11-13).

231. Brewbound. Stormalong Cider Introduces 2 New Styles, Transitions to 16 oz. Cans (product descriptions of Legendary Dry, Mass Appeal, Light of the Sun and Red Skies at Night). https://www.brewbound.com/news/stormalong-cider-introduces-2-new-styles-transitions-16-oz-cans (accessed 2026-11-13).

232. Clemens, R. A.; Jones, J. M.; Kern, M.; Lee, S.; Mayhew, E. J.; Slavin, J. L.; Zivanovic, S. Functionality of Sugars in Foods and Health. *Comprehensive Reviews in Food Science and Food Safety* **2016**, *15* (3), 433–470. https://doi.org/10.1111/1541-4337.12194

233. Brouns, F. Saccharide Characteristics and Their Potential Health Effects in Perspective. *Frontiers in Nutrition* **2020**, *7*, 75. https://doi.org/10.3389/fnut.2020.00075

234. Grembecka, M. Sugar alcohols—their role in the modern world of sweeteners: a review. *European Food Research and Technology* **2015**, *241* (1), 1–14. https://doi.org/10.1007/s00217-015-2437-7

235. Noble, A. C.; Bursick, G. F. The Contribution of Glycerol to Perceived Viscosity and Sweetness in White Wine. *American Journal of Enology and Viticulture* **1984**, *35* (2), 110–112. https://doi.org/10.5344/ajev.1984.35.2.110

236. Emwas, A. H.; Saccenti, E.; Gao, X.; McKay, R. T.; dos Santos, V. A. P. M.; Roy, R.; Wishart, D. S. Recommended strategies for spectral processing and post-processing of 1D ¹H-NMR data of biofluids with a particular focus on urine. *Metabolomics* **2018**, *14* (3), 31. https://doi.org/10.1007/s11306-018-1321-4

237. Worley, B.; Powers, R. Multivariate Analysis in Metabolomics. *Current Metabolomics* **2013**, *1* (1), 92–107. https://doi.org/10.2174/2213235x11301010092

238. Wattenberg, M.; Viégas, F.; Johnson, I. How to Use t-SNE Effectively. *Distill* **2016**, *1* (10), e2. https://doi.org/10.23915/distill.00002

239. Kobak, D.; Berens, P. The art of using t-SNE for single-cell transcriptomics. *Nature Communications* **2019**, *10* (1), 5416. https://doi.org/10.1038/s41467-019-13056-x

240. Gromski, P. S.; Muhamadali, H.; Ellis, D. I.; Xu, Y.; Correa, E.; Turner, M. L.; Goodacre, R. A tutorial review: Metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta* **2015**, *879*, 10–23. https://doi.org/10.1016/j.aca.2015.02.012

Summary

Material. Eleven overlapping workspaces containing 553 ¹H NMR spectrum records were reduced to 285 distinct spectra and to 96 ciders from 36 producers, each counted once, measured between 2013 and 2026 at 300 MHz as received and, for 87 of them, after freeze-drying. The juices of 56 apple varieties, 53 of them from one collection of old varieties, were measured after freeze-drying, and the origin of every variety was compiled.

Background. The history of cider is traced from the Carolingian estates through the Basque, Asturian, Norman and English traditions to the decline and revival of cider in North America. The cider-making process is set out with the constituents by which each step is recognized in the spectrum, and the natural ciders of the Basque Country and Asturias are described together with the bacterial conversion of glycerol to 1,3-propanediol.

Method. Peak areas were measured by integration, constrained line fitting and decomposition of the carbohydrate region with pure-component spectra, and concentrations were calculated against maleic acid as internal standard. The freeze-dried tube was placed on the scale of the native tube through the constituents common to both. In the juices malic acid was determined by a fit of its complete ABX multiplet together with asparagine and aspartic acid.

Apples. Relative to the three sugars the old varieties contain 59 % fructose, 13 % glucose and 28 % sucrose on average, with sucrose shares from 5 to 52 %, and 3.4 to 10.4 g of malic acid per 100 g of sugars; 48 of the 53 lots are sharps. Sorbitol varies more than tenfold and is highest in the russets. Citramalic acid and L-rhamnitol were assigned. PCA, hierarchical clustering, t-SNE and UMAP show a continuum governed by the sucrose share; PLS regression reproduces sugars, sorbitol and malic acid with ratios of performance to deviation of about 10 to 28.

Ciders. Alcohol 3.0 to 10.4 % v/v (median 5.9); residual sugar 0 to 102 g/L; malolactic conversion in 36 ciders; sucrose as the principal sugar in 13, ten of them mass-market; sorbic acid in 24; 1,3-propanediol from bacterial glycerol metabolism in nine, among them one Basque natural cider; two lots with unexpected malolactic conversion and one refermented can.

Dryness and further scales. American Cider Association: 27 dry, 12 semi-dry, 42 semi-sweet, 15 sweet. Spanish designations: 43 extra-seca, 18 seca, 24 semi-seca, 11 dulce. French thresholds of 1953: 59 brut, 19 demi-sec, 18 above the sugar content of doux. Seven further scales are proposed (balance, acidity, acid profile, body, aromatic index, alcohol tier, energy band); by the balance index 12 ciders are bone dry, 10 dry, 11 off-dry, 27 medium and 27 sweet.

Nutrition. Energy 115 to 247 kcal per 12 fl oz of the alcoholic ciders. Alcohol, carbohydrate and energy are ranked by brand for all 96 ciders, and nutrition panels are calculated from the spectra. One of eight declared energy values was exceeded by more than 20 %.

Chemometrics. Absolute intensities, mean-centering and Pareto scaling are recommended for cider; the choice between uniform bins and optimized buckets made no difference. The ciders form no natural clusters. Classes defined by the composition are recognized almost perfectly; market segment with a balanced accuracy above 0.8; adjuncts by their individual markers. PLS regression on buckets reproduces alcohol, sugars, principal acids and energy with ratios of performance to deviation above 7 and estimates total polyols from the native spectrum within 0.6 g/L.

Companion files. A qNMR workbook and a chemometrics workbook with live formulas, the consolidated workspaces of the cider and apple juice spectra, and the code accompany this document.