

Referenzproteome von fünf Weizenarten als Ausgangspunkt für die künftige Entwicklung von Sorten mit geringerem Allergenen Potenzial

Muhammad Afzal 1,5, Malte Sielaff^{2,5}, Ute Distler², Detlef Schuppan^{3,4,6}, Stefan Tenzer^{2,6} and C. Friedrich H. Longin^{1,6}

Weizen ist ein wichtiges Grundnahrungsmittel, und seine Verarbeitungsqualität hängt weitgehend von den Proteinen ab. Allerdings gibt es eine beträchtliche Anzahl von Menschen mit entzündlichen Reaktionen auf Weizenproteine, nämlich Zöliakie, Weizenallergie und das Syndrom der nicht zöliakischen Weizensensitivität. Daher sollten Proteom-Profile für die Akteure entlang der Weizenlieferkette von großer Bedeutung sein. Mit Hilfe der Flüssigchromatographie-Tandem Massenspektrometrie haben wir das Mehlsreferenzproteom für fünf alte bis moderne Weizensorten erstellt, die jeweils auf 10 in drei verschiedenen Umgebungen angebauten Sorten basieren. Wir identifizierten bei jeder Art mindestens 2540 Proteine, und eine Clusteranalyse trennte die Arten anhand ihrer Proteomprofile eindeutig voneinander. Mehr noch, mehr als 50 % der Proteine unterschieden sich signifikant zwischen den Arten - viele von ihnen spielen eine Rolle bei der Produktqualität, der Getreidestärkesynthese, der Regulierung von Pflanzenstress und nachgewiesenen oder potenziellen allergischen Reaktionen beim Menschen. Insbesondere wurde festgestellt, dass die Expression mehrerer wichtiger Weizenproteine hauptsächlich durch genetische Faktoren und nicht durch Umweltfaktoren bestimmt wird, was die Auswahl und Weiterentwicklung verbesserter Weizensorten für die Lieferkette ermöglicht, sofern schnelle Testmethoden entwickelt werden. Insbesondere Einkorn wies 5,4- bzw. 7,2-mal geringere Mengen an potenziellen Allergenen und immunogenen Amylase-Trypsin-Inhibitoren auf als Weichweizen, während der Gehalt an potenziellen Allergenen bei tetraploiden Weizenarten im Mittelfeld lag. Dies rechtfertigt dringend gezielte klinische Studien, bei denen die entwickelten Referenzproteome helfen werden, repräsentative Testnahrungen zu entwerfen. (npj Science of Food (2023) 7:9 ; <https://doi.org/10.1038/s41538-023-00188-0>)

Einleitung

Weizen ist eines der wichtigsten Grundnahrungsmittel mit einer weltweiten Produktion von 765 Millionen Tonnen im Jahr 2019 (<https://www.fao.org/faostat/en/#data/QCL>, abgerufen am 08.12.2021) und liefert 20 % der täglichen Zufuhr von Nahrungsprotein zusammen mit Ballaststoffen, Mineralien und Vitaminen¹. Der Großteil der Produktion entfällt auf die modernen Weizenarten Weichweizen (*Triticum aestivum* ssp. *aestivum*) und Hartweizen (*Triticum turgidum* ssp. *durum*). Während Weichweizen weltweit auf fast 223 Millionen Hektar (<https://apps.fas.usda.gov/psdonline/app/index.html#/app/advQuery>, Zugriff am 29.01.2022) für die Brotherstellung und die Tierernährung angebaut wird, werden mit Hartweizen weltweit 16 Millionen Hektar hauptsächlich für die Nudelherstellung genutzt². Obwohl die alten Weizenarten Dinkel (*Triticum aestivum* ssp. *spelta*), Emmer (*Triticum turgidum* ssp. *dicoccum*) und Einkorn (*Triticum monococcum* ssp. *monococcum*) schon seit Jahrtausenden als Nahrungsmittel genutzt werden^{3,4,5}, werden sie derzeit nur in kleinem Maßstab und nur in bestimmten Regionen angebaut^{6,7,8}. Weizenkörner enthalten ca. 8-15 % Protein des Trockengewichts⁹, das sich in Albumine/Globuline (15-20 %), einschließlich wichtiger essenzieller Aminosäuren¹⁰, und Kleberproteine (80-85 %)^{9,11} unterteilen lässt. Die viskoelastischen und geschmacklichen Eigenschaften als Qualitätsmerkmale für die Brot- und Teigwarenherstellung werden hauptsächlich von den Glutenproteinen bestimmt^{12,13}. Gleichzeitig können einige Weizenproteine Entzündungsreaktionen wie Zöliakie (CeD), klassische Weizenallergie (WA) und Nicht-Zöliakie-Weizensensitivität (NCWS) bei etwa 1 %, weniger als 1 % bzw. bis zu 10 % der Weizenkonsumenten auslösen¹⁴. Spezifische Glutenpeptidsequenzen verursachen CeD^{15,16}, Alpha-Amylase-/Trypsin-Inhibitoren (ATIs) stimulieren angeborene Immunzellen über den Toll-like-Rezeptor 4 (TLR4), um intestinale und extraintestinale Entzündungen in Tiermodellen der Krankheit zu fördern^{17, 18,19,20,21,22,23}, und insbesondere Serpine, Lipid Transfer-Proteine (LTPs), β -Amylasen, ATIs und einige Glutenproteine können durch Immunglobulin E (IgE) vermittelte allergische Reaktionen vom Soforttyp auslösen^{24,25,26,27,28,29}. Darüber hinaus deuten klinische und funktionelle Studien darauf hin, dass Lebensmittelallergien vom Typ 2, die z. B. durch Eosinophile und vor allem durch Weizenproteine ausgelöst werden, eine wichtige Rolle bei der Entstehung des Reizdarmsyndroms spielen^{14,30,31}. Frühere Studien verglichen moderne und alte Weizensorten und beschränkten sich auf Proteine einer bestimmten Familie wie Glutenproteine³² oder ATIs^{33,34,35}, oder untersuchten das immunogene Potenzial von ATIs zwischen Weichweizen und Einkorn^{21,36}. Im Vergleich zu früheren gelbasierten Proteomstudien ermöglichen die neuesten Entwicklungen in der LC-MS-basierten Proteomik die Quantifizierung von Tausenden von Proteinen in weniger als 1,5-2 Stunden pro Probe³⁷. Kürzlich wurde diese Proteomik-Technologie zum Vergleich der Proteome einiger Weizen-, Dinkel- und Roggensorten angewandt^{38,39} und es zeigte sich, dass eine große Anzahl von Proteinen sogar zwischen Weichweizen und Dinkel, beides hexaploide Arten, unterschiedlich exprimiert wurden. Beim Vergleich der Proteinexpressionsniveaus in 150 Weichweizensorten zeigte unsere jüngste Studie einen großen Einfluss der Umwelt und der verschiedenen Sorten auf die Expression einer Reihe von Proteinen⁴⁰. Unseres Wissens nach wurden die Mehlsproteome von Weichweizen, Dinkel, Hartweizen, Emmer und Einkorn jedoch noch nicht mit moderner Proteomik-Technologie verglichen. In der vorliegenden Arbeit haben wir hochauflösende Flüssigchromatographie-Tandem Massenspektrometrie (LC-MS/MS) auf der Grundlage markierungsfreier quantitativer Proteomik (LFQ) eingesetzt,

um das Proteom des Vollkornmehls von zehn Sorten für jede der fünf Weizenarten zu charakterisieren, die alle in drei verschiedenen Umgebungen angebaut wurden. Unsere Ziele waren (i) die Erstellung eines hochauflösenden Referenzproteoms von fünf Weizenarten, (ii) die Quantifizierung der Auswirkungen der Arten, der Sorten innerhalb der Arten und der Umwelt auf die Proteinhäufigkeit bzw. das Expressionsniveau und (iii) die Aufklärung von Ähnlichkeiten und Unterschieden in den Proteomen verschiedener Weizenarten auf der Grundlage von Proteinmustern im Zusammenhang mit Allergien, Immunaktivierung und Ernährungsqualität zur Verbesserung der Gesundheit und der Weizenlieferketten.

Ergebnisse und Diskussion

In unserer Analyse identifizierten wir 17.277 Peptidsequenzen und 2.896 verschiedene Proteine in 150 Mehlproben, was unseres Wissens nach die bisher größte Proteomstudie bei Getreide darstellt. Darüber hinaus ermöglichte die Untersuchung von zehn Sorten für jede Art, die in drei verschiedenen Umgebungen angebaut wurden, eine eingehende Bewertung der Auswirkungen der Art, der Sorten innerhalb der Art und der Umgebung auf die Proteinexpression.

Grundlage für künftige vertiefte proteomische Forschung bei Weizenarten

Bei unseren Proteomanalysen wurden 2706, 2705, 2671, 2687 und 2540 Proteine in Weichweizen, Dinkel, Durum, Emmer und Einkorn identifiziert (Abb. 1a). Interessanterweise waren diese Zahlen zwischen den Arten recht ähnlich, obwohl die Zusammensetzung der Proteinsequenzdatenbank, die für die Suche nach den MS-Spektren verwendet wurde, aufgrund des Fehlens von Referenzproteomen für einige der untersuchten Arten vor allem Einträge aus Weichweizen (38 % aller Einträge) und Hartweizen (51 % aller Einträge) enthielt. Diese Ergebnisse stimmen mit einer Studie überein, in der nur die hexaploiden Arten Weichweizen und Dinkel verglichen wurden³⁸, was auf eine hohe Sequenzhomologie zwischen den Weizenarten hinweist, ungeachtet der unterschiedlichen Ploidie-Stufen.

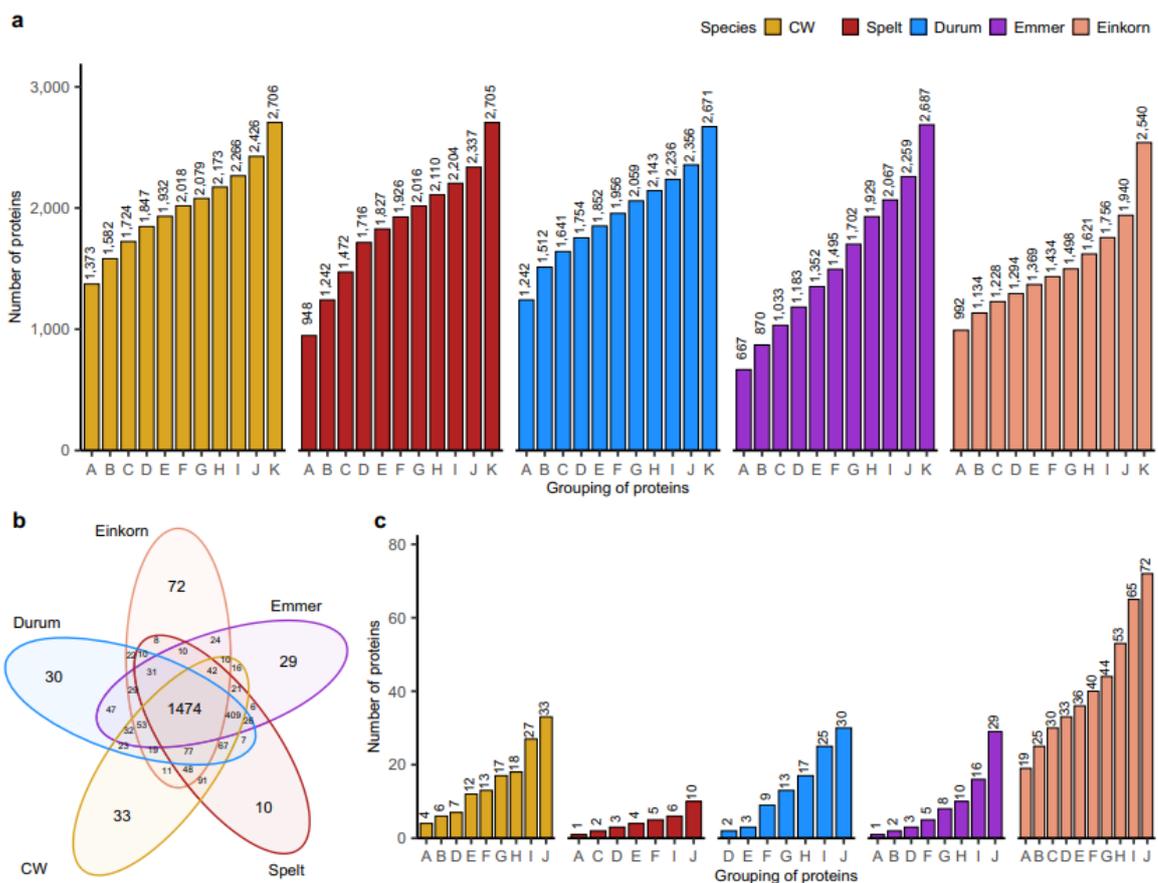


Fig. 1 Overview of the identified proteins in five wheat species. **a** Grouping of proteins based on their presence in a certain number of cultivars across three environments. The capital letters on the x-axis denote groups of proteins. These groups correspond to the proteins that were expressed in all of the three environments in all 10 cultivars (A), in at least 9 (B), 8 (C), 7 (D), 6 (E), 5 (F), 4 (G), 3 (H), 2 (I) cultivars and in at least 1 (J) cultivar. Proteins belonging to group K were found in at least one sample. For further analysis, the proteins from group J were considered. **b** Number of proteins that are unique or common between species. **c** Grouping of the unique proteins (from Fig. 1b) based on their presence in a certain number of cultivars across three environments. The letters on x-axis have the same denotation as in Fig. 1a. CW, common wheat.

Table 1. Potentially interesting proteins for future research and breeding grouped into protein families.

UniProt annotation / keyword	Common wheat	Spelt	Durum	Emmer	Einkorn
Glutenin, HMW, LMW	19 (0.81)	18 (0.78)	15 (0.85)	6 (0.69)	7 (0.78)
Gliadin	23 (0.81)	16 (0.81)	12 (0.81)	2 (0.64)	5 (0.82)
Puroindoline	1 (0.82)	0	0	0	2 (0.54)
Beta-amylase	7 (0.82)	4 (0.93)	2 (0.75)	1 (0.80)	1 (0.62)
1,4-alpha-glucan branching enzyme	3 (0.66)	2 (0.62)	7 (0.74)	2 (0.57)	3 (0.64)
Sucrose synthase, Sucrose_synth	3 (0.78)	3 (0.69)	5 (0.75)	1 (0.65)	0
Alpha-amylase inhibitor, amylase inhibitor	8 (0.71)	7 (0.72)	4 (0.79)	0	0
Trypsin inhibitor	2 (0.83)	3 (0.73)	1 (0.66)	0	1 (0.59)
Serpin	4 (0.75)	4 (0.76)	9 (0.85)	5 (0.60)	1 (0.57)
Non-specific lipid-transfer protein	6 (0.80)	5 (0.89)	4 (0.74)	0	1 (0.77)
Glutathione transferase, S-transferase	2 (0.74)	1 (0.70)	4 (0.80)	2 (0.69)	0
P:defense response ^a	15 (0.74)	11 (0.66)	15 (0.77)	10 (0.71)	8 (0.66)
P:response to water ^a	8 (0.69)	4 (0.60)	4 (0.78)	4 (0.74)	0
Heat shock	5 (0.64)	2 (0.57)	4 (0.73)	5 (0.76)	3 (0.58)
P:response to heat ^a	10 (0.74)	5 (0.65)	17 (0.70)	7 (0.75)	3 (0.63)
Late embryogenesis abundant (LEA)	2 (0.73)	0	0	0	0

HMW high molecular weight, LMW low molecular weight.
^aproteins involved in a given biological process (P) according to the gene ontology (GO) analysis.
 These proteins were filtered by applying a stringent criterion within each species, i.e., heritability >0.50, missing data ≤20%, detected in all environments in ≥50% cultivars and in at least 2 of 3 environments in ≥80% cultivars. For each species, the number of proteins within specific protein families fulfilling the filtering criteria are shown with their average heritability in brackets. For more details, see Supplementary Table1.

Insgesamt wurden die Proteinhäufigkeiten in hohem Maße durch die Auswahl der Sorten innerhalb der Arten und die Umgebung, in der die Sorten angebaut wurden, beeinflusst. So wurden beispielsweise von den insgesamt 2 540 identifizierten Proteinen in Einkorn (Abb. 1a Buchstabe K) nur 1 940 in allen drei Umgebungen in mindestens einer Sorte stabil exprimiert (Abb. 1a Buchstabe J): Somit waren 600 Proteine nur aufgrund von Umwelteinflüssen vorhanden/abwesend. Für die übrigen Proteine betrug die mittlere Heritabilität 0,24, wobei nur 380 Proteine eine Heritabilität von mehr als 0,5 aufwiesen (ergänzende Abb. 1a). Die Heritabilität quantifiziert den Einfluss der Sorte auf die Gesamtausprägung eines Merkmals und reicht von 0 bis 1. Je niedriger die Heritabilität, desto größer ist der Einfluss der Umwelt gegenüber der Sorte auf die Ausprägung eines Merkmals. Folglich ist der Einfluss der Umwelt, einschließlich des Bodens, der klimatischen Faktoren und der Anbaupraktiken, auf die Proteinexpression sehr hoch, was mit der Literatur übereinstimmt^{32,34,38,40}. Für die weitere Diskussion ließen wir Proteine außer Acht, die nur von der Umwelt beeinflusst wurden, insbesondere Proteine, die nicht in allen Umgebungen in mindestens einer Sorte einer Art stabil vorhanden waren (Abb. 1a Buchstabe K). Im Allgemeinen unterschieden sich die verschiedenen Kultivare innerhalb einer Art erheblich in ihrer Proteinexpression, was durch das Vorhandensein/Fehlen von Proteinen (Abb. 1a) oder einen großen Variationskoeffizienten zwischen den Kultivaren für ein Protein (ergänzende Abb. 1b) belegt wird. Zum Beispiel wurden bei Einkorn 1.940 Proteine identifiziert, die in mindestens einer Sorte in allen drei Umgebungen stabil vorhanden waren (Abb. 1a Buchstabe J), aber nur 992 Proteine waren in allen 10 Sorten und allen Umgebungen vorhanden (Abb. 1a Buchstabe A). Dies stimmt mit den Ergebnissen einer früheren Studie überein, in der Weichweizen und Dinkel verglichen wurden³⁸, und unterstreicht die Notwendigkeit, bei der Messung der Proteinhäufigkeiten repräsentative Sätze von Sorten innerhalb von Arten zu verwenden, die in verschiedenen Umgebungen angebaut wurden. In Anbetracht des großen Einflusses der Umwelt auf die Proteinexpression haben wir für die künftige Forschung und Züchtung eine Liste von Proteinen erstellt, die hauptsächlich von der Genetik beeinflusst werden. Diese Proteine könnten in zukünftigen Weizenlieferketten durch die Wahl der Sorte erfolgreich manipuliert werden. Wir wählten daher innerhalb jeder Art nur Proteine aus, die (i) eine Heritabilität >0,50 aufwiesen, (ii) fehlende Daten ≤20% hatten und (iii) in allen Umgebungen in ≥50% der Sorten und in mindestens 2 von 3 Umgebungen in ≥80% der Sorten nachgewiesen wurden. Es handelte sich um 845, 611, 863, 262 und 296 Proteine in Weichweizen-, Dinkel-, Hartweizen-, Emmer- bzw. Einkornsorten (ergänzende Tabelle 1). Diese Liste enthielt Proteine aus wichtigen Familien wie Proteine, die für die Backqualität entscheidend sind (Glutene, Gliadine), den Stärkeweg (Beta-Amylasen, Glucan-verzweigende Enzyme, Saccharose-Synthasen), bestätigte Allergene (Enzym-Inhibitoren, Serpine, Lipid-Transfer-Proteine) und die Reaktion der Pflanzen auf Feldbedingungen (Hitzeschock-, Hitze- und Trockenheitsreaktionsproteine, reichlich vorhandene Proteine der späten Embryogenese) und andere, die zum Teil in verschiedenen Weizenarten vorhanden waren (Tabelle 1). Darüber hinaus haben viele Proteine auf unserer Liste der "heißen Kandidatenproteine" für künftige Weizenversorgungsketten noch unbekannte oder eher beschreibende Namen, die dringende künftige Forschung rechtfertigen. Zusammenfassend lässt sich sagen, dass unsere aktuelle Proteomics-Studie mit hoher Abdeckung eine solide Grundlage für künftige eingehende Untersuchungen des Proteoms und der Proteinfunktionen bei verschiedenen Weizensorten bietet.

Fünf Weizenarten lassen sich anhand ihrer Proteinexpressionsmuster unterscheiden

Das hierarchische Clustering der 50 Sorten von fünf Weizenarten anhand von 2.774 Proteinen trennte die Sorten eindeutig in fünf Gruppen, die den fünf Arten entsprechen (Abb. 2). Das Clustering spiegelte den genetischen Abstand zwischen den Arten wider, indem es geringere Abstände zwischen Arten mit gleichem Ploidiegrad anzeigte. So lagen beispielsweise die beiden hexaploiden Arten Weichweizen und Dinkel enger beieinander als Einkorn und Weichweizen, ebenso wie die beiden tetraploiden Arten Hartweizen und Emmer, was die Validität unseres proteomischen Workflows unterstreicht.

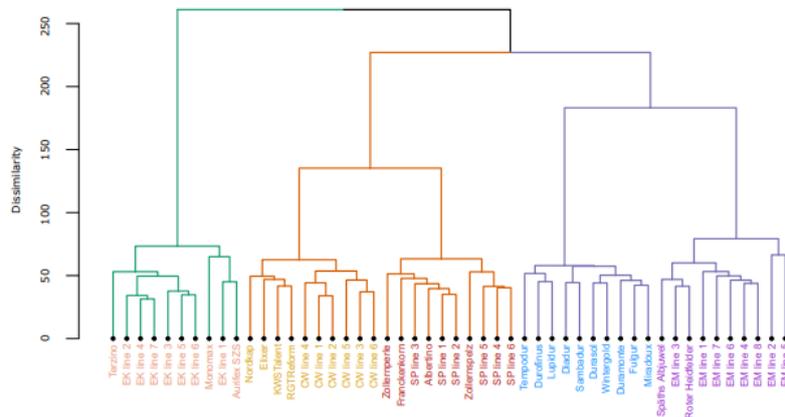


Fig. 2 Hierarchical clustering of 50 cultivars belonging to five wheat species. The colors of the lines (branches) highlight three clusters of cultivars corresponding to their ploidy level. The colors of the cultivar labels correspond to the color legend for species in Fig. 1a. In total, 2774 proteins were used for clustering, which included proteins unique to one of the five species and proteins common between species. All those proteins were identified across all three environments in at least one cultivar and hence belonged to the group J in Fig. 1a.

Diese Trennung wurde auch durch Proteine bestätigt, die nur in einer einzelnen Art vorkommen oder nur in wenigen, aber nicht in allen Arten vorhanden sind (Abb. 1b, c; ergänzende Abb. 2). So wies das diploide Einkorn die höchste Anzahl einzigartiger Proteine auf, wobei ≥ 40 einzigartige Proteine in mindestens vier Einkornsorten vorhanden waren. Im Gegensatz dazu wurden 1.474 Proteine in allen fünf Weizensorten gemeinsam exprimiert (Abb. 1b), wobei die höchste (niedrigste) Anzahl an gemeinsam exprimierten Proteinen bei paarweisen Vergleichen zwischen Weichweizen und Dinkel (Einkorn und Emmer; ergänzende Abb. 2) festgestellt wurde. Allerdings wurden mehr als 50 % der gemeinsamen Proteine zwischen einem beliebigen Artenpaar mit statistisch signifikant unterschiedlicher Häufigkeit exprimiert (Abb. 3), was die Tendenz zeigt, dass der Prozentsatz der unterschiedlich exprimierten Proteine umso höher ist, je größer der Unterschied zwischen den Ploidiestufen der Arten ist. So wiesen beispielsweise 52 % der gemeinsamen Proteine von Dinkel und Weichweizen (Abb. 3a) eine signifikant unterschiedliche Expression auf, die bei Einkorn gegenüber Weichweizen 78 % erreichte (Abb. 3g). Dabei wurde bei Einkorn im Vergleich zu den anderen Weizenarten eine größere Anzahl von Proteinen herunterreguliert als hochreguliert (ergänzende Abb. 3). Insgesamt 254 Proteine wiesen nicht nur eine statistisch unterschiedliche Expression zwischen den verschiedenen Artenpaaren auf, sondern übertrafen auch einen strengen Schwellenwert von $\pm 3 \log_2$ -facher Veränderung (entspricht einer 8-fachen Auf-/Abregulation) (Abb. 3 orangefarbene Punkte, ergänzende Tabelle 2). Diese Proteine gehörten zu den oben erwähnten wichtigen Proteinfamilien (Backqualität, Stärkeweg, Allergene, Stressreaktion der Pflanzen) (Tabelle 2). Abgesehen von unserem kürzlich durchgeführten Vergleich zwischen Weichweizen und Dinkel³⁸ ist bisher keine derartige Analyse veröffentlicht worden. Aufgrund des Fehlens von Referenzproteomen für Dinkel, Emmer und Einkorn könnte unsere Studie insofern verzerrt sein, als Proteine, die für diese Arten einzigartig sind, nicht entdeckt wurden, was die Unterschiede zwischen den Arten noch vergrößern könnte. Bislang bietet unsere Studie jedoch die größte Proteomabdeckung hinsichtlich der identifizierten und quantifizierten Proteine bei fünf Weizenarten.

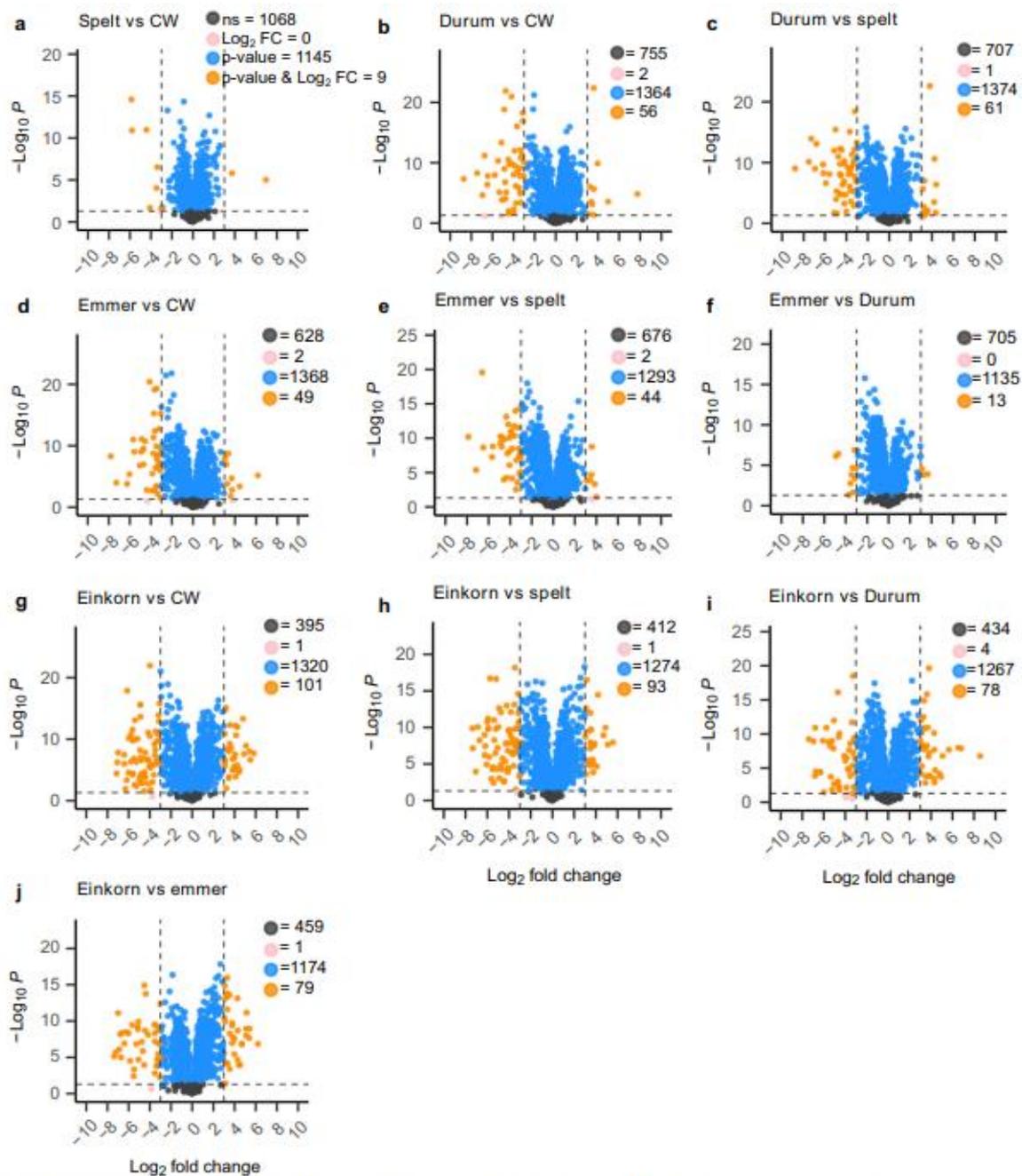


fig. 3 Results of the comparisons for individual proteins between different pairs of wheat species. **a–j** Volcano plots of proteins that were common between different pairs of wheat species. The x-axis shows the log₂-scale fold change (Log_2FC) in the LFQ abundance of proteins between two species. The y-axis shows the negative log₁₀-scale p value from the t-test. The vertical dashed lines denote $\pm 3 \text{Log}_2\text{FC}$ threshold, while the horizontal dashed line denotes p value threshold of 0.05, e.g., the orange points in (a) show proteins, which are downregulated ($\text{Log}_2\text{FC} < -3$) or upregulated ($\text{Log}_2\text{FC} > 3$) in spelt compared to common wheat and passed both thresholds i.e., p value and Log_2FC . Color of the points represents significance status of proteins for p value (< 0.05) and Log_2FC (± 3): gray, non-significant for both thresholds; pink, significant Log_2FC only; blue, significant p value only; orange, significant for both thresholds. CW, common wheat.

Table 2. Proteins with potentially important functions, which were significantly differently expressed (t-test, *p* value <0.05) between species.

UniProt annotation / keyword	NPS	GO annotations ^a	NPU				
			Cw	Sp	Du	Em	Ei
Glutenin, HMW, LMW	50	F:nutrient reservoir activity	2	-	1	1	3
Gliadin	41	F:nutrient reservoir activity	6	2	2	-	4
Puroindoline	7	F:nutrient reservoir activity	-	-	-	-	-
Beta-amylase	7	F:beta-amylase activity	1	-	-	1	-
1,4-alpha-glucan branching enzyme	12	P:glycogen biosynthetic process	1	-	1	-	-
Starch synthase	12	11 F:glycogen (starch) synthase activity	-	-	1	-	1
Sucrose synthase, Sucrose_synth	8	F:sucrose synthase activity	-	-	-	-	-
Alpha-amylase inhibitor, amylase inhibitor, Trypsin inhibitor, BBTI, CTI	30	F:serine-type endopeptidase inhibitor activity	1	-	-	-	1
Serpin	16	F:serine-type endopeptidase inhibitor activity	1	-	-	-	-
Nonspecific lipid-transfer protein	10	P:lipid transport	-	-	-	-	-
ATP synthase	14	P:ATP synthesis coupled proton transport	-	-	-	-	-
Glutathione transferase, S-transferase	8	7 P:glutathione metabolic process, 1 F:glutathione dehydrogenase (ascorbate) activity	-	-	-	-	-
Defensin, Barwin, Thaumatin-like, Thionin, Knot1	13	P:defense response	-	-	1	-	2
Dehydrin	4	P:response to water	-	1	-	-	-
Heat shock, SHSP	42	F:ATP hydrolysis activity, P:response to heat	4	-	7	3	6
Late embryogenesis abundant, LEA	4	-	-	-	-	-	-

NPS number of proteins, which were significantly differently expressed (t-test, *p* value <0.05) between species, *GO* gene ontology, *NPU* number of proteins, which were unique to only one of the five species, *Cw* common wheat, *Sp* spelt, *Du* durum, *Em* emmer, *Ei* einkorn, *HMW* high molecular weight, *LMW* low molecular weight, *BBTI* Bowman-Birk type trypsin inhibitor, *CTI* chymotrypsin inhibitor.
For more details, see Supplementary Table 2.
^aGO annotations provided as either molecular function (F:) or biological process (P:) obtained from the Gene Ontology (GO) analysis.

Potenzielle und bekannte allergene Proteine sind in Einkorn weitgehend reduziert

Weizen ist ein wichtiges und in der Regel gesundes Grundnahrungsmittel für Mensch und Tier, aber eine beträchtliche Anzahl von Menschen leidet unter entzündlichen Weizensensibilitäten. Dabei handelt es sich um Zöliakie, IgE-vermittelte und nicht IgE-vermittelte (Typ 2) Weizenallergie und angeborene Immunaktivierung durch ATI-Proteine, wobei die beiden letzteren möglicherweise zur Nicht-Zöliakie-Weizensensitivität (NCWS) beitragen^{14,17,18,19,20,21,22,23,25,30,31,41,42,43,44,45}. Da es sich bei den meisten potenziellen Allergenen um Proteine handelt und große Unterschiede in den Proteomen der verschiedenen Weizenspezies bestehen (siehe oben), haben wir die Verteilung potenziell allergener Proteine auf die verschiedenen Weizenspezies genauer untersucht. Wir folgten dem Ansatz von Zimmermann et al.³⁹ und stellten eine Liste von Allergenen zusammen, die auf den Daten über samenbürtige Weizenallergene²⁷ und der Allergome Database (<http://www.allergome.org/index.php>)⁴⁶ basiert, und bezogen zusätzlich ATIs mit ein. Die Summe all dieser potenziell allergenen Proteine unterschied sich deutlich zwischen den Arten und entsprach nahezu perfekt dem Ploidiegrad (Abb. 4a). Während ihre Gesamthäufigkeit bei hexaploidem Weichweizen und Dinkel ähnlich war, waren sie bei tetraploidem Durum und Emmer etwa um das Zweifache und bei diploidem Einkorn um das 5,4-fache reduziert. Diese Unterschiede sind sowohl auf die unterschiedliche Anzahl potenzieller Allergene als auch auf unterschiedliche Proteinhäufigkeiten zurückzuführen. Im Gegensatz dazu war der Gesamtproteingehalt des Korns (GPC) bei Einkorn, Dinkel und Emmer etwas höher als bei Weichweizen (ergänzende Abb. 4), was mit früheren Erkenntnissen^{32,47} übereinstimmt, wobei die Unterschiede im GPC der anderen Arten im Vergleich zu Weichweizen zwischen 1 % und 20 % lagen. Bei Weichweizen, Dinkel und Hartweizen wies fast die Hälfte der potenziell allergenen Proteine eine Heritabilität >0,5 auf, und bei allen Arten lag der Variationskoeffizient mit einer Heritabilität >0,5 zwischen 7 % und 261 % (Abb. 4a). Daher kann die Häufigkeit dieser potenziellen Allergene mit Hilfe der Proteomik gezielt reduziert werden, um die Züchtung und die Sortenwahl zu überwachen, was die jüngsten Erkenntnisse über die Gluten- und ATI-Zusammensetzung bei verschiedenen Weizensorten bestätigt^{33,34,48}. Dies würde jedoch die Entwicklung von Schnelltestverfahren erfordern, die im täglichen Geschäft in der gesamten Weizenlieferkette eingesetzt werden können. Unser Referenzproteom kann als Ausgangspunkt dienen, z. B. indem wir uns auf Allergene mit hoher Heritabilität und hohem Variationskoeffizienten zwischen den Sorten innerhalb einer Art konzentrieren.

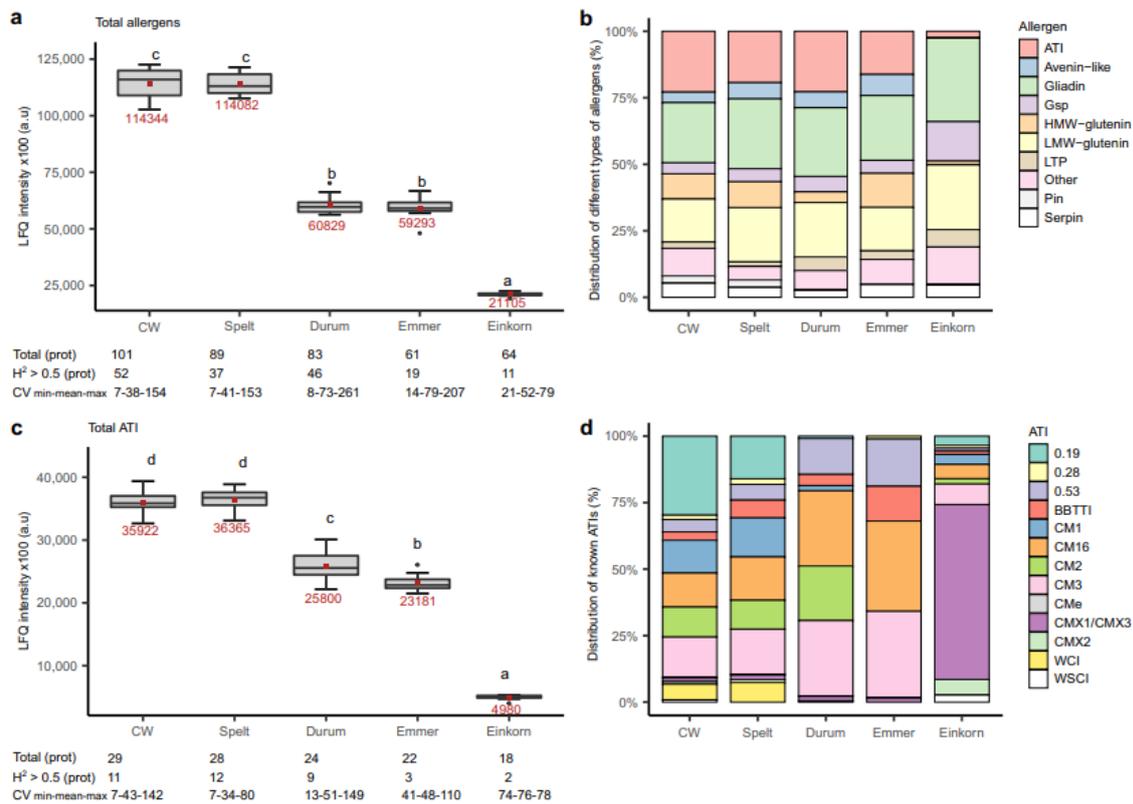


Fig. 4 Allergenic proteins in wheat species. LFO intensity/abundance of different families of proteins per cultivar in different wheat species for (a) potential allergens, and (c) proteins identified as ATIs. Boxes with different letters indicate significant differences between the wheat species (ANOVA with Tukey's test, $p < 0.05$). Each box summarizes the content of 30 samples (ten cultivars grown at three locations). Mean values for each species are highlighted in red color. Furthermore, within each species the distribution of (b) allergens, and (d) ATIs with a known type is shown. CW, common wheat; Total (prot), number of allergenic proteins or ATIs used to calculate total LFO intensity (%); H² > 0.5 (prot), number of allergenic proteins or ATIs with heritability >0.5; H², heritability; CV_{min-mean-max}, minimum-mean-maximum values of the coefficient of variation (%) for allergenic proteins or ATIs with heritability >0.5; LFO, label-free quantification; ATI, amylase/trypsin inhibitor; Gsp, grain softness protein; HMW, high molecular weight; LMW, low molecular weight; LTP, lipid transfer protein; Pin, puroindoline; BBTI, Bowman-Birk type trypsin inhibitor; CM, chloroform/methanol soluble protein; WCI, wheat chymotrypsin inhibitor; WSCI, wheat subtilisin-chymotrypsin inhibitor.

Neben der Menge der Allergene variierte auch ihre Verteilung innerhalb der Arten erheblich (Abb. 4b). Während die Allergene des Weichweizens größtenteils aus ATIs, Gliadinen, HMW- und LMW Gluteninen bestanden, waren mehr als 50 % der Einkornallergene (wesentlich geringere Mengen als bei anderen Weizenarten, Abb. 4a) Gliadine und LMW-Glutene. In Übereinstimmung mit der veröffentlichten Literatur^{33,48} wies Einkorn im Vergleich zu den anderen Weizenarten eine deutlich geringere Menge an ATIs auf, nämlich 7,2, 7,3, 5,2 bzw. 4,7-mal weniger als in Weichweizen, Dinkel, Durum und Emmer (Abb. 4c). Die ATIs werden als Hauptallergene und auch als Aktivatoren von TLR4 in Tiermodellen von Krankheiten beschrieben^{17,18,19,21,22,23,33,48}. Die ATIs in Einkorn wurden hauptsächlich CMX1/CMX3 (UniProt-Zugang M8A1S2) zugeordnet (Abb. 4d). Interessanterweise wurde für Einkorn nicht beschrieben, dass die vorherrschenden ATIs vom Typ CMX1/CMX2/CMX3 die Amylaseaktivität hemmen⁴⁹, und CMX1/CMX3 sind weder als samenbürtige Weizenallergene²⁷ noch in der Allergome Database (<http://www.allergome.org/index.php>)⁴⁶ aufgeführt. Darüber hinaus wurde ein weiteres ATI (UniProt-Zugang C5J3R4; Beschreibung, Trypsin-Inhibitor OS = *Triticum monococcum* subsp. *monococcum* OX = 408188 GN = Eti-Am1) nur in Emmer und Einkorn gefunden, was unsere jüngsten Analysen mit Schwerpunkt auf ATIs bestätigt³⁵. Im Gegensatz dazu waren die ATIs von hexaploidem Weichweizen und Dinkel hauptsächlich durch 0,19, CM1, CM2, CM3 und CM16 repräsentiert, während die ATIs von tetraploidem Durum und Emmer hauptsächlich 0,53, CM3 und CM16 waren (Abb. 4d). Iacomino et al.³⁶ zeigten, dass die ATIs von Einkorn anfälliger für enzymatische In-vitro-Hydrolyse sind als die ziemlich Pepsin-Trypsin-resistenten ATIs von tetra- und hexaploiden Weizen^{19,21,22,23,50} und daher während der Lebensmittelverarbeitung und insbesondere während der oberen Magen-Darm-Passage weitgehend proteolytisch abgebaut werden, was dazu führt, dass sie keine oder eine geringere Fähigkeit zur Auslösung der angeborenen Immunität besitzen³⁶. In ähnlicher Weise haben Sievers et al.⁵¹ gezeigt, dass Einkorn für Menschen, die nur auf ATIs empfindlich reagieren, vorteilhaft sein könnte, aber für Personen mit einer Weizenallergie möglicherweise nicht sicher ist. Zusammenfassend lässt sich sagen, dass unsere Studie eine wesentlich geringere Häufigkeit potenzieller Allergene und ATIs in Einkorn sowie in Durum und Emmer im Vergleich zu hexaploidem Weichweizen und Dinkel zeigt. Allergendatenbanken scheinen jedoch mehr Allergene von Weichweizen als von anderen Weizenarten zu enthalten, was wahrscheinlich auf die begrenzte Verwendung und den Anbau alternativer Weizenarten wie Dinkel, Emmer und Einkorn zurückzuführen ist. Außerdem enthielt die verwendete Proteom-Referenz nur die Referenzproteome von *T. turgidum* ssp. *durum* (51 % aller Datenbankeinträge), *T. aestivum* ssp. *aestivum*

(38 % aller Datenbankeinträge) und T. urartu (9 % aller Datenbankeinträge). Folglich hat unser Ansatz möglicherweise nicht ausgereicht, um alle potenziellen Allergene von Einkorn, Emmer, Durum und Dinkel zu identifizieren. Dennoch vermuten wir, dass Emmer, Durum und Einkorn eine wesentlich geringere Anzahl und Häufigkeit allergener Proteine aufweisen, da die identifizierten Unterschiede sehr groß sind und die Sequenzhomologie hoch war, so dass ungeachtet der oben genannten Einschränkungen eine große Anzahl von Proteinen in allen Weizenarten identifiziert werden konnte. Die Notwendigkeit, alternative Getreidearten mit geringerer Allergenität oder ATIs zu finden, wird durch die zunehmende Zahl präklinischer und insbesondere klinischer Studien zum Vergleich verschiedener Weizensorten unterstrichen. Erste explorative präklinische Studien deuten auf eine potenziell bessere Verträglichkeit von alten gegenüber modernen Weizensorten bei Patienten mit Weizenallergie und NCWS hin. In diesen Studien wurden alte diploide und auch tetraploide Weizensorten wie Einkorn und Emmer von NCWS-Patienten, von denen viele entweder eine IgE positive oder eine IgE-negative (Typ 2) Weizenallergie hatten, tatsächlich besser vertragen als moderne hexaploide Weizensorten^{18,52,53,54,55,56,57,58}. In ähnlicher Weise untersuchten Picascia et al.⁵⁹ die Auswirkungen einer aus Einkorn- und Weichweizenmehl zubereiteten Diät auf die Immunantwort bei Zöliakiepatienten und kamen zu dem Schluss, dass Einkorn im Vergleich zu Weichweizen eine geringere In-vivo-T-Zell-Reaktion hervorruft. Diese Studien basierten jedoch nicht auf repräsentativen Proben, die verschiedene Sorten verschiedener Weizenarten enthielten, die unter vergleichbaren Umweltbedingungen angebaut wurden, die alle einen großen Einfluss auf die proteomischen Profile der Mehlproben haben, wie in dieser Studie gezeigt wurde. Folglich sind gezieltere kontrollierte klinische Studien mit Weizen mit einem definierten niedrigen Gehalt an potenziellen Allergenen dringend erforderlich und können auf der Grundlage der in dieser Studie gelieferten Referenzproteome durchgeführt werden. Neben Einkorn und anderen Weizensorten mit geringem Gehalt an potenziellen Allergenen sollte auch die Auswirkung verschiedener Mehlverarbeitungs- und Brotbackverfahren, wie z. B. der langen Sauerteiggärung, auf den Gehalt und die Aktivität von Allergenen untersucht werden, um gesündere Weizenprodukte insbesondere für Menschen mit weizenbedingten Erkrankungen zu finden.

Ausblick: Einkorn als nachhaltige Kulturpflanze für marginale Umgebungen, aber mögliche gesundheitliche Vorteile müssen dringend validiert werden

Neben der oben beschriebenen geringeren Menge an potenziellen Allergenen, für die es noch keine klinischen Nachweise gibt, enthält Einkorn im Vergleich zu Weichweizen mehr Eiweiß und deutlich höhere Mengen an wertvollen Spurenstoffen wie Vitamin E, Lutein, Sterylferulat^{60,61}, Mineralien wie Fe und Zn⁶² sowie verschiedene andere Mineralien wie Ca, Cu, K, Mg, Mn, P und S⁶³ - allesamt Verbindungen, die für eine gesunde Ernährung wichtig sind^{61,62}. Die Bioverfügbarkeit dieser Verbindungen in Brot und anderen Getreideerzeugnissen und ihre Auswirkungen auf die menschliche Gesundheit sind jedoch noch nicht untersucht worden, so dass weitere Forschungen dringend erforderlich sind. Aus agronomischer Sicht weist Einkorn eine wesentlich höhere Eiweißausbeute als Weichweizen³², eine fast vollständige Resistenz gegen Pilze⁸ und die Flexibilität auf, dieselbe Sorte vor oder nach dem Winter anzusäen, was bei anderen Getreidearten nicht der Fall ist. Im Vergleich zu Weichweizen sind Einkornpflanzen jedoch höher und damit anfälliger für Lagerschäden als Weichweizen. Außerdem hat Einkorn unter guten Bodenbedingungen einen um fast 70 % geringeren Kornertrag als Weichweizen⁴⁷. Angesichts der Notwendigkeit, die landwirtschaftliche Produktivität pro verfügbare Anbaufläche zu steigern, um eine wachsende Weltbevölkerung zu ernähren, kann Einkorn den weit verbreiteten Weichweizen nicht ersetzen. In Grenzertragslagen ist die Produktivität von Weichweizen jedoch deutlich geringer⁶⁴, während Einkorn gute Ergebnisse erzielt⁶⁵. Zu den Grenzertragsstandorten gehören sandige Böden und höhere Lagen in Gebirgsregionen und/oder der Mangel an Stickstoffdünger aufgrund von hohen Kosten, Umweltauflagen oder ökologischer Landwirtschaft. In Anbetracht des geringsten Allergen- und ATI-Gehalts und des hohen Gehalts an nährstoffreichen Inhaltsstoffen von Einkorn sind klinische Studien dringend erforderlich, um diese potenziellen gesundheitlichen Vorteile zu validieren, da Einkorn eine vielversprechende, nachhaltige alternative Kulturpflanze für Randgebiete sein könnte, die die Agrobiodiversität fördert.

Methoden

Pflanzenmaterial und Feldversuche

Wir untersuchten jeweils 10 Sorten der fünf Weizenarten Weichweizen (*T. aestivum* ssp. *aestivum*, $2n = 6 \times = 42$, AuAuBBDD), Dinkel (*T. aestivum* ssp. *spelta*, $2n = 6 \times = 42$, AuAuBBDD), Durum (*T. turgidum* ssp. *durum*, $2n = 4 \times = 28$, AuAuBB), Emmer (*T. turgidum* ssp. *dicoccum*, $2n = 4 \times = 28$, AuAuBB) und Einkorn (*T. monococcum* ssp. *monococcum*, $2n = 2 \times = 14$, AmAm). Unter den ausgewählten Sorten der einzelnen Weizenarten befinden sich sehr wichtige Sorten, die den aktuellen Markt in Deutschland repräsentieren. Darüber hinaus wurden den ausgewählten Sorten die besten und neuesten Zuchtlinien aus Feldversuchen unter verschiedenen Bedingungen hinzugefügt (ergänzende Tabelle 3). Die Feldversuche wurden als Winteranbau, d.h. Aussaat im Oktober 2018 und Ernte im Juli 2019, an drei verschiedenen Standorten für jede Art in

Deutschland/Österreich durchgeführt. Die Versuchsstandorte mit ihren GPS-Koordinaten für jede Art sind in Klammern angegeben, gefolgt vom Namen der Art, z. B. Weichweizen (DSV-Leutewitz - 51°8'58"N, 13°21'46.908 "E; Eckartsweier - 48°31'45 "N, 7°51'18 "E; Stuttgart-Hohenheim - 48°42'50 "N, 9°12'58 "E), Dinkel (Eckartsweier - 48°31'45 "N, 7°51'18 "E; Stuttgart-Hohenheim - 48°42'50 "N, 9°12'58 "E; Oberer Lindenhof - 48°28'26 "N, 9°18'12 "E), Hartweizen (Eckartsweier - 48°31'45 "N, 7°51'18 "E; Stuttgart-Hohenheim - 48°42'50 "N, 9°12'58 "E; Probstdorf - 48°10'19 "N, 16°37'13 "E), Emmer (Stuttgart-Hohenheim - 48°42'50 "N, 9°12'58 "E; Ihingerhof - 48°44'44 "N, 8°55'23 "E; Oberer Lindenhof - 48°28'26 "N, 9°18'12 "E), Einkorn (Eckartsweier - 48°31'45 "N, 7°51'18 "E; Ihingerhof - 48°44'44 "N, 8°55'23 "E; Oberer Lindenhof - 48°28'26 "N, 9°18'12 "E). Die Weizenarten wurden in getrennten, aber benachbarten Versuchen an einzelnen Standorten untersucht, wobei ein nicht repliziertes Felddesign verwendet wurde, das für jede Art an jedem Versuchsstandort separat randomisiert wurde. In allen Versuchen wurden die gleichen Feldbehandlungen mit intensiven konventionellen Anbaumethoden durchgeführt, mit Ausnahme der Stickstoffdüngung, bei der Weichweizen, Hartweizen, Dinkel, Emmer und Einkorn gedüngt wurden, um die folgenden N-Gehalte einschließlich des im Boden gemessenen N zu erreichen ("Nmin"): 180, 180, 160, 60 bzw. 60 kg N/ha. Diese individuelle Anpassung wurde vorgenommen, um der aktuellen landwirtschaftlichen Praxis im konventionellen Anbau Rechnung zu tragen. Außerdem wurde Einkorn aufgrund seiner Feldresistenz gegen Pilzkrankheiten im Gegensatz zu allen anderen Arten nicht mit Fungiziden behandelt. Die Größe der Feldnetzparzellen betrug an allen Standorten 5 m². Alle Parzellen wurden maschinell gesät und mit Mähdreschern geerntet. Alle Proben von Dinkel, Emmer und Einkorn wurden mit einem Mini-Petkus-Saatgutreiniger (Röber, Bad Oeynhausen, Deutschland) geschält und gereinigt, um Spelzen, Stroh und beschädigte Körner zu entfernen. Die Schälung erfolgte mit einer klassischen Steinmühle, bei der der Stein durch Hartgummi ersetzt wurde. Bei Weichweizen wurde außerdem eine Saatgutreinigung mit dem Mini Petkus-Saatgutreiniger durchgeführt, um Spreu und Strohhpartikel zu entfernen, die nach der Mähdruschernte noch vorhanden waren.

Laboruntersuchungen

Drei Beobachtungen (Proben) von drei verschiedenen Standorten wurden verwendet, um die mittlere Häufigkeit eines Proteins pro Sorte jeder Art zu berechnen. Für die Laboranalyse verwendeten wir jedoch für jede Probe eine technische Wiederholung. Obwohl die Verwendung mehrerer technischer Replikate vorzuziehen ist, mussten wir aufgrund der großen Anzahl von Proben, die innerhalb eines bestimmten Budgets und Zeitlimits analysiert werden mussten, einen Kompromiss eingehen. Dieser Ansatz ist gerechtfertigt, da aus zahlreichen Studien zu verschiedenen Merkmalen in Feldversuchen bekannt ist und von der wissenschaftlichen Gemeinschaft akzeptiert wird, dass eine große Varianz in den Daten aufgrund von Unterschieden in den Bedingungen zwischen verschiedenen Anbaustandorten entsteht. Angesichts dieser standortabhängigen Variabilität verbessert sich die Datenqualität und wird repräsentativer für allgemeine Aussagen über die Ausprägung eines Merkmals, wie z. B. Proteine bei verschiedenen Arten, wenn die Anzahl der Standorte auf Kosten der Anzahl der technischen Wiederholungen erhöht wird und nicht umgekehrt. Darüber hinaus wurde die geringe Anzahl der technischen Wiederholungen bei der statistischen Analyse zur Schätzung der Mittelwerte über drei Standorte hinweg berücksichtigt, indem sie um Feldversuche und Effekte von Laboranalysen bereinigt wurde.

Protein-Extraktion

Von jeder Weizensorte (Weichweizen, Dinkel, Hartweizen, Emmer und Einkorn) wurden 20 mg Vollkornmehl in 1,5-mL-Kunststoffröhrchen (Protein LoBind Röhrchen, Eppendorf, Hamburg, Deutschland) eingewogen. Anschließend wurden 50 µL Wasser in LC-MS-Qualität hinzugefügt und die Röhrchen geschüttelt, bis das Mehl vollständig resuspendiert war. Unmittelbar danach wurden 950 µL Extraktionspuffer, bestehend aus 7 M Harnstoff, 2 M Thioharnstoff, 2 % (w/v) CHAPS, 5 mM Dithiothreitol (DTT) und LC-MS-Wasser, zugegeben und die Röhrchen erneut geschüttelt. Nach einer 10-minütigen Inkubation bei 22 °C wurden die Proben 10 Minuten lang bei 16.000 x g und 22 °C zentrifugiert. Die klare obere Schicht des Überstandes wurde für alle weiteren Schritte verwendet. Die durchschnittlichen Proteinkonzentrationen wurden aus den Extrakten speziesspezifischer Mehlmischungen mit dem Pierce 660 nm Protein Assay (Thermo Scientific, Rockford, IL, USA; ergänzende Abb. 4) bestimmt.

Probenvorbereitung

Die Proteine wurden gereinigt und zu LC-MS-kompatiblen tryptischen Peptiden verdaut, wobei ein filtergestütztes Probenvorbereitungsprotokoll (FASP)^{66,67} wie zuvor beschrieben⁶⁸ mit geringfügigen Änderungen verwendet wurde. Kurz gesagt wurden 30 µg Proteinextrakte auf Zentrifugal-Ultrafiltrationsgeräte (Nanosep mit 30 K MWCO Omega-Membran, Pall, Port Washington, NY, USA) geladen und bei 16.000 x g und 22 °C 15-30 Minuten lang zentrifugiert, bis die Flüssigkeit vollständig durch die Membran lief. Die Disulfidbindungen wurden mit DTT reduziert und anschließend die freien Cysteine mit Iodacetamid (IAA) alkyliert. Die Alkylierungsreaktion wurde durch Zugabe von DTT gequencht. Nach jedem Schritt wurde die Membran einmal mit einem Puffer aus 8 M Harnstoff und 100 mM Tris-HCl (pH 8,5) gewaschen. Schließlich wurde die Membran dreimal mit einem Puffer gewaschen, der 50 mM Ammoniumbicarbonat und LC-MS-Wasser enthielt. Der tryptische Verdau auf dem Filter wurde durch Zugabe von Trypsin (Trypsin Gold, Promega, Madison, WI, USA) in einem Protease-zu-

Protein-Verhältnis von 1:50 (w/w) und Inkubation über Nacht bei 37 °C durchgeführt. Die tryptischen Peptide wurden durch Zentrifugation und zusätzliches Waschen der Membran mit 50 mM Ammoniumbicarbonat in frischen Röhrchen gesammelt. Der Durchfluss wurde mit Trifluoressigsäure (TFA) angesäuert, um eine Endkonzentration von 0,5 % (v/v) zu erreichen. Die Peptide wurden auf Sep-Pak tC18-Kartuschen mit 96 Vertiefungen (Waters Corporation, Milford, MA, USA) geladen und unter Verwendung von 0,1 % (v/v) TFA in LC-MS-Wasser als Waschlösungsmittel und 0,1 % (v/v) TFA in 50 % (v/v) Acetonitril/Wasser als Elutionslösungsmittel entsalzt. Gereinigte Peptide wurden gefriergetrocknet und vor der LC-MS-Analyse in 20 µL 0,1 % (v/v) Ameisensäure (FA) in LC-MS-Wasser rekonstituiert.

Flüssigchromatographie-Massenspektrometrie

Die tryptischen Peptide jeder Probe wurden nacheinander mittels Flüssigchromatographie-Tandem Massenspektrometrie (LC-MS/MS) unter Verwendung eines nanoACQUITY UPLC-Systems (Waters Corporation) analysiert, das über eine NanoLockSpray-Dual-Elektrospray-Ionisationsquelle (Waters Corporation) an ein SYNAPT G2-S-Massenspektrometer (Waters Corporation) gekoppelt war. Die Microflow-LC und die Schnittstelle zur Quelle wurden wie zuvor beschrieben⁶⁹ eingerichtet. Peptide in Mengen von 1,5 µL wurden auf eine HSS-T3 C18 Reversed-Phase-Säule (Waters Corporation) mit einer Länge von 250 mm und einem Innendurchmesser von 300 µm geladen und durch Gradientenelution mit LC-Flussraten von 5 µL/min und 60 min LC-Methoden getrennt. Vorläufer- und Fragment-Ionen-Massenspektren wurden mit einer ionenbewegungsunterstützten, datenunabhängigen Aufnahme-strategie (UDMSE)⁶⁸ aufgenommen.

Datenverarbeitung und markierungsfreie Quantifizierung

Die UDMSE-Rohdaten wurden mit ProteinLynx Global Server v3.0.2 (PLGS, Waters Corporation) verarbeitet und mit einer Datenbank abgeglichen, die aus Proteinen der Gattung *Triticum* (UniProtKB Release 2020_05, Taxon-ID: 4564, 367.831 Einträge) zusammengestellt wurde und die Referenzproteome von *T. turgidum* ssp. *durum* (51 % aller Datenbankeinträge), *T. aestivum* ssp. *aestivum* (38 % aller Datenbankeinträge) und *T. urartu* (9 % aller Datenbankeinträge) sowie 171 übliche MS-Verunreinigungen unter Verwendung folgender Parameter: Trypsin wurde als Verdauungsenzym festgelegt, zwei verpasste Spaltungen pro Peptid wurden für die anfängliche Datenbanksuche zugelassen, die Carbamidomethylierung von Cysteinen wurde als feste und die Methioninoxidation als variable Modifikation festgelegt. Die Falschdeckungsrate (FDR) wurde in PLGS berechnet, indem eine Datenbank mit umgekehrten Proteinsequenzen durchsucht und ein Cutoff-Wert von 0,01 angewandt wurde. Die markierungsfreie Quantifizierung (LFQ), einschließlich Retentionszeit-Alignment, Feature Clustering, Cross-Run-Normalisierung und Protein-Inferenz, wurde mit ISOQuant v1.868 durchgeführt. Für die Quantifizierung wurden nur Peptide ohne verpasste Spaltungen, mit einer Mindestsequenzlänge von sieben Aminosäuren, einem PLGS-Score von mindestens 6,0 und ohne variable Modifikation berücksichtigt. Auf Peptid- und Proteinebene wurde in ISOQuant ein FDR Cutoff von 0,01 angewendet, so dass auf Datensatzebene ein FDR von 1 % gewährleistet war. Proteine, die durch mindestens zwei verschiedene Peptide identifiziert wurden, wurden quantifiziert, indem die Intensitäten der drei Peptide mit den höchsten Intensitäten, die zu dem jeweiligen Protein gehören, gemittelt wurden (Top3-Methode)⁷⁰. Die Häufigkeiten der gemeinsam genutzten Peptide wurden auf der Grundlage der relativen Häufigkeiten der eindeutig zugeordneten Peptide auf die Proteine umverteilt (Einzelheiten siehe ISOQuant-Handbuch; <http://www.immunologie.uni-mainz.de/isoquant/index.php?slab=user-manual#x1-400006.7>). Die TOP3-basierte Quantifizierung liefert einen Schätzwert für die Gesamtmenge jedes Proteins in einer Probe. Durch Aufsummieren aller detektierten und quantifizierten Proteine kann die relative Menge jedes Proteins im jeweiligen Proteom (d. h. Teile pro Million des Gesamtproteins) bestimmt werden. Dieser Wert ist unabhängig von der Gesamtmenge des Proteins in der Probe oder auf der Säule.

Datenbanksuche und Homologiefilterung - detaillierte Vorgehensweise

Bei der PLGS-Datenbanksuche wird jedes erkannte Peptid allen Proteinen in der Datenbank zugeordnet, die das jeweilige Peptid enthalten. Dies wird zunächst in PLGS laufend durchgeführt. Anschließend werden bei der Datenverarbeitung in ISOQuant die Proteingruppen auf der Basis des gesamten Datensatzes gefiltert (d. h. es werden Peptid- und Proteininformationen aus allen Läufen genommen, zunächst auf Peptidebene mit einer FDR von 1 % gefiltert und nur Peptide ohne verpasste Spaltungen, mit einer Mindestsequenzlänge von sieben Aminosäuren, einem PLGS-Score von mindestens 6.0 und keine variable Modifikation), dann auf der Grundlage des Occam's Razor Prinzip, was zu einer reduzierten Proteinliste führt (ebenfalls mit 1% FDR gefiltert), die alle Peptide erklären kann, die die oben genannten Kriterien im Datensatz erfüllen.

Blast2GO-Analyse

Die automatische funktionelle Annotation der durch LC-MS/MS identifizierten Proteine wurde mit dem Softwaretool Blast2GO v5.2.571 durchgeführt. Blast2GO verwendet den BLAST-Algorithmus zur Identifizierung ähnlicher Proteine und überträgt bereits vorhandene Gene Ontology (GO)-Annotationen auf die abgefragten Proteinsequenzen. Zusätzlich wird

InterProScan verwendet, um Proteinfamilien- und Domäneninformationen zu erhalten, die in GOs umgewandelt und zusammengeführt werden. Blast-Suchen wurden gegen die NCBI-Datenbank nicht-redundanter Proteinsequenzen (nr) mit dem Blastp-Algorithmus durchgeführt. Ansonsten wurde der Blast2GO Arbeitsablauf (Blast, InterProScan, Mapping und Annotation) mit Standardparametern durchgeführt.

Analyse phänotypischer Daten

Die Analyse der phänotypischen Daten wurde für jede Art getrennt nach dem linearen gemischten Modell in Gleichung (1) durchgeführt:

$$y_{ik} = u + v_i + env_k + e_{ik}, \quad (1)$$

wobei y_{ik} die phänotypische (= gemessene) Beobachtung für die i -te Sorte ist, die in der k -ten Umwelt getestet wurde, u der allgemeine Mittelwert, v_i der Effekt der i -ten Sorte, env_k der Effekt der k -ten Umwelt und e_{ik} der Restfehler ist. Die Varianzkomponenten, d. h. die Varianzen aufgrund der Sorten, der Umwelten und des Restfehlers, wurden mit der Methode der eingeschränkten maximalen Wahrscheinlichkeit (REML) unter Annahme eines Zufallsmodells in einer klassischen einstufigen Analyse geschätzt⁷². Ein Likelihood-Ratio-Test mit Modellvergleichen wurde durchgeführt⁷³, um die Signifikanz der Varianzkomponenten zu überprüfen. Die Durchschnittswerte der Proteine in den verschiedenen Umgebungen wurden als beste lineare unvoreingenommene Schätzungen (BLUEs) unter der Annahme fester genetischer (Sorten-) Effekte ermittelt. Die Erblichkeitsschätzungen (h^2) wurden nach Piepho und Möhring⁷⁴ gemäß Gleichung (2) berechnet:

$$h^2 = 1 - \vartheta / 2\sigma^2G, \quad (2)$$

wobei ϑ die mittlere Varianz einer Differenz von zwei besten linearen unverzerrten Prädiktoren und σ^2G die genotypische Varianz (Kultivarvarianz). Alle Analysen wurden mit Hilfe der Statistiksoftware R⁷⁵ und ASReml 3.076 durchgeführt.

Vergleich der Proteome der verschiedenen Weizenarten

Für den t-Test wurden die BLUEs der Proteine pro Sorte in jeder Art verwendet. Wir führten den Student's t-Test ($\alpha = 0,05$)⁷⁷ durch, um die Häufigkeit der Proteine zwischen den einzelnen Weizenarten zu vergleichen. Für den t-Test wurde die Annahme der Gleichheit der Varianzen zwischen den Gruppen durch Anwendung des Levene-Tests⁷⁸ überprüft. Wenn der Levene-Test signifikant war ($p < 0,05$, was bedeutet, dass die Varianzen zwischen den Gruppen nicht gleich waren), wurde anstelle des normalen t-Tests der robustere Welch's t-Test mit korrigierten Freiheitsgraden durchgeführt. Student's t-Test und Levene-Test wurden mit der Statistiksoftware R⁷⁵ durchgeführt. Für jedes Artenpaar wurden Vulkankurven erstellt, um die Proteine zu identifizieren, deren Häufigkeit statistisch signifikant unterschiedlich war (t-Test, $p < 0,05$) und über/unter einem willkürlichen Schwellenwert von $\pm 3 \log_2$ fold change (\log_2FC) lag. Das R-Paket EnhancedVolcano wurde zur Erstellung von Vulkangrafiken verwendet. Die \log_2FC wurde nach der Formel in Gleichung (3) berechnet:

$$\log_2FC_p = u_i - u_j, \quad (3)$$

wobei \log_2FC_p die Häufigkeit von Protein p in Spezies i im Verhältnis zu Spezies j ist, u_i und u_j die \log_2 mittlere Häufigkeit von Protein p in Spezies i bzw. Spezies j sind. Für das hierarchische Clustering wurden die Daten skaliert und die Euklidische Distanz berechnet. Das hierarchische Clustering wurde mit der Funktion "hclust" in der Statistiksoftware R⁷⁵ unter Anwendung der Methode von Ward⁷⁹ durchgeführt.

Identifizierung der allergenen Proteine

Wir verwendeten die Liste der samenbürtigen Weizenallergene²⁷, die auf der Grundlage der Datenbanken für allergene Proteinfamilien AllFam (www.meduniwien.ac.at/allfam) und AllergenOnline (www.allergenonline.org) kuratiert wurde, um potenziell allergene Proteine in Mehlproben von Weizenarten in der aktuellen Studie zu identifizieren. Wir erweiterten die Identifizierung allergener Proteine mit Hilfe der umfassenden Allergom-Datenbank (<http://www.allergome.org/index.php>)⁴⁶, die identifizierte, charakterisierte und von Experten begutachtete allergene Proteine enthält, darunter Proteine aus der offiziellen Allergen-Nomenklatur der Weltgesundheitsorganisation (WHO) und der International Union of Immunological Societies (IUIS) (<http://www.allergen.org/index.php>). Die UniProt-Zugänge wurden verwendet, um die Proteine über den Frontend-Browser der UniProt-Datenbank (<https://www.uniprot.org/>) der Allergome-Datenbank zuzuordnen, und die entsprechenden Allergome-IDs wurden abgerufen (siehe ergänzende Tabelle 4). Zusätzlich zur Identifizierung allergener Proteine mit Hilfe der oben genannten Datenbanken nutzten wir die Proteinannotationen der UniProt-Datenbank, um nach Amylase/Trypsin-Inhibitoren (ATIs) unter den in dieser Studie quantifizierten Proteinen zu suchen. Die Liste der ATIs findet sich in der ergänzenden Tabelle 4.

Zusammenfassung der Berichterstattung

Weitere Informationen zum Forschungsdesign finden Sie in der Nature Research Reporting Summary, die mit diesem Artikel verlinkt ist.

Verfügbarkeit der Daten

Die Massenspektrometrie-Proteomikdaten, einschließlich Rohdateien, Peptid- und Proteinquantifizierungsberichte, wurden beim ProteomeXchange Consortium über das PRIDE Partner-Repository⁸⁰ mit der Datensatzkennung PXD028676 hinterlegt

Quellen

- Weegels, P. L. The Future of Bread in View of its Contribution to Nutrient Intake as a Starchy Staple Food. *Plant Foods Hum. Nutr.* 74, 1–9 (2019).
- Beres, B. L. et al. A Systematic Review of Durum Wheat: Enhancing Production Systems by Exploring Genotype, Environment, and Management (G × E × M) Synergies. *Front. Plant Sci.* 11; <https://doi.org/10.3389/fpls.2020.568657> (2020).
- Lev-Yadun, S., Gopher, A. & Abbo, S. The cradle of agriculture. *Science* 288, 1602–1603 (2000).
- Zaharieva, M., Ayana, N. G., Hakimi, A. A., Misra, S. C. & Monneveux, P. Cultivated emmer wheat (*Triticum dicoccon* Schrank), an old crop with promising future: a review. *Genet. Resour. Crop Evol.* 57, 937–962 (2010).
- Zaharieva, M. & Monneveux, P. Cultivated einkorn wheat (*Triticum monococcum* L. subsp. *monococcum*): the long life of a founder crop of agriculture. *Genet. Resour. Crop Evol.* 61, 677–706 (2014).
- Hidalgo, A. & Brandolini, A. Nutritional properties of einkorn wheat (*Triticum monococcum* L.). *J. Sci. Food Agric.* 94, 601–612 (2014).
- Longin, C. F. H. & Würschum, T. Back to the Future – Tapping into Ancient Grains for Food Diversity. *Trends Plant Sci.* 21, 731–737 (2016).
- Miedaner, T. & Longin, C. F. H. Neglected cereals. From ancient grains to superfood (Erling, Clenze, 2017).
- Shewry, P. R. Wheat. *J. Exp. Bot.* 60, 1537–1553 (2009).
- Siddiqi, R. A., Singh, T. P., Rani, M., Sogi, D. S. & Bhat, M. A. Diversity in Grain, Flour, Amino Acid Composition, Protein Profiling, and Proportion of Total Flour Proteins of Different Wheat Cultivars of North India. *Front. Nutr.* 7, 141 (2020).
- Veraverbeke, W. S. & Delcour, J. A. Wheat Protein Composition and Properties of Wheat Glutenin in Relation to Breadmaking Functionality. *Crit. Rev. Food Sci. Nutr.* 42, 179–208 (2002).
- Shewry, P. R. & Halford, N. G. Cereal seed storage proteins: structures, properties and role in grain utilization. *J. Exp. Bot.* 53, 947–958 (2002).
- Shewry, P. R. What Is Gluten-Why Is It Special? *Front. Nutr.* 6, 101 (2019).
- Catassi, C. et al. The Overlapping Area of Non-Celiac Gluten Sensitivity (NCGS) and Wheat-Sensitive Irritable Bowel Syndrome (IBS): An Update. *Nutrients* 9, 1268 (2017).
- Schuppan, D., Junker, Y. & Barisani, D. Celiac Disease: From Pathogenesis to Novel Therapies. *Gastroenterology* 137, 1912–1933 (2009).
- Sollid, L. M. et al. Update 2020: nomenclature and listing of celiac disease-relevant gluten epitopes recognized by CD4 + T cells. *Immunogenetics* 72, 85–88 (2020).
- Ashfaq-Khan, M. et al. Dietary wheat amylase trypsin inhibitors promote features of murine non-alcoholic fatty liver disease. *Sci. Rep.* 9, 1–14 (2019).
- Bellinghausen, I. et al. Wheat amylase-trypsin inhibitors exacerbate intestinal and airway allergic immune responses in humanized mice. *J. Allergy Clin. Immunol.* 143, 201–212.e4 (2019).
- Junker, Y. et al. Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *J. Exp. Med.* 209, 2395–2408 (2012).
- Schuppan, D. & Zevallos, V. Wheat amylase trypsin inhibitors as nutritional activators of innate immunity. *Dig. Dis.* 33, 260–263 (2015).
- Zevallos, V. F. et al. Nutritional Wheat Amylase-Trypsin Inhibitors Promote Intestinal Inflammation via Activation of Myeloid Cells. *Gastroenterology* 152, 1100–1113.e12 (2017).
- Zevallos, V. F. et al. Dietary wheat amylase trypsin inhibitors exacerbate murine allergic airway inflammation. *Eur. J. Nutr.* 58, 1507–1514 (2019).
- Caminero, A. et al. Lactobacilli Degrade Wheat Amylase Trypsin Inhibitors to Reduce Intestinal Dysfunction Induced by Immunogenic Wheat Proteins. *Gastroenterology* 156, 2266–2280 (2019).
- Dahl, S. W., Rasmussen, S. K. & Hejgaard, J. Heterologous Expression of Three Plant Serpins with Distinct Inhibitory Specificities*. *J. Biol. Chem.* 271, 25083–25088 (1996).
- Fasano, A., Sapone, A., Zevallos, V. & Schuppan, D. Nonceliac gluten sensitivity. *Gastroenterology* 148, 1195–1204 (2015).
- Salcedo, G., Quirce, S. & Diaz-Perales, A. Wheat allergens associated with Baker’s asthma. *J. Investig. Allergol. Clin. Immunol.* 21, 81–92 (2011).
- Juhász, A. et al. Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci. Adv.* 4, eaar8602 (2018).
- Mameri, H. et al. Molecular and immunological characterization of wheat serpin (Tri a 33). *Mol. Nutr. Food Res.* 56, 1874–1883 (2012).
- van Winkle, R. C. & Chang, C. The Biochemical Basis and Clinical Evidence of Food Allergy Due to Lipid Transfer Proteins: A Comprehensive Review. *Clin. Rev. Allerg. Immunol.* 46, 211–224 (2014).
- Fritscher-Ravens, A. et al. Many Patients With Irritable Bowel Syndrome Have Atypical Food Allergies Not Associated With Immunoglobulin E. *Gastroenterology* 157, 109–118 (2019).
- Fritscher-Ravens, A. et al. Confocal Endomicroscopy Shows Food-Associated Changes in the Intestinal Mucosa of Patients With Irritable Bowel Syndrome. *Gastroenterology* 147, 1012–1020 (2014).
- Geisslitz, S., Longin, C. F. H., Scherf, K. A. & Koehler, P. Comparative study on gluten protein composition of ancient (einkorn, emmer and spelt) and modern wheat species (durum and common wheat). *Foods* 8, 409 (2019).
- Geisslitz,

S., Longin, C. F. H., Koehler, P. & Scherf, K. A. Comparative quantitative LC-MS/MS analysis of 13 amylase/trypsin inhibitors in ancient and modern Triticum species. *Sci. Rep.* 10, 14570 (2020). 34. El Hassouni, K. et al. Genetic architecture underlying the expression of eight α amylase trypsin inhibitors. *Theor. Appl. Genet.* 134, 3427–3441 (2021). 35. Sielaff, M. et al. Hybrid QconCAT-Based Targeted Absolute and Data-Independent Acquisition-Based Label-Free Quantification Enables In-Depth Proteomic Characterization of Wheat Amylase/Trypsin Inhibitor Extracts. *J. Proteome Res.* 20, 1544–1557 (2021). 36. Iacomino, G. et al. Triticum monococcum amylase trypsin inhibitors possess a reduced potential to elicit innate immune response in celiac patients compared to Triticum aestivum. *Food Res. Int.* 145, 110386 (2021). 37. Khodabocus, I., Li, Q., Mehta, D. & Uhrig, R. G. A Road Map for Undertaking Quantitative Proteomics in Plants: New Opportunities for Cereal Crops. In *Accelerated Breeding of Cereal Crops*, edited by A. Bilichak & J. D. Laurie (Springer US, New York, NY, 2022), pp. 269–292. 38. Afzal, M. et al. High-resolution proteomics reveals differences in the proteome of spelt and bread wheat flour representing targets for research on wheat sensitivities. *Sci. Rep.* 10, 14677 (2020). 39. Zimmermann, J. et al. Comprehensive proteome analysis of bread deciphering the allergenic potential of bread wheat, spelt and rye. *J. Proteom.* 247, 104318 (2021). 40. Afzal, M. et al. Characterization of 150 Wheat Cultivars by LC-MS-Based Label-Free Quantitative Proteomics Unravels Possibilities to Design Wheat Better for Baking Quality and Human Health. *Plants* 10, 424 (2021). 41. Felber, J. et al. Aktualisierte S2k-Leitlinie Zöliakie der Deutschen Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten (DGVS). *Z. für Gastroenterol.* 60, 790–856 (2022). 42. Sergi, C., Villanacci, V. & Carroccio, A. Non-celiac wheat sensitivity: rationality and irrationality of a gluten-free diet in individuals affected with non-celiac disease: a review. *BMC Gastroenterol.* 21, 1–12 (2021). 43. Pinto-Sanchez, M. I. & Verdu, E. F. Non-celiac gluten or wheat sensitivity: It's complicated! *Neurogastroenterol. Motil.: Off. J. Eur. Gastrointest. Motil. Soc.* 30, e13392 (2018). 44. Volta, U. et al. Nonceliac Wheat Sensitivity: An Immune-Mediated Condition with Systemic Manifestations. *Gastroenterol. Clin. North Am.* 48, 165–182 (2019). 45. Aufiero, V. R., Sapone, A. & Mazzarella, G. Diploid Wheats: Are They Less Immunogenic for Non-Celiac Wheat Sensitive Consumers? *Cells* 11, 2389 (2022). 46. Mari, A., Rasi, C., Palazzo, P. & Scala, E. Allergen databases: Current status and perspectives. *Curr. Allergy Asthma Rep.* 9, 376–383 (2009). 47. Longin, C. F. H. et al. Comparative study of hulled (einkorn, emmer, and spelt) and naked wheats (durum and bread wheat): agronomic performance and quality traits. *Crop Sci.* 56, 302–311 (2016). 48. Call, L. et al. Effects of species and breeding on wheat protein composition. *J. Cereal Sci.* 93, 102974 (2020). 49. Geisslitz, S. et al. Wheat ATIs: Characteristics and Role in Human Disease. *Front. Nutr.* 8, <https://doi.org/10.3389/fnut.2021.667370> (2021). 50. Pickert, G. et al. Wheat Consumption Aggravates Colitis in Mice via Amylase Trypsin Inhibitor-mediated Dysbiosis. *Gastroenterology* 159, <https://doi.org/10.1053/j.gastro.2020.03.064> (2020). 51. Sievers, S., Rohrbach, A. & Beyer, K. Wheat-induced food allergy in childhood: ancient grains seem no way out. *Eur. J. Nutr.* 59, 2693–2707 (2020). 52. Larré, C. et al. Assessment of allergenicity of diploid and hexaploid wheat genotypes: Identification of allergens in the albumin/globulin fraction. *J. Proteom.* 74, 1279–1289 (2011). 53. Lombardo, C. et al. Study on the Immunoreactivity of Triticum monococcum (Einkorn) Wheat in Patients with Wheat-Dependent Exercise-Induced Anaphylaxis for the Production of Hypoallergenic Foods. *J. Agric. Food Chem.* 63, 8299–8306 (2015). 54. Alvisi, P. et al. Responses of blood mononucleated cells and clinical outcome of non-celiac gluten sensitive pediatric patients to various cereal sources: a pilot study. *Int. J. Food Sci. Nutr.* 68, 1005–1012 (2017). 55. Ianiro, G. et al. A Durum Wheat Variety-Based Product Is Effective in Reducing Symptoms in Patients with Non-Celiac Gluten Sensitivity: A Double-Blind Randomized Cross-Over Trial. *Nutrients* 11, <https://doi.org/10.3390/nu11040712> (2019). 56. Bordoni, A., Danesi, F., Di Nunzio, M., Taccari, A. & Valli, V. Ancient wheat and health: a legend or the reality? A review on KAMUT khorasan wheat. *Int. J. Food Sci. Nutr.* 68, 278–286 (2017). 57. Shewry, P. R. Do ancient types of wheat have health benefits compared with modern bread wheat. *J. Cereal Sci.* 79, 469–476 (2018). 58. Carroccio, A. et al. Wheat Consumption Leads to Immune Activation and Symptom Worsening in Patients with Familial Mediterranean Fever: A Pilot Randomized Trial. *Nutrients* 12, <https://doi.org/10.3390/nu12041127> (2020). 59. Picascia, S. et al. In Celiac Disease Patients the In Vivo Challenge with the Diploid Triticum monococcum Elicits a Reduced Immune Response Compared to Hexaploid Wheat. *Mol. Nutr. Food Res.* 64, 1901032 (2020). 60. Ziegler, J. U. et al. Lutein and Lutein Esters in Whole Grain Flours Made from 75 Genotypes of 5 Triticum Species Grown at Multiple Sites. *J. Agric. Food Chem.* 63, 5061–5071 (2015). 61. Ziegler, J. U., Schweiggert, R. M., Würschum, T., Longin, C. F. H. & Carle, R. Lipophilic antioxidants in wheat (Triticum spp.): a target for breeding new varieties for future functional cereal products. *J. Funct. Foods* 20, 594–605 (2016). 62. Zeibig, F., Kilian, B. & Frei, M. The grain quality of wheat wild relatives in the evolutionary context. *Theor. Appl. Genet.*, 1–20, <https://doi.org/10.1007/s00122-021-04013-8> (2021). 63. Longin, C. F. H. et al. Mineral and Phytic Acid Content as Well as Phytase Activity in Flours and Breads Made from Different Wheat Species. *Int. J. Mol. Sci.* 24, 2770 (2023). 64. CIMMYT. 1990-91 CIMMYT World Wheat Facts and Trends: Wheat and Barley Production in Rainfed Marginal Environments of the Developing World (CIMMYT, Mexico, D.F., 1991). 65. Bencze, S. et al. Re-Introduction of Ancient Wheat Cultivars into Organic Agriculture—Emmer and Einkorn Cultivation Experiences under Marginal Conditions. *Sustainability* 12, 1584 (2020). 66. Manza, L. L., Stamer, S. L., Ham, A.-J. L., Codreanu, S. G. & Liebler, D. C. Sample preparation and digestion for proteomic analyses using spin filters. *Proteomics* 5, 1742–1745 (2005). 67. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* 6, 359–362 (2009). 68. Distler, U. et al. Drift time-specific collision energies enable deep-coverage data independent acquisition proteomics. *Nat. Methods* 11, 167–170 (2013). 69. Distler, U., Łacki, M.

K., Schumann, S., Wanninger, M. & Tenzer, S. Enhancing Sensitivity of Microflow-Based Bottom-Up Proteomics through Postcolumn Solvent Addition. *Anal. Chem.* 91, 7510–7515 (2019). 70. Silva, J. C., Gorenstein, M. V., Li, G.-Z., Vissers, J. P. & Geromanos, S. J. Absolute Quantification of Proteins by LCMSE: A Virtue of Parallel ms Acquisition *S. *Mol. Cell. Proteom.* 5, 144–156 (2006). 71. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676 (2005). 72. Cochran, W. G. & Cox, G. *Experimental Designs*. 2nd ed. (Wiley, New York, 1957). 73. Stram, D. O. & Lee, J. W. Variance components testing in the longitudinal mixed effects model. *Biometrics* 50, 1171–1177 (1994). 74. Piepho, H.-P. & Möhring, J. Computing heritability and selection response from unbalanced plant breeding trials. *Genetics* 177, 1881–1888 (2007). 75. R Core Team. *R: a language and environment for statistical computing* (R Foundation for Statistical Computing, Vienna, Austria, 2018). 76. Gilmour, A. R., Gogel, B., Cullis, B. R. & Thompson, R. *ASReml User Guide Release 3.0* (VSN International Ltd, Hemel Hempstead, UK, 2009). 77. Student. The probable error of a mean. *Biometrika* 6, 1–25 (1908). 78. Levene, H. Robust tests for the equality of variance. In *Contributions to Probability and Statistics*, edited by I. Olkin (Stanford University Press 1960), pp. 278–292. 79. Ward, J. H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244 (1963). 80. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47, D442–D450 (2019)