GE Free New Zealand

In Food And Environment Inc.

PO Box 13402, Wellington, NZ

Soil & Health Association

proud publishers of OrganicNZ magazine

7 July 2017

Food Standards Australia New Zealand
PO Box 10559
Wellington

Joint submission on Application A1139

Introduction

GE Free New Zealand in Food and Environment and the Soil & Health Association of New Zealand ("we" and "the submitters") welcome the opportunity to comment on the application A1139 Food derived from Potato Lines F10, J3, W8, X17 & Y8 ("Application").

GE Free New Zealand in Food and Environment ("GE Free NZ") is an Incorporated Society. It is a non-Governmental Organisation governed by a Board and has a nationwide membership base. It represents its members when making submissions and helps with gathering and disseminating information concerning genetically modified organisms ("GMO") to its members and the wider public through regular newsletters and its website (www.gefree.org.nz).

The Soil & Health Association of New Zealand Inc. ("Soil & Health") is a charitable society registered under the Incorporated Societies Act 1908. It is the largest membership organization supporting organic food and farming in New Zealand and is one of the oldest organic organisations in the world, established in 1941. Soil & Health's objectives are to promote sustainable organic agricultural practices and the principles of good health based on sound nutrition and the maxim: "Healthy soil, healthy food, healthy people". Its membership is chiefly composed of home gardeners and consumers, organic farmers and growers, secondary producers, retailers and restaurateurs. Soil & Health publishes the bi-monthly 'Organic NZ' magazine – New Zealand's leading organics magazine.

We recommend that FSANZ decline the Application. We submit FSANZ cannot approve the potato lines in the Application without a serious breach of its duty of care as well as the principles of its own mission statement.

We note that there are insufficient data on both the sprays and novel proteins detailed in the Application.

We note that FSANZ's legal requirements as stated in its mission statement are:

> *To protect, in collaboration with others, the health and safety of people in Australia and New Zealand through the maintenance of a safe food supply.*

FSANZ Values are:

- To be impartial, open and accountable;
- To use the best available sciences and evidence to guide decision-making; and
- To seek, respect and be responsive to the issues raised by others.

FSANZ Responsibilities are:

- Provide information to consumers to enable better consumer choice;
- Undertake dietary exposure modeling and scientific risk assessments; and
- Provide risk assessment advice on imported food.

We have read the assessments for this Application and consider that FSANZ have led stake holders and consumers astray. We outline our concerns below.

Detailed submission

It is illegal to import viable GE plants or plant parts into New Zealand. Potato plants can readily regenerate from even small parts of raw tubers, therefore making any raw imported GE potatoes equivalent to live GE plant material. This will endanger the biosecurity status of New Zealand. It would be illegal to approve the entry of these GE potatoes/potato pieces into the country.

Not labeling GE-containing foods at the point of sale is a breach of consumer rights. The lack of labeling of GE potatoes sold in any form by restaurants is deceptive, as consumers will be unaware of this. FSANZ should support rather than oppose the enforcement and monitoring of compliance around GE food labeling.

No independent food safety experiments have been carried out on these GE potato lines. Instead FSANZ has relied on data from within the GE industry, i.e., the applicant data. This shows that FSANZ has not been impartial, open or accountable to the public. FSANZ has not required that any independent experimental food safety assessments be undertaken on these potatoes. It can therefore not provide advice on the safety of these imported potato lines, such advice being its core responsibility.

The executive summary of the FSANZ evaluation reads:

> *The changes to levels of free amino acids and reducing sugars are not nutritionally consequential as they do not affect the levels of essential amino acids or other key nutrients important to potato.*

This statement assumes that any changes in the amino acids, free or otherwise, do not affect the levels of all other compounds present in the GE potatoes. The amino acid glutamine, for example, plays an important role in maintaining a healthy immune system, digestive tract and muscle cells. Any changes to amino acid balance may cause alterations to the assimilation of other amino acids. Studies have shown glutamine to reduce morbidity and mortality in periods of critical illness.[1] This demonstrates that any changes to endogenous amino acid levels should not be ignored.

Statement on Compositional analyses:

[1] Lacey, J., & Wilmore, D. (2009). Is Glutamine a Conditionally Essential Amino Acid?. *Nutrition Reviews, 48(8), 297-309.*

A detailed compositional analysis was performed on W8, X17, Y9, F10 and J3 to establish the nutritional adequacy of tubers produced from these lines and to characterise any unintended compositional change. Analyses were done of proximates, fibre, vitamins, minerals, total amino acids, free amino acids, sucrose, reducing sugars (fructose and glucose), and anti-nutrients (glycoalkaloids). These showed that, even with the intended changes to sucrose, reducing sugars and asparagine, the levels of all analytes fell within the natural variation found across the range of conventional potato lines used for human consumption. No conclusion could be reached in relation to line E56 as no compositional data was provided. (iii)

The changes in concentrations of glutamine and asparagine in the GE potato lines may be of some concern. Altered levels of asparagine can result in complications in fetal development, causing brain and neurological problems.[2]

There are concerns over the meaning of "biological relevant differences" as stated in the summary document:

Analysis of the events W8, X17, and Y9 have not revealed any biologically relevant differences compared to the conventional varieties, except for the intended late blight protection, low free asparagine, low reducing sugars, and low polyphenol oxidase activity.

Assuming "biologically relevant differences" translates as food safety, there are no feeding studies to back up this statement, so such assumptions are unable to be made.

Published research on GE potatoes has shown unexpected harmful effects on animals fed with these crops. A 1999 study (Ewen and Puzstai) conducted on rats fed with transgenic potatoes found that abnormalities occurred in the gastrointestinal tract (small intestine and caecum) within a short time.[3] This study found that the GE potatoes caused gut abnormalities with or without (an 'empty construct') the lectin gene. Lectin is a harmless insecticidal compound produced by a number of plants. The authors concluded that:

"(b)ecause caecal thickness was similar in rats given boiled parent potatoes in the presence or absence of spiked GNA (a harmless lectin from the plant species *Galanthus nivalis*)*, we*

[2] Ruzzo,E., Capo-Chichi, J., Ben-Zeev., Chitayat, D., Mao, H., & Pappas,A. et al. (2013). Deficiency Asparagine Synthetase Causes Congenital Microcephaly and a Progressive Form of Encephalopathy. Neuron, 80(2), 429-441.

[3] Ewen, S., & Pusztai, A. (1999). Effect of diets containing genetically modified potatoes expressing Galanthus nivalis lectin on rat small intestine. *The Lancet*, *354*(9187), 1353-1354.

*suggest that the decrease in caecal mucosal thickness seen in rats **fed boiled GM-potato diets was the consequence of the transfer of the GNA gene into the potato."***

These GE potatoes were not subsequently commercialised.

Similar results were obtained in feeding experiments using GE potatoes by Fares *et al.*[4] , who found that there were changes to the mucosal lining and other cells of the ileum of mice. They called for comprehensive feeding tests to avoid any potential risks:

*"Although transgenic crop plants used in food and feed production carry different beneficial transgenes… **before releasing for marketing thorough tests and all possible consequences of these new types of heredity and new genetic structures must be evaluated to avoid any potential risks"***

A 2007 study showed that the consumption of GE potatoes has been observed to cause an increase in immunoglobulin (Ig) levels in human participants.[5]

This is cause for concern, as antibody levels may well have increased as a result of novel proteins present in the GE food.

These afore-mentioned studies are but three of many published studies on the harmful effects of GE foods. Please refer to the submission of the Physicians and Scientists for Global Responsibility (**PSGR**) for a more comprehensive list of publications on the harmful effects of GE crops, as observed in feeding experiments.

The obligation of FSANZ is to make themselves aware of such studies and treat all GE foods as potentially harmful. It would be completely irresponsible to allow these potatoes onto the market. Feeding studies need to be conducted on the GE potato lines in this application before the potatoes are released. This will determine whether these lines have negative health impacts, or even life threatening responses, such as an allergic reaction.

Changes in Metabolites

Cellini *et al*. (2004) reported widespread changes to metabolite levels, both expected and unexpected in GE potato lines. They recommended that data analysis tools need to be used.[6]

[4] Fares, N., & El-Sayed, A. (1998). Fine Structural Changes in the Ileum of Mice Fed on δ-Endotoxin-Treated Potatoes and Transgenic Potatoes. *Natural Toxins*, *6*(6), 219-233.
[5] Tacket, C. O. (2007). Plant-Based Vaccines Against Diarrheal Diseases. *Transactions of the American Clinical and Climatological Association*, *118*, 79–87

A study of potato metabolite production has found that field-grown vs laboratory-grown potato tubers showed a tenfold and greater differences across a range of compounds.[7] The potatoes with modified sucrose metabolism or inhibited starch synthesis revealed unexpected disaccharides (trehalose, maltose and isomaltose).[8] Such changes in metabolites cannot be overlooked, when assessing this application.

A particular cause for concern is that FSANZ has deemed the six GE potato lines in this application as "safe", when there is no compositional data on one line E56.

Applicant data to APHIS[9]

We have outlined the comments from the data provided to APHIS about the potatoes.

7.3 Soft root testing with tubers

> Of the events in that trial (E12, E24, F10, J3, J55, and J78), the only significant difference was that event F10 was more resistant to this disease than the control. (APHIS p.46)

Late blight foliage testing

> Considering both studies, we conclude that the events have similar susceptibility to bacterial soft rot as the controls. (APHIS, p.46)

7.4 Reducing Sugars.

Tubers of the events G11, H37, and H50 contain the same amount of reducing sugars as tubers of their untransformed (non-GE) counterparts. The inability of the silencing construct to limit glucose/fructose formation in H37 and H50 may be due to the fact that the H variety is naturally low in glucose and fructose. Thus, we concluded that silencing of the promoters associated with the PhL/R1 genes effectively lowered reducing sugars near the time of harvest in most events but these differences were not sustained throughout storage for 2-5 months". (APHIS,p. 47)

Disease susceptibility – Appendix 8

---

[6] Cellini, F., Chesson, A., Colquhoun, I., Constable, A., Davies, H., & Engel, K. et al. (2004). Unintended effects and their detection in genetically modified crops. *Food And Chemical Toxicology*, *42*(7), 1089-1125.

[7] Roessner, U., Wagner, C., Kopka, J., Tretheway, N., Willmitzer, L., 2000. Simultaneous analysis of metabolites in potato tubers by gas chromatography-mass spectrometry. Plant Journal 23, 131–142.

[8] Acrylamide Potential and Reduced Black Spot Bruise: Events E12 and E24 (Russet Burbank); F10 and F37 (Ranger Russet); J3, J55, and J78 (Atlantic); G11 (G); H37and H50 (H) – 2013 https://www.aphis.usda.gov/brs/aphisdocs/13_02201p.pdf

[9] www.isaa.org/kc/cropbiotechupdate/article/default.asp?ID=7422

Thus, independent lines of two chipping varieties and two French fry varieties with low Ppo expression in tubers were shown to have similar susceptibility to bacterial soft rot to the corresponding untransformed control for each variety.

Considerations

The APHIS document on the GE potatoes details some significant differences in the compositions between the GE lines and non-GE controls. These GE potatoes are of no nutritional benefit to consumers and could contain higher levels of anti-nutrients.

FSANZ has overlooked three fundamental issues, when allowing this application to proceed:

1. There are currently non-GE potato varieties available that are ideal for chipping and processing .

In **section 2.4.3** of the application it states that the applicant has indicated that reduced blackspot bruising of these GE potato lines can reduce wastage during storage and processing, and that the potatoes are resistant to the fungal disease known as foliar late blight. There are already several non-GE varieties of blight-resistant potatoes (including 'Waneta' and 'Lamoka'), which have been released by plant breeders from the University of Cornell (US). These varieties are ideal for chips, because they store very well and produce a good colour when cut [9]. The Cornell breeding programme develops chipping and tabletop varieties, focussing on colour, size, shape, texture and disease- and pest-resistance.

2. New Zealand has a range of excellent climates and soils in which to grow these non-GE varieties of potatoes. This would support NZ growers and potato processing plants. Furthermore, the importation of potato products from the other side of the world, is an unnecessary source of carbon emissions that will contribute to what is already a major world problem.

3. Acrylamide production can be reduced by the use of sensible cooking methods. There is much information available on this topic.

Conclusions

We ask that FSANZ **decline** approval of A1139.

- An adequate risk assessment and evaluation of the effect/s of novel genes/proteins and subsequent changes in the A1139 potato lines has not been carried out.

- No independent feeding test risk assessments have been undertaken or evaluated by FSANZ.

- The Applicant information provided on safety is insufficient and lacking up to date metabolic profiling using proteomic testing for entry into the food chain.

- The lack of information does not allow the consumer to make informed decisions and removes consumer choice

- By not allowing for labeling of A1139, FSANZ has not provided information to consumers that will enable better consumer choice.

- The assessment has no information about any novel protein/s, which may have been produced during the GE process.

- There is a lack of scientific data necessary to protect and maintain a safe food supply for the health and safety of people in Australia and New Zealand.

  The best available science has not been used to properly guide decision-making.

- The reliance on applicant's data has not shown impartiality, openness and accountability.

Sincerely

Jon Muller
Secretary GE Free NZ

Mischa Davis
Soil and Health

CONCISE COMMUNICATION

Human Immune Responses to a Novel Norwalk Virus Vaccine Delivered in Transgenic Potatoes

Carol O. Tacket,[1] Hugh S. Mason,[3] Genevieve Losonsky,[1] Mary K. Estes,[2] Myron M. Levine,[1] and Charles J. Arntzen[3]

[1]Center for Vaccine Development, University of Maryland School of Medicine, Baltimore; [2]Baylor College of Medicine, Division of Molecular Virology, Houston, Texas; [3]Boyce Thompson Institute for Plant Research, Ithaca, New York

A new approach for delivering vaccine antigens is the use of inexpensive, plentiful, plant-based oral vaccines. Norwalk virus capsid protein (NVCP), assembled into virus-like particles, was used as a test antigen, to determine whether immune responses could be generated in volunteers who ingested transgenic potatoes. Twenty-four healthy adult volunteers received 2 or 3 doses of transgenic potato ($n = 20$) or 3 doses of wild-type potato ($n = 4$). Each dose consisted of 150 g of raw, peeled, diced potato that contained 215–751 μg of NVCP. Nineteen (95%) of 20 volunteers who ingested transgenic potatoes developed significant increases in the numbers of specific IgA antibody–secreting cells. Four (20%) of 20 volunteers developed specific serum IgG, and 6 (30%) of 20 volunteers developed specific stool IgA. Overall, 19 of 20 volunteers developed an immune response of some kind, although the level of serum antibody increases was modest.

One new strategy for mucosal vaccine development is the use of oral vaccines delivered in transgenic plants [1–4]. Such a plant delivery system would be inexpensive and plentiful, and the delivery of vaccine antigen in plant cells may protect the antigen as it passes through the acidic environment of the stomach. In a recent proof-of-principle study, volunteers who ate a transgenic potato that expressed the nontoxic B subunit of enterotoxigenic *Escherichia coli* enterotoxin developed serum and/or mucosal immune responses [5]. The B subunit of *E. coli* enterotoxin is a potent stimulator of mucosal immune responses, in part because of its ganglioside-binding activity, which allows targeting to epithelial cells. It is not known whether the principle of vaccine delivery by an edible plant can be generalized to other antigens, such as viruses. In this study, the Norwalk virus (NV) capsid protein (NVCP) expressed in the potato was used as a test antigen.

NV, a member of the family *Caliciviridae,* causes acute gastroenteritis and has been a recent target for vaccine development. NV is the major cause of nonbacterial epidemic gastroenteritis and is spread via food and water and from person to person. Recent studies using new diagnostic assays developed with recombinant NV (rNV) particles or using reverse transcription–polymerase chain reaction have shown that the epidemiologic significance of NV infections has been greatly underestimated [6–9].

The successful cloning, sequencing, and expression of the single capsid protein of NV in insect cells, by using baculovirus recombinants, led to the discovery that the capsid protein assembles spontaneously into virus-like particles (VLPs) that lack nucleic acid [10]. Such rNV particles are produced in high concentrations in vitro (20 mg of purified particles per 2×10^8 infected cells), and these particles have desirable properties for use as a subunit vaccine. Purified Norwalk VLPs are safe and immunogenic when given to volunteers [11].

The capsid protein of NV can be expressed in plants and causes an immune response when given orally to mice [12]. In our study, we fed transgenic potato expressing rNV capsid protein to adult volunteers, to determine whether similar immune responses could be generated in humans.

Methods

Construction of NV-transgenic potato. Transgenic potato line NV140-13 was created by the transformation of potato (*Solanum tuberosum* L. cv. "Frito Lay" 1607) with pNV140, a binary vector that contains expression cassettes for NVCP and kanamycin resistance [12]. Potato line NV140-13 contains 4 copies of pNV140 T-DNA stably integrated into nuclear chromosomal DNA, as determined by genomic Southern blot (data not shown). The control

Received 7 February 2000; revised 30 March 2000; electronically published 6 July 2000.

Informed consent was obtained from volunteers, and the human experimentation guidelines of the US Department of Health and Human Services and the University of Maryland, Baltimore, were followed in the conduct of this clinical research.

None of the authors has an intellectual property agreement or other association that might pose a conflict of interest with any commercial company related to plant-based edible vaccines.

Financial support: National Institutes of Health (AI-65299).

Reprints or correspondence: Dr. Carol O. Tacket, Center for Vaccine Development, 685 West Baltimore St., Baltimore, MD 21201 (ctacket@medicine.umaryland.edu).

The Journal of Infectious Diseases 2000; 182:302–5

0022-1899/2000/18201-0039$02.00

wild-type potato was *S. tuberosum* L. cv. "Frito Lay" 1607 without any genetic manipulation.

Tissue culture plantlets of transgenic potato line NV140-13 were clonally propagated and transplanted into soil in 11.35-liter plastic pots. Potted plants were maintained under greenhouse conditions, fertilized, and grown for at least 15 weeks. Tubers were harvested, and tubers from each pot were kept separate from each other. Each plant produced 720–1224 g of tubers. One tuber per pot was evaluated by ELISA for NVCP as described [12], with the following modifications. The tuber tissue was ground to a fine powder with a mortar and pestle under liquid N_2. Each gram of powder was thawed and extracted with 4 mL of buffer that contained 50 m*M* sodium phosphate, pH 6.6, 110 m*M* NaCl, 50 m*M* sodium ascorbate, 1 m*M* EDTA, 1 m*M* phenylmethylsulfonyl fluoride, and 0.1% Triton X-100. Aliquots were transferred to microcentrifuge tubes and were centrifuged at 16,000 *g* at 4°C for 3 min, and the supernatants were diluted for ELISA as described [12]. Data were recorded as micrograms of NVCP per gram of peeled potato tuber.

Assay for VLP assembly. NVCP assembled as VLP in potato tuber was assayed by sucrose gradient sedimentation. Fresh, unfrozen tubers were peeled and homogenized using a Ten-Broek (Wheaton Industries, Millville, NJ) ground glass homogenizer (clearance 0.15 mm) in 2 mL per gram ice-cold buffer (containing 25 m*M* sodium phosphate, pH 6.6, 50 m*M* NaCl, 50 m*M* sodium ascorbate, 1 m*M* EDTA, and 1 m*M* phenylmethylsulfonyl fluoride). The homogenate was clarified by centrifugation at 16,000 *g* at 4°C for 3 min, and 0.5 mL was layered on a discontinuous sucrose gradient constructed by layering 2.1 mL each of 50%, 40%, 30% 20%, and 10% sucrose/PBS into a Beckman SW41 Ti tube (Beckman Coulter, Fullerton, CA). Gradients were centrifuged in a rotor (SW41 Ti; Beckman Coulter) at 41,000 rpm at 4°C for 3 h and were fractionated by using a density gradient fractionator (model 640; ISCO, Lincoln, NE). Fractions were assayed for NVCP by ELISA as described [12]. VLP standard produced in insect cells (provided by M. Estes) was sedimented in a similar gradient and was analyzed by ELISA to identify the VLP fraction. The proportion of the total ELISA-positive NVCP material in the potato tuber gradient that co-sedimented with the VLP standard was calculated to obtain the percentage of VLPs.

Clinical study design. Twenty-four healthy adult volunteers were randomized in a double-blind manner to receive 1 of 3 different regimens: (1) 3 doses of transgenic potato on days 0, 7, and 21 ($n = 10$); (2) 2 doses of transgenic potato on days 0 and 21 and 1 dose of wild-type potato on day 7 ($n = 10$); or (3) 3 doses of wild-type potato on days 0, 7, and 21 ($n = 4$).

Preparation of potato vaccine. The potatoes were peeled before being ingested, to remove the skin, which contains solanine, an alkaloid that is present in all raw potatoes and can cause abdominal discomfort or nausea and may have a bitter taste. On the morning of dosing, transgenic and wild-type potatoes were peeled and cut into uniform 1-cm cubes. Transgenic and wild-type potato cubes were pooled into separate bowls, to minimize differences in dose due to variable NVCP expression in transgenic tubers. The cubes were mixed and stored in separate bowls of cold water (to prevent their turning brown) until ready for dosing. Potatoes were drained on a paper towel. Doses of 150 g of raw potato cubes were weighed on a scale and immediately ingested. Volunteers ingested nothing for 90 min before and after ingesting potato. A representative aliquot of potato cubes from each bowl was stored at −70°C and assayed for NVCP by ELISA to document the dose of rNV ingested.

Clinical follow-up and immunologic response measurements. Volunteers kept a diary for 3 days after ingesting each dose of potato. Any symptoms, such as nausea, vomiting, cramps, diarrhea, and fever, were recorded.

Blood was obtained for serum antibody responses to recombinant NVCP on day 0 and on days 7, 14, 21, 28, and 61 after the first dose of potato. The IgG and IgM antibody responses to NVCP were measured by ELISA in a blinded fashion by using coded specimens [11]. Whole blood was collected for IgA antibody–secreting cell (ASC) assays on days 0, 7, 14, 16, 21, 28, and 30 after the first dose of potato. ASCs that produced antibody against NVCP were measured by ELISPOT assays in a blinded fashion [5]. In this assay, microtiter plates were coated with NVCP at a concentration of 1 μg/mL and held overnight at 4°C. Peripheral blood mononuclear cells (PBMC) were added so that each of 4 wells contained 250,000 cells per well, and the assay was done as described elsewhere [5].

Whole stool was collected on days 0, 7, 14, 21, 28, and 61. Coded stool samples were processed to measure total fecal IgA and specific IgA antibodies against NVCP. In brief, stool was suspended in a 10% solution (wt/vol) of supplemented PBS and centrifuged, and the supernatant was assayed for total IgA antibody. The calculation of specific antibody against NVCP was standardized to 20 mg% of total IgA to correct for variability in the amount of total IgA in stool and to allow for comparisons among individuals.

Definitions. Fever was defined as an oral temperature >37.8°C. Diarrhea was defined as the passage of 3 liquid stools within 24 h. A response for each immunologic end point was defined as a 4-fold rise in stool or serum antibodies after immunization. A significant rise in ASCs after immunization was defined as the mean of the number of ASCs in the pre-immunization specimens +3 SD.

Statistical analysis. Two-tailed Fisher's exact test was used to compare the rates of occurrence of symptoms.

Results

NVCP content of transgenic potatoes and assembly into VLPs. At the end of each of the 3 vaccination days, well-mixed aliquots of peeled, diced transgenic potatoes were shipped on dry ice to Boyce Thompson Institute and assayed for NVCP content. From the first vaccination day, 2 aliquots of potatoes from 2 batches of diced, mixed potatoes contained ~215 and 542 μg of NVCP per 150-g dose, respectively; from the second vaccination day, 1 aliquot from the single batch used that day contained ~378 μg of NVCP per 150-g dose; from the third vaccination day, 2 aliquots of potatoes from 2 batches of diced, mixed potatoes contained ~538 and 751 μg of NVCP per 150-g dose, respectively. No NVCP was detected in any aliquot of peeled, diced wild-type potatoes.

Assay for VLP assembly. The proportion of NVCP that was assembled as VLPs was estimated by sucrose gradient sedimentation, by using a VLP standard produced in insect cells. We determined that the standard extraction procedure used for

ELISA, in which frozen tubers were powdered and extracted in buffer that contained Triton X-100, caused substantial dissociation of VLPs to yield mostly free subunits. We thus extracted fresh tuber material in buffer that lacked Triton X-100 for VLP estimation. By calculating the proportion of the total ELISA-positive material that co-sedimented with the VLP standard, we found that the percentage of VLPs in different tubers of potato line NV140-13 varied between 25% and 50%.

Safety. There were no differences in the incidence rates of nausea, vomiting, mild cramps, fever, or diarrhea among volunteers who ingested transgenic or wild-type potatoes in the 3 days after ingesting the first dose of potatoes. The most common symptoms were nausea, which occurred in 4 (20%) of 20 of those who ingested transgenic potatoes and 1 (25%) of 4 who ingested wild-type potatoes, and cramps, which occurred in 5 (25%) of 20 who ingested transgenic potatoes and 2 (50%) of 4 who ingested wild-type potatoes.

Immunogenicity. Nineteen (95%) of 20 volunteers who ingested 2 or 3 doses of transgenic potatoes developed significant increases in the numbers of IgA ASCs (range, 6–280/10^6 PBMC) (table 1). The recipients of wild-type potato had a geometric mean of 2 IgA ASCs/10^6 PBMC after 3 doses. Thirteen of the 19 IgA ASC responses occurred after the first dose of transgenic potato. Six (30%) of 20 volunteers developed significant increases in IgG ASCs (range, 25–115/10^6 PBMC). These circulating cells that secrete specific antibodies reflect immunologic priming of the gut mucosal immune system.

The prevaccination serum IgG antibody titers varied from <1:50 to 1:6400 among these adult volunteers; this variance is expected in an unselected population of adults. The mean prevaccination titer did not differ between vaccinees and control volunteers ($P = .28$, 2-tailed t test). Four (20%) of 20 volunteers developed serum IgG anti-NVCP, and 4 (20%) of 20 volunteers (3 of whom did not develop IgG responses) developed serum IgM anti-NVCP after ingesting transgenic potatoes (table 1). The geometric mean titer of serum IgG anti-NVCP was 1:67 before immunization and 1:757 after immunization among responders. This corresponds to a mean fold rise of 12 in IgG titer among responders. The mean titer of serum IgM was <1:15 before immunization and 1:100 after immunization among responders. This represents a mean fold rise of 7 in IgM titer among responders. Stool IgA anti-NVCP was detected in 6 (30%) of 20 volunteers who ingested transgenic potatoes (geometric mean titer of 1:45 or a mean fold rise of 17 in titer among responders) (table 1).

The serum responses occurred after the second dose in 6 of the 8 IgG and IgM seroconversions, and the stool antibody responses occurred after the second dose in 4 of the 6 responders. The rise in IgG titers persisted at day 61, the latest time point measured, but the IgM titers had decreased to baseline at day 28. Immune responses were not clearly related to the amount of NVCP ingested in each dose, which varied from 215 to 751 μg per dose. Recipients of wild-type potatoes had no immune responses to NVCP.

Discussion

In an earlier study, 100% of 15 volunteers who ingested 2 doses of purified Norwalk VLPs (250 μg/dose) with bicarbonate buffer developed increases in serum IgG antibody [11]. In the current study, 7 (35%) of 20 volunteers developed serum IgG or IgM responses after a similar dose of Norwalk protein delivered without buffer in potatoes. Only a portion of the NVCP was assembled into VLPs in the potato cell. Assembly into a particle may be crucial to obtain maximal mucosal immunogenicity of this protein. Immunogenicity probably is enhanced by the uptake and processing of the virus particle, and little or no response would be expected to the soluble NVCP in the

Table 1. Immune responses to transgenic potatoes that express Norwalk virus capsid protein (NVCP) and to wild-type potatoes.

	Transgenic potatoes			Wild-type potatoes,
Immunoassay	3 doses ($n = 10$)	2 doses ($n = 10$)	Total ($n = 20$)	3 doses ($n = 4$)
IgA ASC anti-NVCP response rate	9/10 (90%)	10/10 (100%)	19/20 (95%)	0/4
Geometric mean peak ASCs per 10^6 PBMC[a]	32	26	28	—
Range IgA ASCs per 10^6 PBMC[a]	6–245	6–280	6–280	—
IgG ASC anti-NVCP response rate	2/10 (20%)	4/10 (40%)	6/20 (30%)	0/4
Geometric mean peak ASCs per 10^6 PBMC[a]	103	34	49	0
Range IgG ASCs per 10^6 PBMC[a]	92–115	25–62	25–115	0
Serum IgG anti-NVCP response rate	3/10 (30%)	1/10 (10%)	4/20 (20%)	0/4
IgG peak geometric mean titer[a]	1:468	1:3200	1:757	—
Mean peak fold rise[a]	13.3	8	12	—
Serum IgM anti-NVCP response rate	4/10 (40%)	0/10 (0%)	4/20 (20%)	0/4
IgM peak geometric mean titer[a]	1:100	—	1:100	—
Mean peak fold rise[a]	7	—	7	—
Stool IgA response rate	4/10 (40%)	2/10 (20%)	6/20 (30%)	0/4
Stool IgA peak geometric mean titer[a]	1:48	1:38	1:45	—
Mean peak fold rise[a]	17.8	16.6	17.4	—

NOTE. ASC, antibody-secreting cell; PBMC, peripheral blood mononuclear cells.
[a] Among responders.

potato cell cytoplasm or even to soluble protein released in the intestine. In addition, the lack of a linear dose-response between the amount of NVCP ingested and the immune response may be related to differences in the degree to which NVCP was assembled into VLPs in the different batches of potatoes ingested by volunteers. The variability in responses among volunteers also could be related to differences in genetics or history of NV infection in these individuals.

The type of immune response required to induce protection against Norwalk-like virus infection and the role of specific antibody are not known [13, 14]. This study has demonstrated that foreign proteins, other than the B subunit of *E. coli* heat-labile toxin, can be immunogenic when presented to the mucosal immune system via an orally delivered transgenic potato. Although 19 of our 20 volunteers developed an immune response of some kind, the rate and titers of serum antibody increases were modest, possibly because of the high prevalence of pre-existing serum antibody to NVCP.

Immune responses to wild-type NV challenge provide a benchmark for potential immune responses to NVCP after experimental infection. In a study in which volunteers were challenged with wild-type NV, the geometric mean IgG anti-NVCP after challenge was 1:62,414, with a mean prechallenge titer similar to that in our volunteers [11]. The immunogenicity of NVCP presented in potatoes may be improved if larger doses of recombinant protein are given and if a larger proportion of the NVCP is assembled into VLPs. Strong plant promoters and tissue-specific promoters should be chosen to efficiently drive the transcription of the vaccine protein gene. Immune responses also could be improved by using a mucosal adjuvant. Although the NVCP assembled into VLPs is immunogenic without adjuvant, mucosal adjuvants have been shown to enhance the immune responses of other particles given by the oral or intranasal route in preclinical studies [15]. Nontoxic derivatives of heat-labile enterotoxin of *E. coli* (LT) or cholera toxin (CT) are potent mucosal adjuvants, and simultaneous administration of vaccine protein and a mutant LT or CT (both delivered in a transgenic food) might greatly enhance the immune responses to the vaccine protein.

Plant-based vaccines have tremendous potential as a means to produce or safely deliver vaccine antigens. A number of plant-base vaccines against human diseases, such as rabies and hepatitis B [1, 2], have been tested in animals, and clinical studies will continue to evaluate this creative approach to mucosal immunization.

Acknowledgments

We are grateful for the invaluable contributions of the staffs of the Adult Clinical Studies and Applied Immunology Sections, Center for Vaccine Development, and of the volunteers from our community. We also thank Bryan Maloney at the Boyd Thompson Institute for expert technical assistance.

References

1. Mason HS, Lam DMK, Arntzen CJ. Expression of hepatitis B surface antigen in transgenic plants. Proc Natl Acad Sci USA **1992**; 89:11745–9.
2. Modelska A, Dietzschold B, Sleysh N, et al. Immunization against rabies with plant-derived antigen. Proc Natl Acad Sci USA **1998**; 95:2481–5.
3. Kusnadi AR, Hood EE, Witcher DR, Howard JA, Nikolov ZL. Production and purification of two recombinant proteins from transgenic corn. Biotechnol Prog **1998**; 14:149–55.
4. Arakawa T, Chong DKX, Langridge WHR. Efficacy of a food plant-based oral cholera toxin B subunit vaccine. Nat Biotechnol **1998**; 16:292–7.
5. Tacket CO, Mason HS, Losonsky G, Clements JD, Levine MM, Arntzen CJ. Immunogenicity in humans of a recombinant bacterial antigen delivered in a transgenic potato. Nat Med **1998**; 4:607–9.
6. Pang XL, Joensuu J, Vesikari T. Human calicivirus–associated sporadic gastroenteritis in Finnish children less than two years of age followed prospectively during a rotavirus vaccine trial. Pediatr Infect Dis J **1999**; 18: 420–6.
7. Fankhauser RL, Noel JS, Monroe SS, Ando T, Glass RI. Molecular epidemiology of "Norwalk-like viruses" in outbreaks of gastroenteritis in the United States. J Infect Dis **1998**; 178:1571–8.
8. Vinjé J, Altena SA, Koopmans MPG. The incidence and genetic variability of small round structured viruses in outbreaks of gastroenteritis in The Netherlands. J Infect Dis **1997**; 176:1374–8.
9. Sharp TW, Hyams KC, Watts D, et al. The epidemiology of Norwalk virus during an outbreak of acute gastroenteritis aboard a US aircraft carrier. J Med Virol **1995**; 45:61–7.
10. Jiang X, Wang M, Graham DY, Estes MK. Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. J Virol **1992**; 66:6527–32.
11. Ball JM, Graham DY, Opekun AR, Gilger MA, Guerrero RA, Estes MK. Recombinant Norwalk virus–like particles given orally to volunteers: phase 1 study. Gastroenterology **1999**; 117:40–8.
12. Mason HS, Ball JM, Shi JJ, Estes MK, Arntzen CJ. Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. Proc Natl Acad Sci USA **1996**; 93:5335–40.
13. Parrino TA, Schreiber DS, Trier JS, Kapikian AZ, Blacklow NR. Clinical immunity in acute gastroenteritis caused by Norwalk agent. N Engl J Med **1977**; 297:86–9.
14. Johnson PC, Mathewson JJ, DuPont HL, Greenberg HB. Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults. J Infect Dis **1990**; 161:18–21.
15. Ball JM, Hardy ME, Atmar RL, Conner ME, Estes MK. Oral immunization with recombinant Norwalk virus–like particles induces a systemic and mucosal immune response in mice. J Virol **1998**; 72:1345–53.

Addendum. To determine whether ingestion of the transgenic potato vaccine affected the response to a potato antigen, we measured serum IgG antibodies to potatin, the major protein contained in potatoes, by ELISA by using serum from volunteers who received a potato treated with the Norwalk virus capsid protein or those who received a potato treated with the B subunit of *Escherichia coli* heat labile enterotoxin in a previous study [5]. None of 38 such volunteers had a change in serum anti-potatin IgG after ingesting a transgenic potato vaccine.

PLANT-BASED VACCINES AGAINST DIARRHEAL DISEASES

CAROL O. TACKET, MD

BALTIMORE, MARYLAND

ABSTRACT

Every year 1.6 million deaths occur due to diarrhea related to unsafe water and inadequate sanitation—the vast majority in children under 5 years old. Safe and effective vaccines against enteric infections could contribute to control of these diseases. However, purification of protective antigens for inclusion in vaccines using traditional expression systems is expensive and unattractive to vaccine manufacturers who see the vaccine market as economically uninviting. Cost is one of the persistent barriers to deployment of new vaccines to populations that need them most urgently.

Transgenic plant-derived vaccines offer a new strategy for development of safe, inexpensive vaccines against diarrheal diseases. In phase 1 clinical studies, these vaccines have been safe and immunogenic without the need for a buffer or vehicle other than the plant cell. This paper describes early clinical studies evaluating oral transgenic plant vaccines against enteric infections such as enterotoxigenic *E. coli* infection and norovirus.

Introduction

Transgenic plants have been proposed as a strategy for producing and delivering oral vaccine antigens to protect against a variety of infectious diseases (1–4). An oral plant-based delivery system would be especially useful to prevent diarrheal diseases for which mucosal immune responses at the site of colonization are likely to be protective. Transgenic plant-derived vaccines would be inexpensive to produce, free of pathogenic animal viruses, and scaleable for mass production.

Two important diarrheal pathogens have been targets of vaccine development in recent years—enterotoxigenic *E. coli* (ETEC) and norovirus (NV). ETEC diarrhea is mediated by its heat labile enterotoxin, designated LT. The B subunit of the ETEC toxin (LT-B) binds epithelial cell GM_1 gangliosides and is a potent stimulator of mucosal immune responses when given orally, but is not itself diarrheagenic. Norovirus is a Calicivirus with a single capsid protein that can assem-

From the Center for Vaccine Development, University of Maryland School of Medicine, 685 W. Baltimore Street, Baltimore, MD 21201. 410-706-8437; FAX 410-706-4171; ctacket@medicine.umaryland.edu

ble spontaneously into virus-like particles (VLP) and stimulate immune responses (5). Both LT-B and the capsid protein of norovirus were successfully expressed in plants and induced immune responses in mice after ingestion of transgenic plant material (6,7). This paper describes clinical studies in which transgenic potato or transgenic corn expressing LT-B or NV capsid protein were fed to adult volunteers to determine whether similar immune responses could be generated in humans.

Materials and Methods

Transgenic potato expressing LT-B. A transgenic line of potatoes, designated #TH110-51 derived from Frito Lay variety 1607, was selected for production of potatoes for clinical studies based on a comparatively high level of LT-B expression. Potato tubers were planted in soil, grown to maturity, harvested, and stored at 4°C (8).

Transgenic corn expressing LT-B. Transgenic corn grain was grown and processed using standard commercial techniques. The grain was fractionated and the germ, or embryo, fraction was further treated to remove oil and fat and ground to a small particle size. Isolation of the germ fraction had the effect of concentrating the LT-B molecule. The grinding of the defatted germ resulted in a uniform small particle size, and thus ensured the homogeneity of LT-B in the sample (9).

Transgenic potato expressing NV capsid protein. Transgenic potato line NV140-13 was created by transformation of potato (Frito Lay 1607) with a vector containing an expression cassette for norovirus capsid protein (10). These potatoes were grown and harvested in a manner similar to the transgenic potatoes expressing LT-B.

Clinical studies. Three separate clinical studies were done to evaluate the safety and immunogenicity of (i) transgenic potato expressing LT-B, (ii) transgenic corn expressing LT-B, and (iii) transgenic potato expressing NV capsid protein. All three protocols were reviewed and approved by the University of Maryland Human Institutional Review Board. Volunteers were counseled about research participation and gave informed, written consent.

In the first study, 14 healthy adult volunteers ingested either (i) 100 gm of transgenic potato expressing LT-B ($n = 6$), (ii) 50 gm of transgenic potato expressing LT-B ($n = 5$), or (iii) 50 gm of wild-type potato ($n = 3$) after having been randomized in a double-blind fashion (8). A second and third dose were given on days 7 and 21. Each transgenic potato contained from 3.7 to 15.7 μg/gm of LT-B. This variability may have been due to the tissue specificity of the promoter so that the cloned gene was expressed unevenly in the different tissues of the

potato. The actual amount of LT-B ingested per 50 or 100 gm dose ranged from approximately 0.4 mg/dose to 1.1 mg/dose (mean 0.75 mg/dose). The potatoes were peeled and diced before ingested raw.

In the second study, 13 volunteers were randomized in a double-blind manner to receive either (i) 2.1 gm of defatted LT-B corn germ meal made from transgenic corn containing 1 mg of LT-B (n = 9) or (ii) 2.1 gm of control defatted corn germ meal (n = 4) (9). Additional doses of transgenic or wild-type defatted corn germ meal were given on days 7 and 21. Each 2.1 gm dose of corn germ meal was suspended in 5 ounces of tap water and the suspension ingested.

In the third study, 24 healthy adult volunteers were randomized in a double-blind manner to receive one of three different regimens of transgenic potato expressing NV capsid protein: (i) three doses of transgenic potato on days 0, 7, and 21 (n = 11); (ii) two doses of transgenic potato on days 0 and 21 and a dose of wild-type potato on day 7 (n = 10); or (iii) three doses of wild-type potato on days 0, 7, and 21 (n = 4) (10). These potatoes were also peeled and diced and ingested raw on the day of vaccination.

The volunteers in all 3 studies completed a diary each day for 7 days after ingesting each dose of the transgenic potato or corn germ meal to record the occurrence of nausea, vomiting, cramps, diarrhea, or other symptoms. Blood was collected before and at 7, 14, 21, 28, and 60 days after the first dose of transgenic plant for measurement of serum antibodies to LT or NV capsid protein. Whole blood was collected for antibody secreting cell (ASC) assays on day 0, 7, 14, 21, and 28.

Results

Tolerability. The transgenic potatoes and corn germ meal vaccines were generally well tolerated. There were no differences in the rates of nausea, vomiting, mild cramps, fever, or diarrhea among volunteers who ingested raw transgenic or wild-type potatoes or corn germ meal in the 3 days after ingesting each dose.

Antibody secreting cell responses to LT delivered in potatoes or corn germ meal. Gut-derived antibody secreting cells (ASC) are peripheral blood mononuclear cells (PBMC) that secrete specific antibodies and appear in the circulation about 7 days after mucosal immunization. ASC are no longer detectable in the peripheral blood after about 14 days because they have migrated to a mucosal tissue site. The presence of these cells at 7–10 days after immunization reflects immunologic priming of the gut mucosal immune system.

Before ingestion of potato- or corn-based vaccines, such antigen-

specific cells could not be detected in the peripheral blood of volunteers (Figure 1). At day 7 after ingesting a single dose of transgenic potato expressing LT-B, antibody secreting cells (geometric mean of 18.4 IgA anti-LT antibody secreting cells per 10^6 PBMC; median 41) were detected (Figure 1). At 7 days after the second dose, cells were still detectable, although fewer in number (geometric mean of 6.6 IgA anti-LT ASC per 10^6 PBMC; median 6; range 0–60). On day 28, seven days after the third dose, the number of cells increased (geometric mean of 19.1 IgA anti-LT ASC per 10^6 PBMC; median 20).

A similar magnitude of ASC responses were seen among those who ingested transgenic corn germ meal expressing LT-B, although the kinetics were different in that two doses were needed to see the most striking response (Figure 1). It is important to keep in mind that direct comparison of the ASC immunogenicity of the potato-based and corn-based vaccines is not entirely appropriate since the volunteers were not randomized to receive potato or corn nor were these assays performed in the laboratory on the same day.

Serologic responses to LT delivered in potatoes or corn germ meal. Serologic responses were also detected in volunteers who ingested either potato- and corn-based LT-B vaccines. Ten (91%) of 11 volunteers who ingested transgenic potato and none of those who ingested wild-type potato developed 4-fold rises in IgG anti-LT at some point after immunization. Six (55%) of 11 volunteers who ingested transgenic potato developed 4-fold rises in IgA anti-LT. The IgG antibodies in all responders remained elevated when measured 59 days after ingestion of the first dose. To test the function of the serum antibodies after transgenic potato ingestion, LT neutralization assays

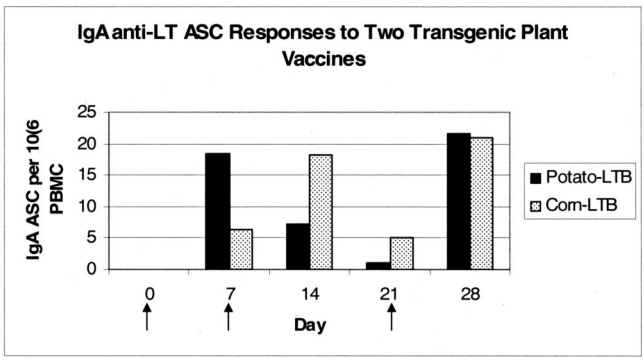


FIG. 1. IgA anti-LT antibody secreting cell responses (ASC) among volunteers who ingested transgenic potato or transgenic corn expressing *E. coli* LT-B in two different clinical studies. Vaccine was administered on days 0, 7, and 21.

were performed in Y-1 adrenal cells. Eight (73%) of 11 volunteers who ingested the transgenic potatoes developed neutralization titers >1:100 at some point after immunization. The peak geometric mean titer (1:267) among transgenic potato eaters was measured on day 28, one week after the third dose (Figure 2).

Seven (78%) of 9 vaccinees who ingested transgenic corn expressing LT-B developed at least 4-fold rises in serum IgG anti-LT after vaccination, as measured by ELISA, usually after the second or third dose of transgenic corn germ meal vaccine. The geometric mean IgG titer peaked at day 56 (1:1165 among responders) (Figure 3). Four (44%) of 9 developed 4-fold rises in serum IgA anti-LT; the IgA titer peaked at day 14 (1:4237 among responders) (Figure 3).

ASC responses to NV capsid protein delivered in potatoes (Figure 4). Among those who ingested 2 or 3 doses of transgenic potatoes expressing NV capsid protein, 19 (95%) of 20 volunteers developed significant rises in the numbers of IgA ASC (range 6–280/10^6 peripheral blood mononuclear cells, PBMC). Thirteen of the 19 IgA ASC responses occurred after the first dose of transgenic potato.

Serologic responses to NV capsid protein delivered in potatoes. The pre-vaccination serum IgG antibody titers varied among these adult volunteers from <1:50 to 1:6400, as expected in an unselected population of adults. The mean pre-vaccination titer did not differ between vaccinees and control volunteers ($p = 0.28$, 2-tailed t test). Four (20%) of 20 volunteers developed serum IgG anti-NVCP, and 4 (20%) of 20 volunteers (3 of whom did not develop IgG responses) developed serum IgM anti-NVCP after ingesting transgenic potatoes. The geometric mean titer of serum IgG anti-NVCP before immunization was 1:67 and after immunization, 1:757 among responders. For serum IgM, the mean titer before immunization was <1:15, and after

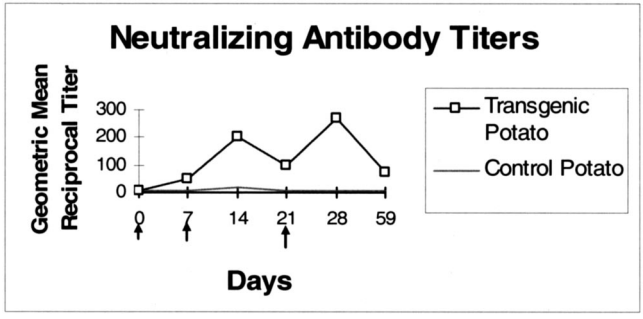


FIG. 2. Geometric mean titer of LT neutralizing antibodies after ingesting transgenic potatoes expressing LT-B. Arrows indicate vaccination days.

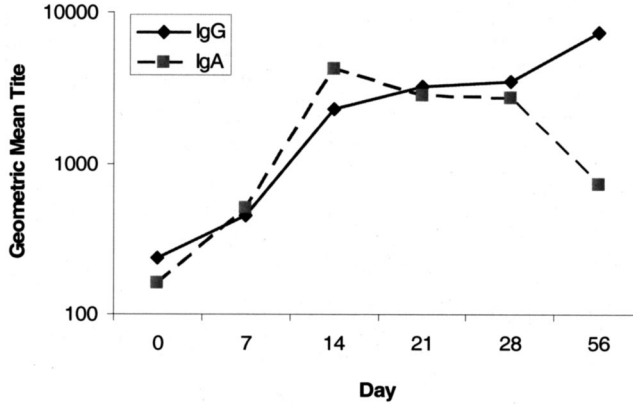


FIG. 3. Geometric mean serum ELISA antibody titers after ingesting transgenic corn germ meal expressing LT-B. Vaccine was administered on days 0, 7, and 21.

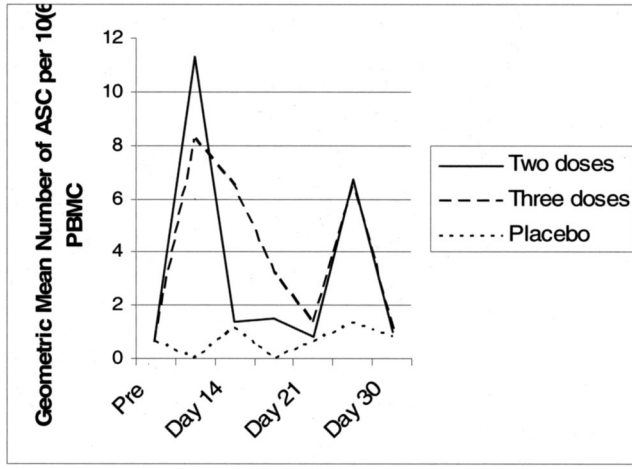


FIG. 4. IgA antibody secreting cell responses to transgenic potatoes expressing NV capsid protein. Vaccine was administered on days 0, 7, and 21.

immunization, 1:100 among responders. Immune responses were not clearly related to the amount of NVCP ingested in each dose, which varied from 215–751 μg per dose. Recipients of wild-type potatoes had no immune responses to NVCP.

Discussion

In this series of three independent studies, two vaccine antigens, *E. coli* LT-B and Norwalk virus capsid protein, delivered by an edible

transgenic plant were processed by the human immune system in a way that resulted in mucosal and systemic immune responses. The ASC responses after ingestion of transgenic potatoes expressing LT-B were not unlike those measured when volunteers were challenged with 10^9 virulent enterotoxigenic *E. coli* organisms given with 2 gm of sodium bicarbonate in a previous study (11). It is remarkable that LT-B delivered in a potato cell, without buffering, induced a roughly similar degree of mucosal B-cell priming as occurs after experimental enterotoxigenic *E. coli* challenge.

These results support a new strategy for production of safe and inexpensive vaccine antigens against diseases for which a protective antigen has been defined, such as tetanus, diphtheria, and hepatitis B. However, ingestion of whole vegetable vaccines, such as transgenic potatoes, is not likely to be a useful way to provide mass vaccinations for several reasons, including: concerns about genetically modified foods in general in some parts of the world, concerns about segregating transgenic food plants from the food supply, and questions about consistency of antigen concentration in different lots of transgenic food. Nevertheless, transgenic plants represent an efficient means of producing purified or partially purified antigens for oral delivery in large amounts and at low cost. For example, gene expression in corn is sufficient to allow commercialization of trypsin and other proteins from corn grain (12), and a number of plants have been used to produce potentially therapeutic monoclonal antibodies (13). The novel vaccines of the future directed against childhood diseases and against newly recognized emerging diseases may include antigens derived from transgenic plant expression systems.

ACKNOWLEDGMENTS

The author is grateful for the contributions of the volunteers who participated in these studies and for the excellent recruiting and nursing staff of the Center for Vaccine Development, University of Maryland School of Medicine. These studies were conducted in collaboration with Hugh S. Mason and Charles J. Arntzen, Arizona State University; John D. Clements at Tulane University; John A. Howard and Stephen Streatfield, ProdiGene; and Mary K. Estes, Baylor College of Medicine. I am also indebted to my colleagues at the Center for Vaccine Development, Genevieve Losonsky, Marcela Pasetti, Robert Edelman, and Myron M. Levine, and to Linda Rosendorf for her invaluable assistance in regulatory affairs.

The studies described herein were supported by NIH contract NO1-AI-45251 to MML; NIH grant RO1-AI36519 to JDC, CJA, and HSM; NIH contract NO1-AI-65299 to COT, and the University of Maryland General Clinical Research Center grant M01 RR165001, NCRR, NIH.

REFERENCES

1. Mason HS, Lam DM-K, Arntzen CJ. Expression of hepatitis B surface antigen in transgenic plants. Proc Natl Acad Sci USA 1992;89:11745–11749.
2. Modelska A, Dietzschold B, Sleysh N, et al. Immunization against rabies with plant-derived antigen. Proc Natl Acad Sci USA 1998;95:2481–2485.
3. Tackaberry ES, Dudani AK, Prior F et al: Development of biopharmaceuticals in plant expression systems: cloning, expression and immunological reactivity of human cytomegalovirus glycoprotein B (UL55) in seeds of transgenic tobacco. Vaccine 1999;17:3020–3029.
4. Arakawa T, Chong DKX, Langridge WHR. Efficacy of a food plant-based oral cholera toxin B subunit vaccine. Nat Biotechnol 1998;16:292–297.
5. Jiang X, Wang M, Graham DY et al. Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. J Virol 1992;66:6527–6532.
6. Mason HS, Ball JM, Shi J-J et al. Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. Proc Natl Acad Sci USA 1996;93:5335–5340.
7. Haq TA, Mason HS, Clements JD et al. Oral immunization with a recombinant bacterial antigen produced in transgenic plants. Science 1995;268:714–716.
8. Tacket CO, Mason HS, Losonsky G et al. Immunogenicity in humans of a recombinant bacterial antigen delivered in a transgenic potato. Nature Medicine 1998;4:607–609.
9. Tacket CO, Pasetti MF, Edelman R et al. Immunogenicity of recombinant LT-B delivered orally to humans in transgenic corn. Vaccine 2004;22:4385–4389.
10. Tacket CO, Losonsky G, Estes MK et al. Immunogenicity of recombinant Norwalk virus capsid protein delivered in transgenic potatoes. J Infect Dis 2000;182:302–305.
11. Tacket CO, Reid RH, Boedeker EC et al. Enteral immunization and challenge of volunteers given enterotoxigenic *E. coli* CFA/II encapsulated in biodegradable microspheres. Vaccine 1994;12:1270–1274.
12. Woodard SL, Mayor JM, Bailey MR et al. Maize-derived bovine trypsin: characterization of the first large-scale, commercial protein product from transgenic plants. Biotechnol Appl Biochem 2003;38:123–130.
13. Fischer R, Hoffmann K, Schillberg S et al. Antibody production by molecular farming in plants. J Biol Regul Homeost Agents 2000;4:83–92.

DISCUSSION

DuPont, Houston: Very nice, Carol. There are new approaches in vaccine development that are being taken. The use of transgenic plants is one area. Transcutaneous patch vaccines are also being used and evaluated in the field. Where do we stand with this vaccine technology? You've been working on it for more than 10 years. Is a company getting into this to develop it commercially? I don't get the feeling that it has moved into clinical trials. I just wonder how far along it has moved. The beauty of this is the cost difference and whereas diarrhea inconveniences people in the United States, it kills infants in developing countries and they don't have much money to be able to afford vaccines. Transgenic plant vaccines may be affordable if they can be developed.

Tacket, Baltimore: Any enteric vaccine that is directed at childhood diarrhea in developing countries, where the vaccines are most desperately needed, is confronted with the problem that there is not a commercial market to drive the development, beyond the travelers' and potentially the military markets. That's been one of the problems with using transgenic plant vaccines as well as live attenuated vaccines for

enteric infections. The edible plant strategy, where a whole vegetable or pieces of a vegetable are consumed, is not logistically feasible because of the reasons that I mentioned, namely, concerns about regulatory issues related to separation of transgenic plants from the food chain and how to be sure that the lot-to-lot consistency requirements are met. But a second strategy, using transgenic plants to produce large amounts of inexpensive purified antigen, is something that could be commercially developed.

Blaser, New York: Carol, that has been very exciting work. I want to go back to your first of the two alternatives—the development of edible vaccines. It seems to me that the big problem is the platform and because Agrobacterium is a pathogen of plants, it causes what you showed were small potatoes and lot-to-lot inconsistencies. So if one could develop a particular line that is adapted so that everything is well understood and controlled, then you could have a generic platform for any number of products.

Tacket: The Agrobacterium plasmid inserts the foreign gene randomly into the plant chromosome. So expression is variable depending on the site of insertion. Expression may also be different in the different tissues within the potato. So, your points are very well taken. The best application of this technology may be in using transgenic plants as an expression system to inexpensively produce antigen. It's a new vaccine development strategy and should be explored more fully.

A census of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions

Melissa D. Michelitsch* and Jonathan S. Weissman*†‡

*Department of Cellular and Molecular Pharmacology, †Howard Hughes Medical Institute, 513 Parnassus Avenue, University of California, San Francisco, CA 94143-0450

Edited by David S. Eisenberg, University of California, Los Angeles, CA, and approved August 25, 2000 (received for review June 2, 2000)

Glutamine/asparagine (Q/N)-rich domains have a high propensity to form self-propagating amyloid fibrils. This phenomenon underlies both prion-based inheritance in yeast and aggregation of a number of proteins involved in human neurodegenerative diseases. To examine the prevalence of this phenomenon, complete proteomic sequences of 31 organisms and several incomplete proteomic sequences were examined for Q/N-rich regions. We found that Q/N-rich regions are essentially absent from the thermophilic bacterial and archaeal proteomes. Moreover, the average Q/N content of the proteins in these organisms is markedly lower than in mesophilic bacteria and eukaryotes. Mesophilic bacterial proteomes contain a small number (0–4) of proteins with Q/N-rich regions. Remarkably, Q/N-rich domains are found in a much larger number of eukaryotic proteins (107–472 per proteome) with diverse biochemical functions. Analyses of these regions argue they have been evolutionarily selected perhaps as modular ''polar zipper'' protein–protein interaction domains. These data also provide a large pool of potential novel prion-forming proteins, two of which have recently been shown to behave as prions in yeast, thus suggesting that aggregation or prion-like regulation of protein function may be a normal regulatory process for many eukaryotic proteins with a wide variety of functions.

Amyloids are self-propagating, β-sheet-rich, fibrillar protein aggregates associated with a number of diseases (1). A variety of proteins with no apparent sequence similarities can form amyloids or related β-sheet-rich aggregates. For an important subset, however, aggregation is mediated by the presence of a glutamine/asparagine (Q/N)-rich domain. These aggregation-prone Q/N-rich sequences have been observed in two distinct contexts. First, expanded pure glutamine repeats are found in several proteins that cause neurodegenerative diseases (2, 3). Second, Q/N-rich domains are found in two *Saccharomyces cerevisiae* proteins, Sup35p and Ure2p, that are responsible for the prion-like inheritance of the yeast non-Mendelian factors [*PSI*$^+$] and [*URE3*], respectively (4–6). In both of these phenomena, an altered conformation of a protein results in the loss of its normal function and the acquisition of the ability to convert the normally functioning form of the protein to the abnormal (prion) form.

In the case of several neurodegenerative diseases involving the deposition of protein aggregates, abnormally long pure glutamine repeats result from the unstable expansion of CAG codons at the DNA level. In all of these diseases, which include Huntington's and Kennedy's diseases and several spinocerebellar ataxias, expanded glutamine repeats in the affected proteins form intranuclear neuronal inclusions (2, 3). Codon expansion beyond a threshold repeat length, typically 37 uninterrupted glutamines, causes aggregation and the onset of disease, both repeat length and disease severity often increasing with successive generations (7, 8).

In contrast to the codon expansion-related neurodegenerative diseases, the Q/N-rich regions of Sup35p and Ure2p are highly interspersed with other amino acids. Several lines of evidence suggest a causative relationship between Q/N content and the formation of prion aggregates. First, these Q/N-rich sequences form the minimal domains required to cause self-propagating aggregation and prion-like inheritance in both of these proteins, although they bear no other sequence similarity (5, 9–12). These Q/N-rich regions drive prion formation by causing self-propagating aggregation and loss of function of the normally soluble forms of these proteins. Second, Q/N content is highly conserved across Sup35p sequences from a number of distantly related yeasts and, as with the *S. cerevisiae* protein, these domains are necessary and sufficient for prion-based inheritance (13–15). Finally, mutations of many glutamines and asparagines in *S. cerevisiae* Sup35p, most often to charged residues, lead to prion curing and protein solubilization (4).

A number of observations from structural studies have led to a model that explains why Q/N-rich regions have such a high propensity to aggregate and/or form fibrils. X-ray diffraction studies of oligoglutamine fibrils have demonstrated that they, like other unrelated amyloids, appear to consist of a cross-β-helix formed with β strands radiating from the helical axis (16, 17). Although no high-resolution data are available for any amyloid fibrils, Perutz and coworkers have created a sterically reasonable model of polar side-chain interactions with main-chain amides in the context of antiparallel β-sheet structures (16).

High Q/N content can be sufficient to cause protein aggregation, and a subset of these Q/N-rich aggregating proteins is capable of stably propagating in a prion-like manner. Therefore, it should be possible to predict novel prion- and/or aggregate-forming proteins on the basis of their Q/N content. In this study, an algorithm that identifies proteins with Q/N-rich regions was devised. By using this algorithm, all of the translated ORFs containing Q/N-rich regions across the available completed proteomic sequences were identified. These data have provided evidence for a conserved function of Q/N-rich domains in eukaryotes, perhaps as a mediator of specific protein–protein interactions. Furthermore, they provide a pool of potential yeast prions and/or amyloid-forming proteins, a few of which have already been tested and shown to act as prions and/or aggregates (13, 18).

Methods

Sequences and Homology Identification. The complete proteomic sequences and the GENPEPT nonredundant protein sequence database for all of the organisms examined in this study were obtained as FASTA format files from either the National Center for Biotechnology Information (NCBI) or the European Bioinformatics Institute. The incomplete proteomic sequences were obtained from SwissProt. Homologies were identified by using PSI-BLAST at the NCBI web site by using the default settings (19).

This paper was submitted directly (Track II) to the PNAS office.

‡To whom reprint requests should be addressed. E-mail: jsw1@itsa.ucsf.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Because of the high Q/N content of the proteins being studied, a filter for low-complexity regions was always used, thereby reducing the probability of predicting false homologues.

Biochemical functions for the yeast proteins identified were assigned on the basis of their classifications at the Yeast Protein Databank (YPD). When none was listed at YPD, putative functions were assigned on the basis of homologies determined by PSI-BLAST by using full-length proteins as query sequences.

Identification of Q/N-Rich Regions. The program DIANA moves through each ORF by single amino acid steps by using a window of arbitrary size determined by the user, returning the sequence of the most Q/N-rich region along with a number of statistics including its amino acid composition, the ORF title, size, and the amino acid composition of the full length protein. These records are then imported into a database for further sorting. When multiple regions occur within a protein that are equally Q/N rich, the algorithm returns the last one. Although high Q/N content was the target of the search described here, DIANA can be modified to identify regions rich in any amino acid or combination of amino acids. The code for DIANA can be obtained at http://itsa.ucsf.edu/~mmichel/DIANA. DIANA was written in PERL 5 and executed on an SGI Octane IRIX6 workstation.

Probability of Occurrence of Q/N Repeats and Q/N-Rich Regions. The probability of random occurrence of a Q/N-rich region was calculated by using the Poisson distribution:

$$f(i) = \frac{e^{-m}m^{i}}{i!},$$

where $f(i)$ is the probability of an event happening i times, and m is the mean percent occurrence of glutamine plus asparagine per predicted ORF in the yeast proteome. The value of i used is 30, the definition of a Q/N-rich region in this study. The value of m is 7.68, the average Q/N content per 80 mer. The resulting value of $f(i)$ implies that one in every 6.336×10^{10} 80 mers will have a Q/N content of at least 30. $f(i)$ was then multiplied by $n = 2.4477283 \times 10^{6}$, the number of consecutive 80 mers in yeast moving in single amino acid increments through each predicted ORF. The resulting value implies that the probability of a Q/N-rich domain occurring randomly given the Q/N content of *S. cerevisiae* is one in every 13,629 proteomes.

Results and Discussion

The Frequency of Q/N-Rich Regions in Eukaryotic, Archaeal, and Bacterial Proteomes. To predict candidates for novel prions or amyloidogenic proteins, an algorithm named DIANA (Defined Interval Amino acid Numerating Algorithm) was designed to identify proteins containing regions of consecutive amino acids with exceptionally high Q/N content. DIANA moves a window of arbitrary size through each translated ORF of a proteome by single amino acid steps. At each step, the total Q/N content within the window is determined. A database is then created containing the sequence of the most Q/N-rich region per protein, its amino acid composition, and the size and amino acid composition of the full length protein.

We used DIANA to examine the predicted translated ORFs for the three eukaryotic, six thermophilic archaeal, two thermophilic bacterial, and 20 mesophilic bacterial genomes available at the National Center for Biotechnology Information at the time of analysis (Fig. 1). In addition, the incomplete proteomic sequences from humans, mice, the plant *Arabidopsis thaliana*, and the GENPEPT nonredundant protein sequence database were examined. The sequences in these incomplete databases are biased toward commonly studied classes of proteins, and therefore they are not included in the following analyses unless specifically stated.

For the purpose of the present analyses, "Q/N-rich regions" are defined as 80 consecutive residues containing at least 30 glutamines and/or asparagines. We chose a window length of 80 amino acids because it is somewhat smaller than the known prion-forming domain of Sup35p (9–12), comparable to the size of prion-forming domain of Ure2p (5) and about twice the size of the minimum glutamine expansion needed to cause disease (2, 3). Therefore, long nonrepeating sequences rich in glutamine and asparagine can be identified without biasing the search against more concentrated and compact pure glutamine or asparagine repeats. For a region to be considered "Q/N-rich," we required that it have a minimum Q/N content of 30. This value was chosen because it is only one less than that of the most Q/N-rich 80 mer of the *Pichia pastoris* Sup35p homologue (13). Furthermore, this value lies beyond the expected distribution of Q/N content for the most Q/N-rich 80 mers for all of the organisms examined (Fig. 2 *a–d*; see below). This conservative definition of a Q/N-rich region allows us to focus on only the most promising candidates for proteins capable of forming prions and/or amyloids.

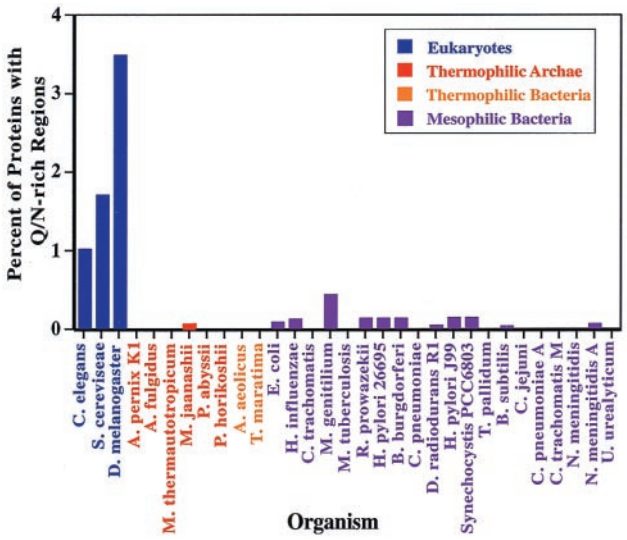


Fig. 1. Many eukaryotic proteins contain Q/N-rich regions. The percent of ORFs with a region of 80 consecutive amino acids containing a Q/N content of at least 30 was determined for the indicated proteomes.

Even with these relatively stringent criteria, a surprisingly large number of proteins with Q/N-rich regions were identified in eukaryotes (Figs. 1 and 2*a*). In *S. cerevisiae*, 107 polypeptides with Q/N-rich regions (1.69% of all ORFs) were identified, 143 were found in *Caenorhabditis elegans* (1.00% of all ORFs), and 472 were found in *Drosophila melanogaster* (3.47% of all ORFs).

A number of the eukaryotic sequences examined contained sequences with exceptionally high Q/N content. The region identified with the greatest total Q/N content was found in an adenylyl cyclase from *Dictyostelium discoideum*, with a total Q/N content of 72. This region also contains the highest glutamine content, with 71 glutamines and the longest uninterrupted glutamine repeat with 45 tandem glutamines. The proteins with the most asparagine-rich region and the longest asparagine repeat are also found in *Dictyostelium.* A prespore-specific protein-containing Q/N-rich region with 68 asparagines is the most asparagine-rich domain identified; the longest uninterrupted asparagine repeat consisted of 49 tandem asparagines and is found in a protein–tyrosine phosphatase.

All of the eukaryotic proteomes examined contained a subset of sequences with relatively pure glutamine or asparagine repeats reminiscent of the codon repeat disorders. Interestingly, none of these regions can be attributed solely to trinucleotide expansion, as all exhibit substantial codon variation. However, roughly three-fourths of the Q/N-rich sequences have a com-

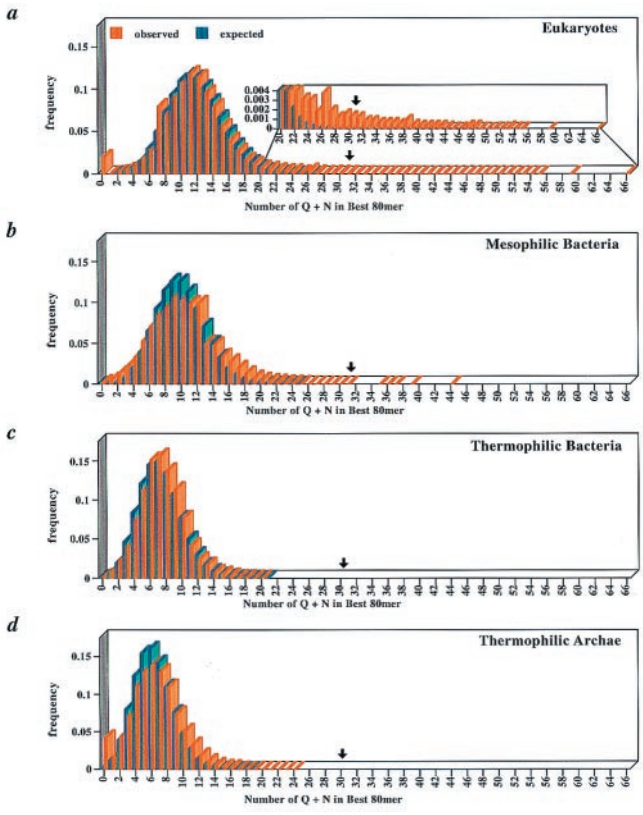


Fig. 2. Expected and observed distributions of Q/N content of the best 80 mers across mesophilic and thermophilic proteomes. A comparison of the expected and observed distributions of Q/N contents across the most Q/N-rich regions of (*a*) eukaryotes, (*b*) mesophilic bacteria, (*c*) thermophilic bacteria, and (*d*) archaea. Arrows indicate the minimum Q/N content required for a domain to be considered Q/N rich. (*a Inset*) Magnification of the observed and expected distributions of Q/N-rich regions. Expected values were calculated on the basis of the average Q/N content per best 80 mer for each of the four types of organisms.

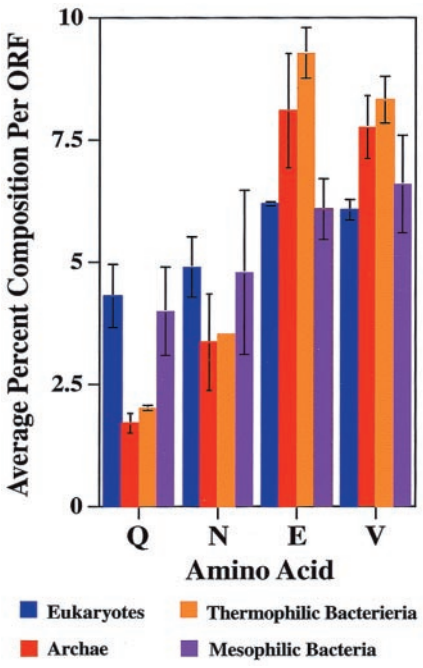


Fig. 3. Thermophiles have a lower Q/N content and higher glutamate and valine content than mesophiles. The average percent amino acid content per ORF per proteome was determined for the completed proteomes. Shown are glutamine (Q), asparagine (N), glutamate (E), and valine (V). The other amino acids showed little variation. Error bars indicate the observed variance among the indicated class.

position more similar to the yeast prion-forming domains in that they are interspersed with other amino acids.

In eleven of twenty of the mesophilic bacterial proteomes queried, Q/N-rich regions were present in a small number of proteins (one to four) (Figs. 1 and 2*b*). The other nine proteomes contain no Q/N-rich regions. In contrast to the eukaryotic Q/N-rich proteins that are predominantly cytoplasmic, a number of bacterial outer-membrane proteins were identified. None of these bacterial proteins contain long uninterrupted glutamine or asparagine repeats. Interestingly, over half of the bacterial Q/N-rich regions contain imperfect repeats of glutamine and/or asparagine in combination with other amino acids. Imperfectly repeating sequences have been suggested to promote and stabilize prion formation in the yeast protein Sup35p, although at present there is no evidence for prion-like behavior in bacteria (20).

In marked contrast to the eukaryotes and mesophilic bacteria, Q/N-rich regions are nearly absent from the proteomes of all of the thermophilic organisms tested, including all six archaea and both thermophilic bacteria (Figs. 1 and 2 *c* and *d*). Among these eight proteomes, only one predicted protein sequence was found that contained a Q/N-rich region. Furthermore, short tandem repeats of glutamine or asparagine residues rarely occur in thermophiles. The longest glutamine or asparagine repeat in three of the complete thermophilic proteomes examined was four residues long and only three residues long in the other five proteomes, whereas the longest consecutive glutamine or asparagine repeat is thirty-seven residues long and seven across all of the complete eukaryotic and mesophilic bacterial proteomes, respectively. These observations suggest that tandem Q/N repeats may be specifically selected against in thermophiles, whereas there is a strong positive selection for both Q/N-rich regions and tandem repeats in eukaryotes (see below).

Glutamine and Asparagine Are Less Abundant in Proteins from Thermophiles. One possible explanation for the lack of Q/N-rich domains in thermophiles might be that these amino acids are underrepresented in these organisms. Indeed, Haney and coworkers have noted that a set of proteins from thermophilic *Methanococcus* species tend to have fewer polar residues (serine, threonine, glutamine, and asparagine) and more charged residues (especially glutamate, aspartate, and lysine) than homologous proteins from mesophilic *Methanococcus* species (20). However, studies comparing the relative amino acid frequencies in proteins across entire proteomes of thermophiles and mesophiles have not previously been reported.

We determined the relative abundance of the different amino acids for all of the completely sequenced proteomes to determine whether the difference in the number of Q/N-rich domains found in thermophiles and mesophiles could simply result from differences in overall amino acid compositions (Figs. 2 and 3). Interestingly, the amino acid that differs the most in abundance between mesophiles and thermophiles is glutamine, with its relative frequency in thermophilic archaea and bacteria being less than half of that of the mesophilic bacteria and eukaryotes (1.85% vs. 4.31%). Asparagine is also on average less abundant in thermophiles, although the asparagine content shows greater variability in the archaea. By contrast, in thermophiles, glutamate is significantly enriched and valine is modestly enriched, 1.5- and 1.3-fold, respectively. There were no systematic differences in the relative abundance of the other amino acids (data not shown). Thus, although Q/N-rich regions would not have been expected to occur at random in either thermophilic or mesophilic organism, their chance occurrence in thermophiles is even less likely (Fig. 2 *a–d* and see below).

Why would thermophiles have evolved to have lower Q/N content than mesophiles? One possibility is chemical lability. Glutamine/asparagine deamination and glutamate/aspartate succinamide formation occur readily in peptides at 100°C (22). However,

there is evidence that these modifications rarely occur in well-ordered proteins. Thus, chemical instability should be cited as a reason for reduced Q/N content only if considered in the context of structure. Because Q/N-rich regions tend to form relatively unstructured loops in unliganded native proteins, they would be expected to be particularly susceptible to deamination (23). In addition to promoting chemical stability at high temperatures, the lower Q/N content could promote fold stability. Thermophilic proteins have been shown to have fewer loops than their mesophilic counterparts, presumably to attain a more compact structure to achieve greater thermostability (24). Finally, lower Q/N content might reduce protein aggregation. Although Q/N-rich regions are aggregation prone at lower temperatures as well, this property is likely to be exacerbated in thermophiles. In contrast, the presence of charged residues such as glutamate has been shown to reduce or prevent the aggregation of Q/N-rich regions (4).

Amino Acid Conservation with Codon Variation Suggests a Conserved Function for Q/N-Rich Regions. Although there are several plausible reasons why Q/N-rich regions and even short tandem repeats would be disfavored in thermophiles, the question remains whether there is appreciable selection for Q/N-rich regions in proteins from mesophilic organisms. The Poisson distribution function is used to predict the frequencies with which independent events will occur given a known mean frequency of occurrence. Deviation from this expected distribution is an indicator of nonrandom distribution (25). Poisson distributions were calculated for the eukaryotes, archaea, and mesophilic and thermophilic bacteria by using the average Q/N content of the best 80 mer from each protein within the proteomes as the expected mean. The distributions predicted from the Poisson function were then compared with the actual distribution of the Q/N contents of the best 80 mers across all of the proteomes examined (Fig. 2). In every case, the bulk of the distribution closely follows the expected distribution. However, for the mesophilic organisms, especially the eukaryotes, there are exceptionally long tails extending far beyond the expected values. Given the average Q/N content of *S. cerevisiae*, a consecutive 80 amino acid regions with a Q/N content of 30 would randomly occur only once in every 13,629 proteomes of the same size. Furthermore, the probability of a region having a Q/N content of 66 like the Q/N-rich region of Snf5p is vanishingly small (1 in 5.64×10^{32} proteomes). The probabilities of these very Q/N-rich regions occurring are so miniscule that it is extraordinarily unlikely that the presence of these domains is merely because of chance.

An alternative explanation for the presence of these domains is that they are the result of codon expansion rather than functional selection at the protein level. However, few of the polypeptides identified in this study contain long uninterrupted tracts of polyglutamine or polyasparagine. Moreover, the small number of repeat sequences that are present have codon variation. For example, the 37 consecutive glutamines of Snf5p are encoded by interspersed CAA and CAG codons with no more than eight consecutive CAGs and nine consecutive CAAs. This observation suggests that they are conserved sequences and are not merely trinucleotide expansions resulting from aberrant DNA metabolism. Finally, the presence of a large number of Q/N-rich regions that are enriched in both glutamine and asparagine suggests there is positive evolutionary selection for the chemical nature of their amido side chains.

Have these Q/N-rich regions been conserved because they play a common role in a particular class of proteins? The known or predicted biochemical functions of all of the proteins containing Q/N-rich regions in *S. cerevisiae* are listed in Table 1. (see supplementary material at www.pnas.org for a complete list of the ORFs identified.) Interestingly, proteins with a broad range of biochemical functions were found to contain Q/N-rich domains, although the identified proteins can be placed in functional clusters including transcription and translation factors, nucleoporins, DNA- and RNA-binding proteins, and proteins involved in vesicular trafficking. These observations indicate that the function of Q/N-rich domains is not specific to any single class of protein, but that their presence has been conserved across several of the diverse protein classes in which they occur.

The extreme improbability of these Q/N-rich regions occurring randomly, the lack of evidence for codon expansion, and the observation that their occurrence is conserved in a broad spectrum of proteins strongly support the conclusion that the Q/N content of these regions is because of positive evolutionary selection. Furthermore, because these regions are compact and can be found fused internally or at either terminus to a variety of different unrelated proteins, they are likely to behave as functionally conserved modular domains.

Identification of Potential Prion- and Amyloid-Forming Proteins. The two extensively characterized yeast prion phenomena were not discovered in directed screens for novel prions; rather, their existence was deduced only after decades of careful observation of their non-Mendelian inheritance (26, 27). Given the past lack of methodology for systematically identifying prions and the large pool of candidate Q/N-rich regions, it seems probable that significantly more prions remain to be identified.

It is likely that many of the proteins containing Q/N-rich regions will aggregate when overexpressed. Which of these will propagate stably, thus allowing prion-based inheritance, is less clear and needs be established experimentally on a case-by-case basis. Recently, Santoso and coworkers and, independently, Sondheimer and Lindquist described two similar genetic systems that should now allow the rapid testing of Q/N-rich domains for the ability to support stable prion-like inheritance (13, 18).

In our laboratory, two of the *S. cerevisiae* proteins identified by DIANA were tested for their ability to behave as prions and/or aggregates. The first protein examined, New1p, was previously uncharacterized but contains a Q/N-rich region with very similar imperfect repeats to those in the prion-forming domain of Sup35p, including the identical pattern YQQGGYXYN. The second protein tested, Pan1p, is involved in actin cytoskeletal organization that, apart from its Q/N richness, bears no similarity to either known yeast prion (28). Both were found to aggregate when overexpressed in *S. cerevisiae*, but with the experimental system used, only New1p was observed to behave as a prion (13) (L. Z. Osherovich and J.S.W., unpublished data).

DIANA also identified other sequences with strong similarities beyond high Q/N content to the prion-forming domain of the Sup35 protein. These proteins include YBR016, YDR210W, and Rnq1p. Indeed, while this study was being completed, Sondheimer and Lindquist demonstrated that a chimera consisting of the Q/N-rich domain of Rnq1p fused to the nonprion-forming domain of Sup35p is capable of exhibiting prion-like behavior (18). Moreover, the endogenous protein acts as a prion. In addition, they showed that a number of proteins with sequence similarities to Sup35p are capable of causing the aggregation of GFP *in vivo*. There are also many promising candidates whose Q/N-rich domains strongly resemble the Q/N-rich domain of Ure2p, including YIL130 and YLR278. Most of these domains contain long stretches of asparagine interspersed with other amino acids, particularly serine.

Although proteins with statistically significant similarities to Sup35p or Ure2p are especially promising candidates, it is likely that a wide array of Q/N-rich sequences can behave as prions, because Ure2p and Sup35p are not homologous to each other, and their prion-forming domains have very dissimilar amino acid compositions. For example, YOR197 is a predicted ORF of unknown function whose Q/N-rich region contains imperfect repeats of QQYG. These repeats are reminiscent of the imperfect PQGGYQQYN repeats present in Sup35p that have been shown to promote conversion to the prion state (20). In addition

Table 1. Proteins containing Q/N-rich regions in *S. cerevisiae* fall into a broad spectrum of functional clusters

Cluster	Protein	Q/N
Nonkinase signaling proteins		
	GPR1	55
	RPI1	35
	PHO81	34
Kinases		
	NPR1, ARP1	52
	CBK1	49
	YCK1	40
	YCK2	36
	HRR25	36
	APG13	36
	YAK1	33
	SKY1	32
	KSP1	32
	SCH9	31
	SKS1	31
	YBR059	38
DNA/RNA binding		
	MPT5, HTR1	31
	YPR042	51
	YBR180	34
	YGL014	34
	YLL013	33
RNA processing		
	NAB1	48
	LSM4	38
	PCF11	37
	NAB3	34
	PUB1	33
	RNA15	30
	PSP2	33
	NGR1	31
	YBL051	30
Nucleoporins		
	NUP116	36
	NUP100	32
	NUP57	31
	NUP49	30
Transcription		
	SNF5	66
	CYC8	54
	MCM1	49
	IXR1, ORD1	49
	MED2	48
	GAL11	47
	SWI11	43
	TAF61	43
	PDR1	38
	URE2	38
	DAL81	38
	MOT3, HMS1	38
	RLM1	36
	SNF2, GAM1	36
	SWI4	35
	AZF1	35
	CCR4	34
	CRZ1	32
	PDC2	32
	DAT1	31
	SFP1	30
	SPT20	30
	CDC39	30
	YIL130	55
	YEL007	46
	YDR409	37
	POP2	37
	YMR263	35
	EPL1	32
	TBS1	31
	YLR373	30
Translation		
	VAR1	40
	SUP35	38
	NEW1, YPL226	32
	TIF4632	30
Vessicular trafficing/secretory pathway		
	ENT2	48
	SEC61	45
	YAP1801	39
	PAN1	38
	SLA2	38
	ENT1	36
	SCD5	36
	FAB1	32
	ANP1	31
	YKL054	53
Outer membrane proteins		
	YDR213	44
	YPR022	42
	PSP1	31
	YLR177	30
Other		
	RNQ1	42
	CDC27	35
	MAD1	34
	GRR1	32
	JSN1	31
	SLF1	30
	VAC7	30
	YKR096	38
	YGL066	37
	YKL088	32
Unknown		
	MSS11	52
	HOT1	33
	YBR016	41
	YBL081	38
	YIL105	36
	YBR238	33
	YBL029	31
	YOR197	30
	YML053	30

S. cerevisiae proteins containing Q/N-rich regions were placed into functional groups on the basis of either known biochemical function or sequence homology to proteins of known function (indicated by italics).

to the presence of repeats, YOR197 is intriguing because, as in huntingtin, a proline-rich domain follows its Q/N-rich domain.

It will also be interesting to study the function and aggregation properties of some proteins independently of whether they also prove to be novel prions. For example, YGL066 is an uncharacterized protein with significant sequence homology ($E = 2 \times 10^{-6}$) to human ataxin-7, a protein that causes a human spinocerebellar ataxia in its mutant glutamine-expanded form (29). This homology is external to the Q/N-rich region, suggesting a true conserved biochemical function rather than a mere reflection of similar amino acid content in their Q/N-rich regions. Studying the function of this protein in yeast could help elucidate the currently unknown function of its disease-causing human counterpart.

The growing number of known yeast prions together with the large pool of Q/N-rich candidates raises the possibility that the formation of prion-like aggregates is a conserved and beneficial

cellular process that occurs in a broad range of proteins. Consistent with this proposal, recent studies reveal that the ability of Sup35p to support prion-based inheritance is strongly conserved across diverse species of budding yeast (13–15). In addition to the established role of [*PSI*$^+$] in regulating translation termination (6), Tuite and coworkers recently showed that in some strains, [*PSI*$^+$] yeast show an increased tolerance to thermal stress and high ethanol concentrations compared with isogenic strains lacking the [*PSI*$^+$] prion (30). Moreover, the [Het-s] prion phenomenon has a well-characterized role in regulating heterokaryon incompatibility in the fungus *Podospora anserina* (31).

Why might prion-based inheritance be advantageous to an organism? Prion formation provides a mechanism for a protein to inhibit its own activity by specifically self aggregating in response to overexpression or to other changes in the environment (30, 32). Potentially, a major advantage of prion-based functional inhibition would be that, unlike protein inactivation caused by DNA mutations, the inhibited prion-like state can be propagated indefinitely while retaining the ability to revert to the original functional state. Finally, because prion-forming domains and other Q/N-rich domains occur in proteins with a variety of functions, it is possible that a broad range of proteins can be regulated by the aggregation of Q/N-rich domains. In this regard, it is significant that prion-like aggregate formation is highly self specific, thereby allowing multiple different prion-states to propagate independently within a single cell (13).

The challenge now is to determine experimentally which of these Q/N-rich domains can behave as prions and/or aggregates. This information should help both in refining prion prediction algorithms and in understanding the physiologic role for prions and/or other self-specific protein aggregates.

Q/N-Rich Domains as Modulators of Specific Protein–Protein Interactions? It is clear that there has been a positive selective pressure for certain proteins to contain Q/N-rich domains. What function besides prion-based inheritance do these domains impart that would make their presence advantageous?

Perutz has suggested that Q/N-rich regions (and possibly regions rich in other polar residues) might behave as modular mediators of protein–protein interactions termed "polar zippers" because of the capacity of their side chains to form hydrogen bond networks (16). Supporting this hypothesis, there have been several recent experimental reports of domains containing Q/N-rich regions mediating specific protein–protein interactions. First, the glutamine-rich region in transcription factor Sp1 was shown by functional mapping *in vivo* to bind the dTAFII110 component of the *Drosophila* TFIID complex (33). Second, Pan1p and Sla1p, two proteins that act in complex to promote cytoskeletal organization in yeast, were shown by coimmunoprecipitation and two-hybrid analysis to interact via their domains containing Q/N-rich regions (28). Finally, The Sla1p Q/N-rich domain and Sup35p prion-forming domains interact in two-hybrid studies (34). Moreover, this interaction appears to influence the rate of conversion to the prion state by Sup35p. It will be important to directly demonstrate the role of glutamine and asparagine residues in stabilizing these interactions.

The roles of Q/N-rich domains in mediating both protein interactions and in promoting aggregation could be different manifestations of the same modular functionality. Indeed, it is possible that prion and/or aggregate formation is a conserved function of certain Q/N-rich regions that results from their ability to mediate protein–protein interactions in a self-propagating manner. By empirically examining the roles of the Q/N-rich domains identified in this study, the possible intersection of their functions as mediators of protein–protein interactions and the formation of both prion and nonprion aggregates can be elucidated.

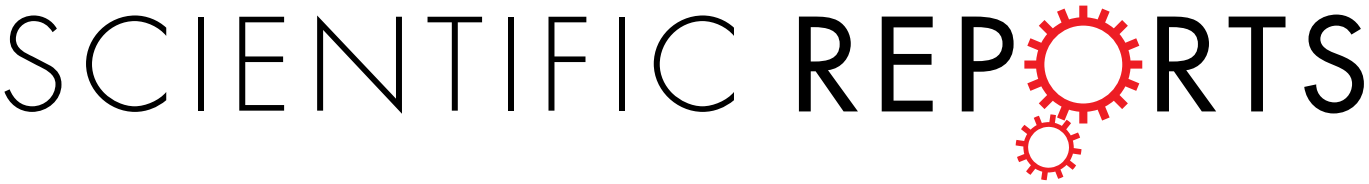
Summary

By identifying all of the Q/N-rich regions across a number of completely sequenced proteomes, we have revealed that Q/N-rich domains are ubiquitous and modular, are associated with a variety of types of proteins, and have been evolutionarily conserved in a number of these proteins. This conservation suggests an independent functional role for these domains. Growing evidence suggests that Q/N-rich domains function as mediators of specific protein–protein interactions. We argue that their conserved ability to form self-specific prion and nonprion aggregates might be simply an extension of their ability to act as interaction domains by forming extensive, self-propagating, hydrogen-bonding networks.

In addition to providing evidence for conserved modular functions for Q/N-rich regions, the identification of Q/N-rich domains in eukaryotes has proven useful for identifying two novel prions and a nonprion-aggregating protein (13, 18) (L. Z. Osherovich and J.S.W., unpublished data). By using a new genetic system and biochemical methods, additional prions and amyloidogenic proteins can now be rapidly identified, enabling refinement of the criteria used to predict prion-like or amyloidogenic behavior. This information will facilitate the analysis of prion-based inheritance in normal cellular physiology. Finally, as the Human Genome Project is completed, the algorithm described here in combination with other experimental observations may be useful for identifying new proteins prone to forming disease-causing aggregates.

We thank P. Babbitt, P. Bosque, F. Cohen, A. DePace, I. Kuntz, J. Nguyen, L. Osherovich, A. Santoso, H. Sparrer, R. Stroud, and members of the Weissman, Cohen, and Babbitt groups for helpful discussions. This work was supported by the Searle Scholars Program, the David and Lucile Packard Foundation, the National Institutes of Health, and a National Science Foundation predoctoral fellowship (M.D.M.).

1. Koo, E. H., Lansbury, P. T., Jr. & Kelly, J. W. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9989–9990.
2. Perutz, M. F. (1999) *Trends Biochem. Sci.* **24,** 58–63.
3. Ross, C. A., Margolis, R. L., Becher, M. W., Wood, J. D., Engelender, S., Cooper, J. K. & Sharp, A. H. (1998) *Prog. Brain Res.* **117,** 397–419.
4. DePace, A. H., Santoso, A., Hillner, P. & Weissman, J. S. (1998) *Cell* **93,** 1241–1252.
5. Maddelein, M. L. & Wickner, R. B. (1999) *Mol. Cell. Biol.* **19,** 4516–4524.
6. Serio, T. R. & Lindquist, S. L. (1999) *Annu. Rev. Cell. Dev. Biol.* **15,** 661–703.
7. Stine, O. C., Pleasant, N., Franz, M. L., Abbott, M. H., Folstein, S. E. & Ross, C. A. (1993) *Hum. Mol. Genet.* **2,** 1547–1549.
8. Klockgether, T. & Evert, B. (1998) *Trends Neurosci.* **21,** 413–418.
9. Ter-Avanesyan, M. D., Kushnirov, V. V., Dagkesamanskaya, A. R., Didichenko, S. A., Chernoff, Y. O., Inge-Vechtomov, S. G. & Smirnov, V. N. (1993) *Mol. Microbiol.* **7,** 683–692.
10. Patino, M. M., Liu, J. J., Glover, J. R. & Lindquist, S. (1996) *Science* **273,** 622–626.
11. Ter-Avanesyan, M. D., Dagkesamanskaya, A. R., Kushnirov, V. V. & Smirnov, V. N. (1994) *Genetics* **137,** 671–676.
12. Derkatch, I. L., Chernoff, Y. O., Kushnirov, V. V., Inge-Vechtomov, S. G. & Liebman, S. W. (1996) *Genetics* **144,** 1375–1386.
13. Santoso, A., Chien, P., Osherovich, L. Z. & Weissman, J. S. (2000) *Cell* **100,** 277–288.
14. Chernoff, Y. O., Galkin, A., Lewitin, E., Chernova, T. A., Newnam, G. P. & Belenkiy, S. M. (2000) *Mol. Microbiol.* **35,** 865–876.
15. Kushnirov, V. V., Kochneva-Pervukhova, N., Chechenova, M. B., Frolova, N. S. & Ter-Avanesyan, M. D. (2000) *EMBO J.* **19,** 324–331.
16. Perutz, M. F., Johnson, T., Suzuki, M. & Finch, J. T. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5355–5358.
17. Sunde, M. & Black, C. (1997) *Adv. Protein Chem.* **50,** 123–159.
18. Sondheimer, N. & Lindquist, S. (2000) *Mol. Cell* **5,** 163–172.
19. Liu, J. J. & Lindquist, S. (1999) *Nature (London)* **400,** 573–576.
20. Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R. & Olsen, G. J. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3578–3583.
21. Daniel, R. M., Dines, M. & Petach, H. H. (1996) *Biochem. J.* **317,** 1–11.
22. Wright, P. E. & Dyson, H. J. (1999) *J. Mol. Biol.* **293,** 321–331.
23. Thompson, M. J. & Eisenberg, D. (1999) *J. Mol. Biol.* **290,** 595–604.
24. Bulmer, M. (1979) *Principles of Statistics* (Dover, New York).
25. Cox, B. S., Tuite, M. F. & McLaughlin, C. S. (1988) *Yeast* **4,** 159–178.
26. Wickner, R. B. (1994) *Science* **264,** 566–569.
27. Tang, H. Y., Xu, J. & Cai, M. J. (2000) *Mol. Cell. Biol.* **20,** 12–25.
28. David, G., Abbas, N., Stevanin, G., Dürr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E., *et al.* (1997) *Nat. Genet.* **17,** 65–70.
29. Eaglestone, S. S., Cox, B. S. & Tuite, M. F. (1999) *EMBO J.* **18,** 1974–1981.
30. Coustou, V., Deleu, C., Saupe, S. & Begueret, J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 9773–9778.
31. Chernoff, Y. O., Derkach, I. L. & Inge-Vechtomov, S. G. (1993) *Curr. Genet.* **24,** 268–270.
32. Gill, G., Pascal, E., Tseng, Z. H. & Tjian, R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 192–196.
33. Bailleul, P. A., Newnam, G. P., Steenbergen, J. N. & Chernoff, Y. O. (1999) *Genetics* **153,** 81–94.
34. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.



OPEN

Received: 17 August 2016
Accepted: 02 November 2016
Published: 19 December 2016

An integrated multi-omics analysis of the NK603 Roundup-tolerant GM maize reveals metabolism disturbances caused by the transformation process

Robin Mesnage[1,*], Sarah Z. Agapito-Tenfen[2,*], Vinicius Vilperte[3], George Renney[4], Malcolm Ward[4], Gilles-Eric Séralini[5], Rubens O. Nodari[3] & Michael N. Antoniou[1]

Glyphosate tolerant genetically modified (GM) maize NK603 was assessed as 'substantially equivalent' to its isogenic counterpart by a nutrient composition analysis in order to be granted market approval. We have applied contemporary in depth molecular profiling methods of NK603 maize kernels (sprayed or unsprayed with Roundup) and the isogenic corn to reassess its substantial equivalence status. Proteome profiles of the maize kernels revealed alterations in the levels of enzymes of glycolysis and TCA cycle pathways, which were reflective of an imbalance in energy metabolism. Changes in proteins and metabolites of glutathione metabolism were indicative of increased oxidative stress. The most pronounced metabolome differences between NK603 and its isogenic counterpart consisted of an increase in polyamines including N-acetyl-cadaverine (2.9-fold), N-acetylputrescine (1.8-fold), putrescine (2.7-fold) and cadaverine (28-fold), which depending on context can be either protective or a cause of toxicity. Our molecular profiling results show that NK603 and its isogenic control are not substantially equivalent.

The application of genetic engineering (GE) to modify edible crops is often advocated as one of the most important scientific advances to improve farming systems and feed the world in a more sustainable manner[1]. GE has been used to create crops adapted to abiotic stress, resistant to pathogens, with a longer shelf life, or with enhanced nutritional properties. However, commercialization of these traits is currently minor. Agricultural genetically modified (GM) crops are dominated by plants engineered to tolerate application of a herbicide or/and to produce their own insecticides[2]. A total of 180 million hectares of GM crops are currently cultivated worldwide on around 1.5 billion hectares constituting approximately 10% of global arable land[3]. Approximately 80% of GM crops have been modified to tolerate application of and thus accumulate glyphosate-based herbicide residues without dying in order to facilitate weed management.

Regulations for the release of genetically modified organisms (GMOs) of any kind in a country are covered by the national biosafety regulations of that nation. Guidance on risk assessment (RA) aim at identifying and avoiding adverse effects by early detection and proper evaluation of intended and potential unintended changes in a GMO. These should be detected and identified at early stages of RA, often referred to as "hazard identification". Hazard identification is essential to the RA process as it sets the foundation of what is considered or observed in later steps in the risk assessment process[4]. In the US, the Food and Drug Administration considers

[1]Gene Expression and Therapy Group, King's College London, Faculty of Life Sciences & Medicine, Department of Medical and Molecular Genetics, 8th Floor, Tower Wing, Guy's Hospital, Great Maze Pond, London SE1 9RT, United Kingdom. [2]Genøk, Center for Biosafety, The Science Park, P.O. Box 6418 Tromsø 9294, Norway. [3]CropScience Department, Federal University of Santa Catarina, Rod. Admar Gonzaga 1346, 88034-000 Florianópolis, Brazil. [4]Proteomics Facility, King's College London, Institute of Psychiatry, London SE5 8AF, United Kingdom. [5]University of Caen, Institute of Biology, EA 2608 and Network on Risks, Quality and Sustainable Environment, MRSH, Esplanade de la Paix, University of Caen, Caen 14032, Cedex, France. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.N.A. (email: michael.antoniou@kcl.ac.uk)

GM technology as an extension of conventional breeding and GMO crops are deregulated once nutritional and compositional "substantial equivalence" is demonstrated[5]. The set of parameters and analyses necessary to declare a GMO as substantially equivalent to its conventional counterpart is still vague and focuses on a restricted set of compositional variables, such as the amounts of protein, carbohydrate, vitamins and minerals. GMOs are then declared substantially equivalent when sufficient similarities appear for those selected variables[6]. Remarkably, while a majority of GMO crops have been modified to withstand and thus accumulate a herbicide without dying, analysis for residues for such pesticides are neglected in compositional assessment[7].

Recent technologies used to ascertain the molecular compositional profile of a system, such as transcriptomics, proteomics, metabolomics, epigenomics and mirnomics, collectively referred to as "*omics* technologies", are used extensively in basic and applied science[8]. Comparative *omics* analyses have been performed comparing GMO crops and their isogenic counterpart. A number of them have shown metabolic disturbances from potential unintended effects of the GM transformation process in Bt maize[9–12], glyphosate-tolerant soybean[13–15], potato[16], cotton[17] and rice[18]. However, these studies do not report consistent or coherent results, which can be explained by the use of a variety of genetic backgrounds and/or different growth conditions, as well as variations in the technologies and threshold levels applied[19]. Indeed, the majority of authors of these types of studies conclude that the statistically significant changes observed between the conventional and the GM varieties are not biologically significant because they fall into the range of variations obtained in the comparisons between different conventionally-bred varieties, and under different environmental conditions[11]. However, other authors conclude that observed differences could reflect biologically significant, GM transformation process induced changes in protein profiles[12] or metabolism[20] when appropriate near-isogenic controls were applied and test crops grown at the same time and location to avoid differences brought about by variable environmental conditions[20]. Currently, no regulatory authority requires mandatory untargeted molecular profiling *omics* analysis to be performed but some acknowledge their potential relevance for food and feed derived from GM plants with specific metabolic pathways modified, or in situations where a suitable comparator is not available[4,21].

Despite being declared to be 'substantially equivalent', off target effects have been observed in non-target species for Bt toxin-producing GMO crops[22–24]. Additionally, laboratory animal feeding trials performed with some GM plants in comparison to the non-GM counterpart have been proposed to provide evidence of ill-health effects. Several laboratory studies consisting of 90-day feeding trials in rodents have been conducted to evaluate the safety of GMO crop consumption[25,26]. These investigations have frequently resulted in statistically significant differences in parameters reflective of disturbances in various organ systems and in particular liver and kidney biochemistry, but with interpretation of their biological significance, especially with respect to health implications, being controversial[27–29]. Such differences in outcome in such laboratory animal feeding studies could have multiple sources including the presence of GMO-associated pesticide residues[30,31].

In an effort to provide insight into the substantial equivalence classification of a Roundup tolerant NK603 GM maize, we have performed proteomics and metabolomics analyses of NK603 (sprayed or unsprayed with Roundup) and isogenic maize kernels (Fig. 1). We used a TMT10plex™ isobaric mass tag labelling method and quantified proteins by Liquid chromatography-tandem mass spectrometry (LC-MS/MS). The metabolome profile was determined by ultrahigh performance liquid chromatography-tandem mass spectroscopy (UPLC-MS/MS). Altogether, our integrative analysis shows that the GM transformation process used to generate NK603 maize caused deep alterations in the proteome and metabolome profiles of this crop and results in marked metabolic changes. We conclude that NK603 maize is not compositionally equivalent to its non-GM isogenic counterpart as previously claimed.

Results

The objective of this investigation was to obtain a deeper understanding of the biology of the NK603 GM maize by molecular profiling (proteomics and metabolomics) in order to gain insight into its substantial equivalence classification. We began by undertaking an unsupervised exploratory analysis of variance structure. We integrated metabolome and proteome profiles of the NK603, cultivated either with or without Roundup, and its isogenic counterpart, into a two-step multiple co-inertia analysis (MCIA) process. First, a one-table ordination method transforms each multidimensional dataset (hyperspaces) separately into comparable lower dimensional spaces by finding axes maximizing the sum of the variances of the variables. The resulting variance structure can be described by a PCA (Additional file 3). The results show a clear separation of each feed type (NK603, NK603+ Roundup and control) in both platforms. Control samples had the most distinct proteome and metabolome profiles as observed in PCA plots.

In a second step, the variance structure analyses from metabolome and proteome profiles were combined into a single analysis (Fig. 2). This aims to find new axes on which the two hyperspaces are projected by maximizing the square covariance. Figure 2A shows the projection of metabolome and proteome profiles onto the first two principal components of MCIA. Absolute eigenvalues of these components are given by a bar plot (Fig. 2B). The transgenic feed samples NK603 and NK603+Roundup are separated from the non-transgenic control (Isogenic) along the first component (horizontal axis). This clustering accounts for most of the variation (percentage of explained variance of 56.7%). The NK603 maize sprayed with Roundup separates from the unsprayed NK603 maize on the second component (vertical axis, percentage of explained variance of 16.6%). The lines connecting the different dots are proportional to the divergence between the different variables of the dataset. A relatively high correlation is depicted by the short edges. It shows similar trends in metabolome and proteome profiles, and also between the two cultivations, indicating that the most variant sources of biological information were similar. The projection of individual protein or metabolites on a 2-dimensional space (Fig. 2C) showed a mix pattern indicating that no particular subsets of variables are driving the clustering of groups. Finally, Fig. 2D shows the pseudo-eigenvalues space. The proteome samples (blue and green dots) are highly weighted on the horizontal axis indicating that this dataset is the highest contributor of the clustering of the transgenic feed samples from the

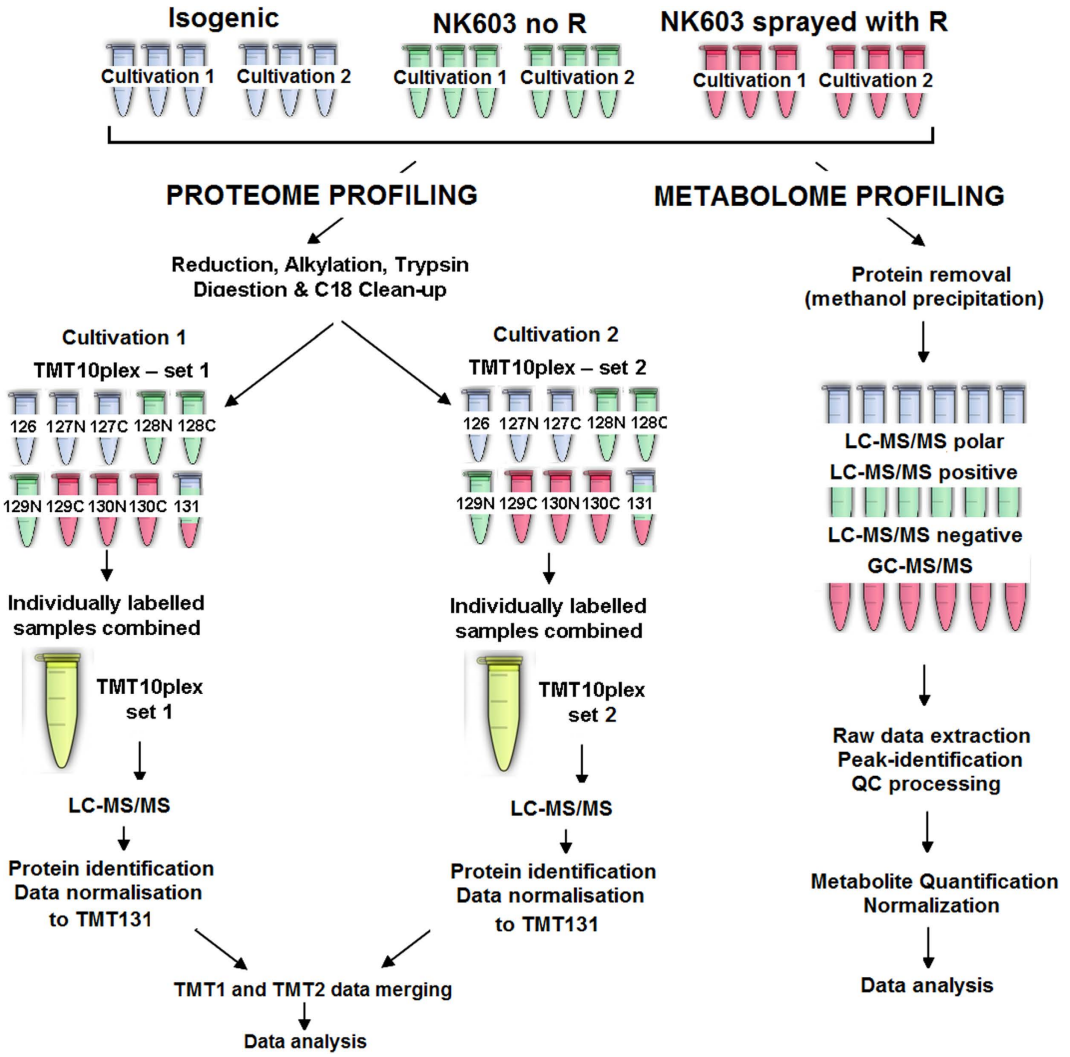


Figure 1. Flowchart of the experimental procedure. Harvested grains from NK603 GM maize cultivations, sprayed (NK603 + R) or not (NK603) with Roundup, were compared to their nearest isogenic non-transgenic control (Isogenic) grown under similar normal conditions. Two biological replicates were obtained by performing two cultivations at the same location in different years. Maize grains were analyzed by different mass spectrometry methods to determine proteome and metabolome profiles in 3 technical replicates.

control. By contrast, the differences between the NK603 maize sprayed with Roundup and the unsprayed NK603 maize are mostly due to the composition of the metabolome since the latest has a high weight on the vertical axis (red and black dots) of the pseudo-eigenvalues space. The fold changes observed in the comparisons of the NK603 maize sprayed with Roundup, the unsprayed NK603 maize and the isogenic control corn were highly correlated between the two cultivations performed during two different growing seasons (Additional file 4). Overall, the MCIA shows that the GM transformation process was the major contributor to variation in the protein and metabolite profiles rather than environmental factors such as the spraying of a pesticide or the growing season.

We next conducted a statistical evaluation of the biological differences resulting from the GM transformation process, as well as from the spraying of Roundup, by pairwise comparisons in order to identify proteins and metabolites associated with possible metabolic alterations. The list of proteins and metabolites having their levels significantly disturbed is given in Additional files 5 and 6, respectively. Figure 3 shows the statistical significance of differential protein/metabolite levels by volcano plots along with respective fold changes. While only one protein is newly produced as a result of the transgene insertion, a total of 117 proteins and 91 metabolites have been altered in maize by the genetic transformation process and insertion of the EPSPS-CP4 cassette (Isogenic vs NK603 panel, Fig. 3). One protein (B4G0K5) and 31 metabolites had their expression significantly altered by the spraying of the Roundup pesticide (NK603 vs NK603+Roundup (R) panel, Fig. 3).

The NK603 maize has been engineered to express a modified version of the *Agrobacterium tumefaciens* strain EPSPS-CP4[32]. Two peptides (lAGGEDVADLR and gLGNASGAAVATHLDHR) from EPSPS-CP4[33] were detected and quantified by undertaking a specific targeted data analysis (Fig. 4). Their location on EPSPS-CP4 is shown by Fig. 4C. Reporter ion intensities for EPSPS-CP4 peptides in the NK603+Roundup and the NK603

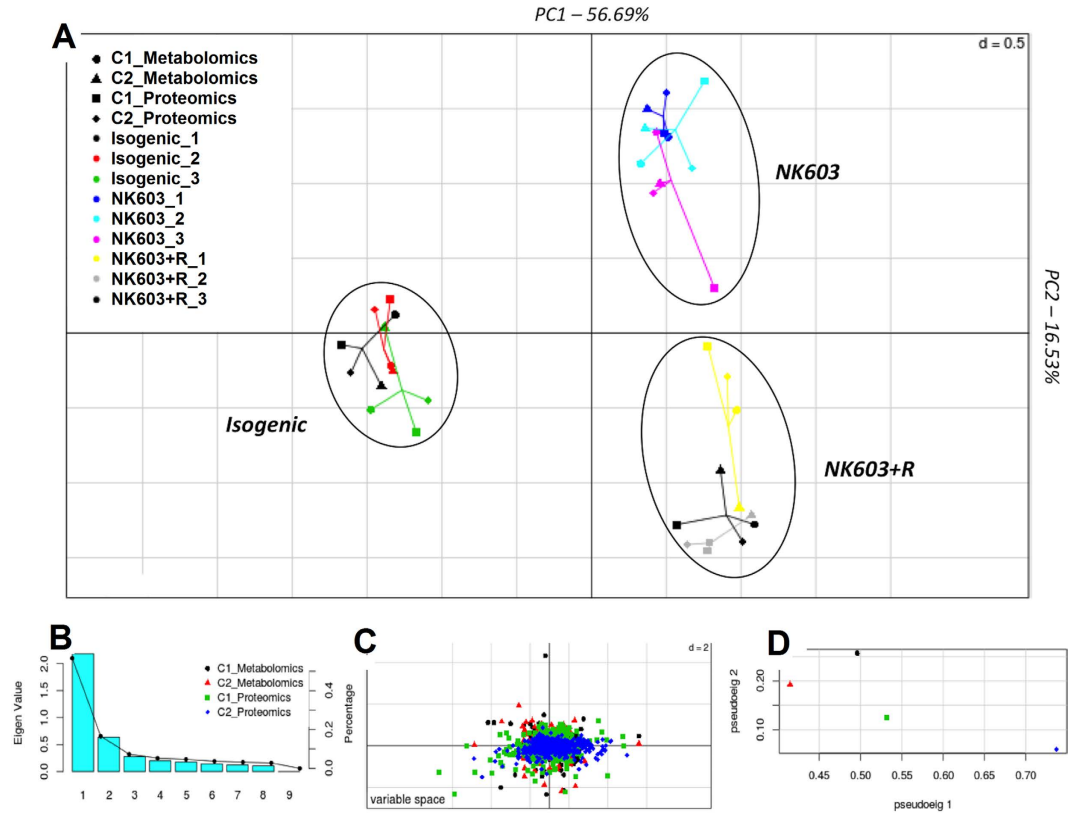


Figure 2. Integration of metabolome and proteome profiles of the NK603 maize and its near-isogenic counterpart into a multiple co-inertia analysis projection plot. (**A**) The first two axes of MCIA represent metabolome and proteomic datasets. Different shapes represent the different variables which are connected by lines, the length of these lines is proportional to the divergence between the data. Lines for each sample are joined at a common point at which the covariance derived from the MCIA analysis is maximal. (**B**) Pseudo-eigenvalue space showing the percentage of variance explained by each of the MCIA component. Each barplot represents the absolute eigenvalues. (**C**) Protein or metabolites (colored dots) are projected on a 2-dimensional space. In this panel, a protein or a metabolite that is particularly highly expressed in a maize variety will be located on the direction of this variety. (**D**) Pseudo-eigenvalues space of all datasets, indicating how much variance of an eigenvalue is contributed by the proteome or the metabolome for cultivations 1 and 2.

were on average respectively 7 and 10 times higher than in the isogenic control. The observed signal for the non-transgenic corn probably represents non-specific background noise since it does not contain the EPSPS-CP4 gene. This would be caused by the co-isolation of other peptides in the corresponding MS/MS experiment, which gives rise to low intensity reporter ions in the control channels.

We analysed the biological information contained in proteome profiles from the NK603 and its isogenic counterpart to see if they bear a signature representative of metabolic disturbances caused by the insertion of the transgene cassette and/or the expression of bacterial EPSPS-CP4. Among different pathway enrichment analysis software tested, String was chosen due to its in-house predictions and homology transfers, as well as its connection to many fine external database resources, and thus its ability to identify a larger number of proteins. Nevertheless, our interpretation remained limited by the quality of protein annotation in such databases. A total of 42.7% (50/117) and 35% (55/156) of the proteins respectively disturbed in the comparison to the unsprayed or the sprayed NK603 maize were uncharacterized or not annotated in the databases (Additional file 5).

Pathway enrichment analysis of differentially expressed proteins in NK603 and NK603+Roundup feed samples was mainly assigned to carbohydrate and energy metabolism (Table 1). Most of the proteins, including enzymes, associated with these pathways were overexpressed in GM samples (Additional file 5). An increased expression of some proteins involved in glycolysis (FDR adjusted p-value = 4.2e-7), and in particular in the synthesis of pyruvate from D-glyceraldehyde 3-phosphate can be indicative of an increased demand for energy. Among them, pyruvate kinase (B4F9G8), enolase (ENO1), and three glyceraldehyde-3-phosphate dehydrogenases (GAPC1, GAPC2, GAPC3) had their levels increased in NK603 maize. Interestingly, gene ontology terms related to metabolic responses to stress were enriched (FDR adjusted p-value = 1.5e-6) and some heat shock proteins (e.g., HSP82) have been overexpressed.

The comparison between Roundup-sprayed NK603 and control samples revealed a similar pattern to that observed in unsprayed samples. However, glutathione metabolism (KEGG ID 480) showed a significant alteration in sprayed NK603. The proteins assigned to that pathway, glutathione S-transferase 1 and 6-phosphogluconate

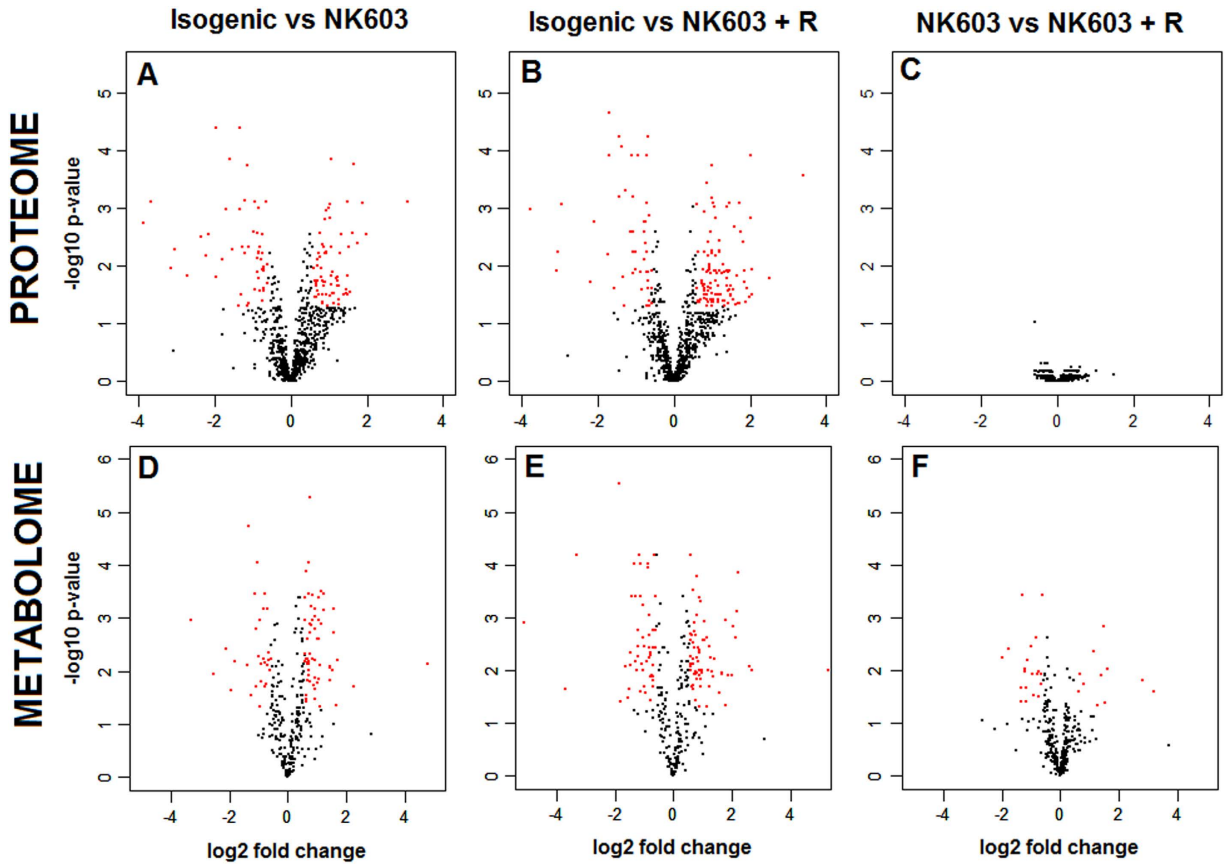


Figure 3. Volcano plots of the maize proteome and metabolome profiles. Volcano plots show the log 2 fold changes and the −log10 adjusted p-values in protein or metabolite level induced by the GM transformation process (isogenic vs NK603, isogenic vs NK603 + R) or by the pesticide spraying (NK603 vs NK603 + R). Data were selected at the cut off values adj-p < 0.05 and fold change >1.5. Red dots represent protein or metabolites having their level significantly altered in the different samples.

dehydrogenase (P12653 and B4FSV6 respectively) were more abundant in sprayed samples while another glutathione transferase isoform GST-5 (A0A0B4J3E6) was less abundant. Additionally, the 1-Cys peroxiredoxin PER and the peroxidase were overexpressed. Although only one protein was statistically significantly altered in a pairwise comparison between NK603+Roundup and NK603 as the effect of Roundup herbicide spray alone, the protein B4G0K5 that has an identified conserved domain of Ricin-type β-trefoil lectin. The Ricin-type β-trefoil is a carbohydrate-binding domain found in a variety of molecules serving diverse functions such as enzymatic activity, inhibitory toxicity and signal transduction[34].

The composition of the metabolome is shown in Additional file 6. The most pronounced differences between the NK603 GM maize and its isogenic counterpart mostly consisted of an increase in the amounts of numerous polyamines. The levels of N-acetyl-cadaverine (2.9-fold), N-acetylputrescine (1.8-fold), putrescine (2.7-fold) and cadaverine (28-fold) were increased in NK603. The metabolome profile also highlighted an impairment of energy metabolism. While metabolites from the first part of the TCA cycle had their levels increased (α-ketoglutarate by 1.65-fold and citrate by 1.49-fold), metabolites from the second part of the TCA cycle had their levels decreased (malate by 0.59-fold, fumarate by 0.60-fold, succinate by 0.80-fold). Additionally, while proteins associated with glycolysis were overexpressed, carbohydrate metabolism is depleted in several metabolites (glucuronate by 0.63-fold, glucose 1-phosphate by 0.56-fold, maltohexaose by 0.28-fold, maltopentaose by 0.51-fold). Differences due to the pesticide spray were subtle: phenylpropanoid such as 4-hydroxycinnamate (0.63-fold), ferulate (0.59-fold) and sinapate (2.9-fold) were significantly changed. While alterations of the shikimate pathway were not detected, intermediates from aromatic amino acid metabolism (PEP derived) had their level increased (phenyllactate by 1.60-fold, phenylpyruvate by 2.71-fold, N-acetylphenylalanine by 2.24-fold and xanthurenate by 1.82-fold). These changes could be indicative of an increase in amino acid catabolism. However, of note is that PEP itself was not detected in the analysis.

Table 2 provides pathway enrichment analysis of metabolites that were found to be statistically significantly altered in the pairwise comparisons. For the metabolome pathways analysis, the profile of NK603 and NK603+R showed a distinct pattern compared to the profiles observed in the proteome analysis. From the 10 most altered pathways, these two samples shared only five altered pathways and these suggest an alteration due to the GM transformation process. These pathways revealed an alteration in aspartate, pyruvate and phenylalanine amino acid downstream processes. The NK603 metabolome profile seems to differ from sprayed samples by fatty acid

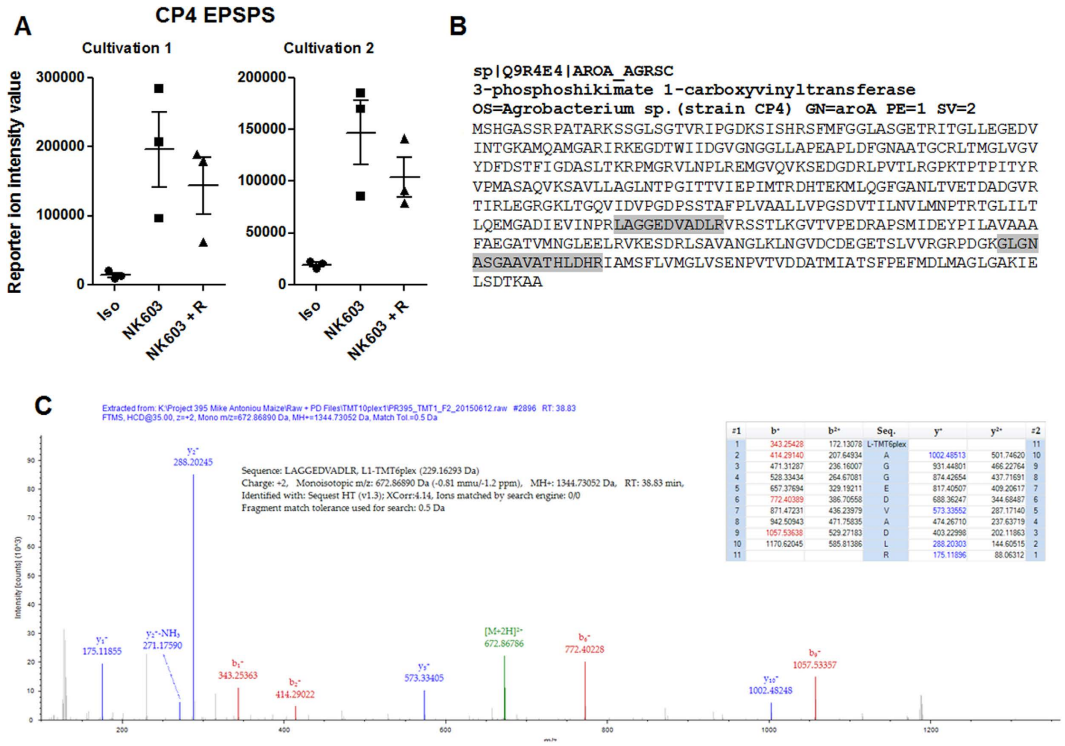


Figure 4. Mass spectrometric detection of CP4 EPSPS in the NK603 genetically modified maize. Two different peptides from Agrobacterium sp. 3-phosphoshikimate 1-carboxyvinyltransferase have been detected (gLGNASGAAVATHLDHR and lAGGEDVADLR) in all different samples allowing semi-quantitation (**A**) Reporter intensity ion values pertaining for CP4 EPSPS in the different samples of the two cultivations. (**B**) Localization of the peptides on the CP4 EPSPS (in grey) (**C**) Spectrum from the detection of the lAGGEDVADLR pertaining to Agrobacterium CP4 EPSPS (cultivation 1 of NK603).

related pathways and choline, nicotinate and nicotinamide metabolism while sprayed samples showed alterations in serine metabolism and other sugar related metabolism.

The STITCH tool was used to provide a visualisation of predicted interactions of chemicals and proteins that might have a link to the transgene-associated EPSPS-CP4 pathway. The interaction network reveals that some proteins or metabolites altered in the NK603 maize are interacting with EPSPS (Fig. 5). The network formed by these proteins/metabolites is centred on some TCA cycle intermediates, among them, the α-ketoglutarate. One should note that EPSPS is using an energy metabolism intermediate (phosphenolpyruvate) as substrate. Overall, our data shows that the expression of a heterologous EPSPS in the NK603 maize is causing a deep alteration in the proteome and metabolome profiles of feed samples and thus resulting in a metabolic imbalance.

Discussion

In this report we present the first *multi-omics* analysis of GM NK603 maize compared to a near isogenic non-GM counterpart. Based on analysis conducted by the developer Monsanto Company, NK603 maize was scored as 'substantially equivalent' to its isogenic control, which was a major contributor to this product being granted market approval for animal and human consumption in the European Union, United States, Brazil and several other nations. Although NK603 had comparable nutritional and compositional profiles when originally accessed by the developer company upon registration of their product, our analysis at a detailed, in-depth molecular profiling level shows that NK603 grains, with or without Roundup spraying during cultivation, are not equivalent to isogenic non-transgenic control samples (Fig. 2).

The concept of substantial equivalence has long being used in safety testing of GMO crops, but the term and the concept has no clear definition[35]. In 1993 the Organization for Economic Co-operation and Development (OECD) stated that the "*concept of substantial equivalence embodies the idea that existing organisms used as food, or as a source of food, can be used as the basis for comparison when assessing the safety of human consumption of a food or food component that has been modified or is new*"[36]. The vagueness of this term generates conflict among stakeholders to determine which compositional differences are sufficient to declare a GMO as non-substantially equivalent. However, the Codex Alimentarius Commission[37] makes it clear that a safety assessment of a new food based on the concept of substantial equivalence *"does not imply absolute safety of the new product; rather, it focuses on assessing the safety of any identified differences so that the safety of the new product can be considered relative to its conventional counterpart."* Thus, the concept of substantial equivalence should not be used as a proof of safety. However, it could be used as a first tier in risk assessment to detect any unintended effects of the GM transformation process. Unintended effects can be understood as the effects that go beyond the primary expected

	Term description	n	p-adj
Isogenic × NK603			
KEGG ID			
1110	Biosynthesis of secondary metabolites	12	6.2E-11
1200	Carbon metabolism	8	1.6E-10
1120	Microbial metabolism in diverse environments	8	5.8E-10
1100	Metabolic pathways	13	1.3E-09
1230	Biosynthesis of amino acids	7	5.7E-09
710	Carbon fixation in photosynthetic organisms	5	3.6E-08
10	Glycolysis/Gluconeogenesis	5	4.2E-07
51	Fructose and mannose metabolism	3	9.4E-05
30	Pentose phosphate pathway	3	9.4E-05
520	Amino sugar and nucleotide sugar metabolism	3	6.6E-04
GO ID			
GO:0008150	Biological process	16	1.3E-18
GO:0008152	Metabolic process	13	5.0E-15
GO:0005975	Carbohydrate metabolic process	8	1.9E-14
GO:0016052	Carbohydrate catabolic process	6	4.0E-14
GO:0009987	Cellular process	12	6.1E-14
GO:0044699	Single-organism process	11	8.4E-14
GO:1901135	Carbohydrate derivative metabolic process	7	5.3E-13
GO:1901575	Organic substance catabolic process	6	5.9E-13
GO:0044238	Primary metabolic process	10	2.8E-12
GO:0071704	Organic substance metabolic process	10	5.9E-12
Isogenic × NK603 + R			
KEGG ID			
1110	Biosynthesis of secondary metabolites	14	2.3E-12
1200	Carbon metabolism	9	1.6E-11
710	Carbon fixation in photosynthetic organisms	7	1.6E-11
1120	Microbial metabolism in diverse environments	9	6.7E-11
1100	Metabolic pathways	15	1.3E-10
1230	Biosynthesis of amino acids	8	5.4E-10
10	Glycolysis/Gluconeogenesis	6	2.1E-08
51	Fructose and mannose metabolism	3	1.9E-04
30	Pentose phosphate pathway	3	1.9E-04
480	Glutathione metabolism	3	1.0E-03
GO ID			
GO:0008150	Biological process	21	2.3E-25
GO:0008152	Metabolic process	16	1.2E-18
GO:0009987	Cellular process	15	2.1E-17
GO:0016052	Carbohydrate catabolic process	7	1.3E-16
GO:0005975	Carbohydrate metabolic process	9	3.3E-16
GO:0044699	Single-organism process	13	4.3E-16
GO:0044723	Single-organism carbohydrate metabolic process	8	3.0E-15
GO:1901575	Organic substance catabolic process	7	3.0E-15
GO:0006757	ATP generation from ADP	6	3.0E-15
GO:0072524	Pyridine-containing compound metabolic process	6	3.0E-15

Table 1. Pathway enrichment analysis in proteome profiles of the maize samples. Among different analytical software, String was chosen as it recognized a maximum number of proteins. The maize genome was used as a background list to calculate the p-values of each term. The 10 most enriched GO biological process terms and KEGG pathways (ranked by p-values) are presented. N, number of protein disturbed in each pathway; p-adj, fdr adjusted p-value.

effects of the genetic modification, and represent statistically significant differences in the GMO compared with an appropriate control[38]. Unintended effects during transgenesis include rearrangements, insertion, or deletions during the genetic transformation or during the tissue culture stages of GMO development[39,40]. A comprehensive characterization of the GM plant at the molecular level could facilitate identification of unintended effects in GMO crops and could be used as a complementary analytical tool to existing safety assessment procedures[41–44].

	n/N	ES	p-value
Iso vs NK603 pathways			
Amines and polyamines	3/4	2.6	0.0370
Nicotinate and nicotinamide metabolism	3/4	2.6	0.0370
Fatty acid, Dicarboxylate	4/8	1.7	0.0797
Aspartate family (OAA derived)	10/27	1.3	0.0908
TCA cycle	3/6	1.7	0.1300
Dipeptide	1/17	0.2	0.1359
Free fatty acid	7/19	1.3	0.1557
Branched Chain Amino Acids (pyruvate derived)	5/13	1.3	0.1780
Phenylpropanoids	3/7	1.5	0.1914
Photorespiration	2/4	1.7	0.2192
Iso vs NK603 + R pathways			
Amino sugar and nucleotide sugar	7/8	2.3	0.0007
Serine family (phosphoglycerate derived)	7/10	1.8	0.0062
Dipeptide	10/17	1.5	0.0096
TCA cycle	4/6	1.7	0.0532
Branched Chain Amino Acids (pyruvate derived)	7/13	1.4	0.0536
Aspartate family (OAA derived)	12/27	1.2	0.0730
Benzenoids	2/2	2.6	0.0769
Phenylpropanoids	4/7	1.5	0.0978
Glycolysis	3/5	1.6	0.1342
Inositol metabolism	2/3	1.7	0.1885
NK603 vs NK603 + R pathways			
Phenylpropanoids	3/7	4.3	0.0180
Amino sugar and nucleotide sugar	3/8	3.8	0.0271
Fatty acid conjugate	1/1	10.1	0.0899
Branched Chain Amino Acids (pyruvate derived)	3/13	2.3	0.1012
Choline metabolism	1/2	5.1	0.1719
gamma-glutamyl	2/10	2.0	0.2238
Serine family (phosphoglycerate derived)	2/10	2.0	0.2238
Fatty acid amide	1/3	3.4	0.2467
Nicotinate and nicotinamide metabolism	1/4	2.5	0.3150
Glycolysis	1/5	2.0	0.3773

Table 2. Pathway enrichment analysis in metabolome profiles of the maize samples. The 10 most altered pathways (ranked by p-values) are presented. The number of metabolites disturbed in each pathway (n) is compared to the total number of metabolites measured for the given pathway (N). Enrichment scores (ES) for each pathway are calculated as follow: ES = (number of significant metabolites in pathway/total number of detected metabolites in pathway)/(total number of significant metabolites/total number of detected metabolites). The p-values were calculated according to a one sided Fisher exact test.

In general, our study design further highlights the importance of restricting comparison to the GMO crop and non-GMO isogenic comparator and cultivation of the two at the same location and season when the objective is to evaluate the effect of the GM transformation process. This is obligatory in order to reduce effects on plant metabolism arising from differing environmental conditions, which can make it difficult to attribute differences that are observed to the procedure of transgenesis. However, even though our experimental design takes into account the effect of the growing season, further experiments made under different environmental conditions would be needed to determine the full range of effects of the GM transformation process on NK603 phenotype. Indeed, virtually all traits are influenced by genotype–environment interactions. Neither genetic differences nor environmental variations alone can account for the production of a particular phenotypic variation. For example, a study of the expression of the transgene encoding a Bt toxin in the MON810 GM maize under different environmental conditions, has shown that the phenotype resulting from the GM transformation process is influenced by stressful environmental conditions[45].

The increasing literature reporting application of *omics* methods to assess proteome, metabolome and transcriptome profiles in GMO crops shows strong evidence of distinct grain proteomes in other GM maize events, such as MON810 Bt insecticide producing maize[11,12,46]. Although the majority of studies have focused on insect-resistant maize (e.g., MON810 event) and most likely because this was the first GM maize to enter the food and feed market, there has also been one previous metabolomics study investigating NK603. Metabolite profiling of NK603 maize kernels were analyzed and approximately 3% of the metabolites detected showed statistically significant differences compared to the respective isogenic lines[47]. Two metabolites (γ-tocopherol and myo-inositol)

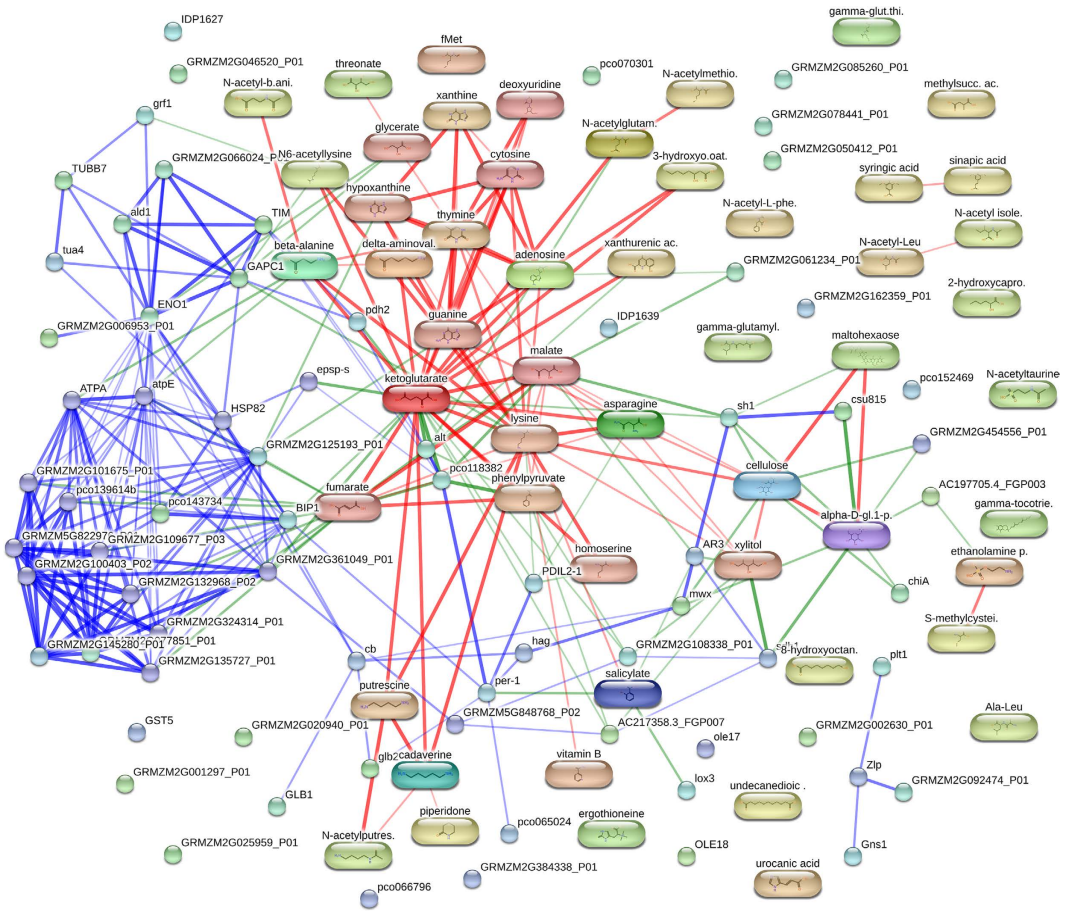


Figure 5. Interaction network of the metabolic effects resulting from the GM transformation process. The STITCH ('Search Tool for Interacting Chemicals') tool was used to provide a visualisation of the consequence of EPSPS insertion. The proteins and the metabolites which were found commonly deregulated in the two comparisons of NK603 and NK603 + R to its isogenic counterpart were used as input list. The top ten interaction partners with the highest scores as well as the maize EPSPS were included to reveal interactions.

were less abundant in NK603[47]. Interestingly, γ-tocotrienol and myo-inositol levels were also found to be significantly reduced in our study, and thus attributable to the genetic transformation. This suggests that some metabolic alterations are consistently reported despite a strong background triggered by environmental influence. In a study of two common MON810/non-GM variety pairs subjected to two farming practices (conventional and low-nitrogen fertilization), it was found that up to 37.4% of the variation was dependent upon the variety, 31.9% were the result of the fertilization treatment, and 9.7% was attributable to the GM character[48].

Alterations can also be found in other plant tissues. For example, analysis of leaves of Brazilian varieties of MON810 Bt maize revealed a total of 32 differentially expressed proteins between GM and non-GM samples that were identified and assigned to carbohydrate and energy metabolism, genetic information processing and stress response[9].

Our study revealed significant metabolome profile differences between NK603 that was either sprayed or not with Roundup during cultivation (Fig. 2). This was surprising since the single application of this herbicide was prior to development of the maize cobs. In addition, we did not detect glyphosate or AMPA residues in the test maize kernel samples (Additional File 1). This indicates that metabolic differences provoked by an early application of Roundup persisted throughout the life of the maize even in the absence of herbicide residues. At present we can only speculate as to the mechanisms that may explain these effects but they may have their basis in epigenetic programming of gene expression patterns with consequent longer term effects. The spraying of Roundup could have acted as a signal causing an alteration in gene expression patterns in the growing maize. A recent study that demonstrated marked epigenetic (DNA methylation) changes in *A thaliana* in response to treatment with carbendazim supports this possibility[49]. In addition, it has been demonstrated that epigenetic (DNA methylation and post-translational histone modification) patterns acquired in one cultivation can be transgenerationally inherited in an *A thaliana* model system[50]. However, further research would be needed to determine if epigenetic alterations provoked by pesticide exposure can hamper plant phenotypes across generations.

The maize kernels analysed in this study were previously used to feed laboratory animals that formed part of a chronic (2 year) study looking at potential toxic effects arising from the consumption of this NK603 Roundup-tolerant GM maize. A dry feed was formulated to contain 11%, 22%, or 33% of NK603 maize, cultivated either with or without Roundup application, or 33% of the near isogenic variety. Sprague Dawley rats fed

for two years on these diets presented blood/urine biochemical changes indicative of an increased incidence of liver and kidney structure and functional pathology in the NK603-containing diet groups compared to non-GM controls[51]. Standard biochemical compositional analysis revealed no particular differences between the different maize types tested[51]. Metabolic disturbances observed in our study may help to understand the negative health effects suggested after the chronic consumption of this GM maize. Alterations in concentrations of metabolites in grains might be directly related to pathogenic effects due to some active compounds that are known to be toxic[52]. For instance, a soybean glycoprotein allergen (Gly m Bd 28 K fragment) was also found overexpressed in a proteomic study of Roundup Ready GM soybean seeds (MSOY 7575 RR event)[13]. In our study, cadaverin levels were significantly increased (Log2FC 4.81 for NK603 and 5.31 for NK603+Roundup). Cadaverin plays important roles in lysine biosynthesis[53] and also glutathione metabolism[54]. Other similar biogenic amines, such as N-acetyl-cadaverine, N-acetylputrescine and putrescine were also found to be present at higher levels in NK603 in our investigation. Different polyamines have been reported to have different effects, which depend on various factors such as age, tissue or disease status[55]. In certain contexts some of these polyamines have been found to be protective whereas in other situations they can be a cause of toxicity. On the one hand, toxicological effects such as nausea, headaches, rashes and changes in blood pressure are provoked by the consumption of foods with high concentrations of polyamines[56]. Putrescine and cadaverine have been reported as potentiators of the effects of histamine, and both have been implicated in the formation of carcinogenic nitrosamines with nitrite in meat products[57]. On the other hand, certain polyamines can also have beneficial anti-inflammatory effects and have been found to be beneficial during aging in some rodent model systems[58]. Noticeably, these polyamines were not measured in the first compositional analysis of NK603 maize performed for regulatory purposes[32]. Overall, whether the increased levels of cadaverine and putrescine found in the NK603 maize samples can account for the signs of potential negative health effects upon its consumption by rats, as implied by the blood/urine biochemical analysis[33], needs to be further analyzed in experiments using more quantitative methods.

Our results suggest that expression of the EPSPS-CP4 transgene alters the oxidative environment in cells, and the increased levels of antioxidant enzymes are likely to be a response to oxidative burst by reactive oxygen species (ROS) in order to maintain proper physiological function. Glutathione metabolism was significantly altered in the NK603 when Roundup was sprayed during cultivation. Glutathione is known to be an important antioxidant in most living organisms, preventing damage to important cellular components caused by several environmental pollutants, including agrochemicals[59]. Plant glutathione S-transferases (GSTs) are also widely known for their role in herbicide detoxification[60]. Enzymes involved in combating reactive oxygen species, ascorbate peroxidase, glutathione reductase, and catalase are expressed at a higher level in transgenic soybean seeds[14]. Levels of ROS and other free radicals in GM food and feed would have to be monitored and quantified by further experiments in order to conclude on their potential impact on the agronomic performances of the plant. Additionally, it is known that polyamines are typically elevated in plants under abiotic stress conditions[61]. Typically, when cellular polyamine content increases, the levels of hydrogen peroxide also increases, activating antioxidant systems. Unintended effects of the inserted EPSPS-CP4 transgene was linked to energy metabolism disturbances in other studies[13–15]. It can be hypothesized that the plant is searching for a new equilibrium to maintain heterologous EPSPS-CP4 metabolism within levels that can be tolerated by the plant.

Glyphosate, the active ingredient of Roundup herbicide, inhibits the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), which is the sixth enzyme of the shikimate pathway, and plays an essential role in the biosynthesis of aromatic amino acids and other aromatic compounds in plants[62]. The EPSPS has a binding site for phosphenolpyruvate (PEP) and it could be hypothesized that an overexpression of a heterologous EPSPS could provoke a metabolic imbalance by altering the metabolism of PEP. Alterations in intermediate metabolism are corroborated in our experiment by the fact that the network formed by altered proteins/metabolites is centred on some TCA cycle intermediates (Fig. 5) such as α-ketoglutarate. In fact, it is also known that EPSPS inhibition by glyphosate impairs carbon metabolism, in particular by inducing alternative respiration and aerobic fermentation[63]. In this latest study, the metabolic switch was explained by an accumulation of pyruvate. Thus, if EPSPS inhibition is able to alter intermediate metabolism, a comparable change in the opposite direction could be expected as a result of EPSPS overexpression.

This study is the first and most detailed *multi-omics* characterization of a widely commercialized GMO crop and its isogenic counterpart. In conclusion, our integrative statistical and bioinformatics analysis allowed us to suggest a mechanistic link between the proteome and metabolome alterations observed and the insertion of a particular transgene. The transformation process and the resulting expression of a transgenic protein cause a general disturbance in the GM plant and it is clear that NK603 maize is markedly different from its non-GM isogenic line at the proteome and metabolome levels. In addition, our data correlates with previous studies, which observed higher amounts of ROS that act as free-radicals promoting oxidative stress in those transgenic plant materials. We also confirm a metabolic imbalance in energy and carbohydrate metabolism. Although a clear mechanistic link between alterations in the GM feed and the possible health effects following long-term consumption of this product remains to be established, the evidence we present clearly shows that NK603 and non-GM isogenic maize are not substantially equivalent and the nutritional quality of GM feed might be hampered by metabolic imbalances related to plant energy and stress metabolism.

Materials and Method

Maize cultivation. The varieties of maize used in this study were DKC 2678 Roundup-tolerant NK603 (Monsanto Corp., USA), and its nearest isogenic non-transgenic control DKC 2675. These two types of maize were grown under similar normal conditions, in the same location and season, spaced at a sufficient distance to avoid cross-contamination. The site of cultivation consisted of an imperfectly drained field with a coarse loam surface texture and fine loam subsoil. A typical soil compositional analysis is provided in Additional File 1. The maize cultivation rows were spaced 75 cm apart, with approximately 30 cm between planted seeds (78,000 seeds/ha).

One pass of the seeder included 4 rows of corn. To avoid edge effects in the field, 2 passes (8 rows) of DKC 2575 (isogenic) were planted as a buffer zone. DKC 2678 (NK603) and DKC 2575 (isogenic) were planted ~85 m apart. Half of the DKC 2678 received the treatment with Roundup WeatherMax.

Fertilization was performed with 26 T/ha liquid dairy manure, 100 kg/ha of 30-0-10 fertilizer was broadcast at planting, and 150 kg/ha of 18-46-0 fertilizer banded with the seed. The corn was harvested when the moisture content was less than 30%. All corn varieties were hand harvested by collecting ears in large tote bags to avoid cross contamination. The corn pickers were instructed to pick every ear of corn so as to avoid any risk of quality differentiation. Each corn variety was shelled (kernels removed from the cob) using a small threshing machine designed for this purpose. Each variety was dried in separate bulk drying bins to avoid any risk of cross contamination. The corn was dried at a low temperature (<30 °C) to avoid drying too rapidly and affecting feed quality. The corn was dried to <14% moisture before bagging.

The genetic nature, as well as the purity of the NK603 maize seeds and harvested material, was confirmed by quantitative PCR analysis of DNA samples. One field of NK603 was sprayed once with Roundup at 3 L ha^{-1} (WeatherMAX, 540 g/L of glyphosate, EPA Reg. 524–537) whilst another field of NK603 was not treated with Roundup. Test samples were produced by two cultivation cycles performed over two growing seasons. All maize samples were analysed for a total of 423 pesticide residues by SGS Institut Fresenius GmbH (Berlin, Germany), including glyphosate and its metabolite AMPA. No pesticide contaminants were detected in any of the samples (Additional file 2). All samples were maintained at −80 °C until processing for analysis. A schematic overview of our experimental design, sampling strategy and analytical approach is provided in Fig. 1.

Proteome analysis. *Sample preparation.* Ground maize kernel samples were lysed in 8M lysis buffer (urea, NaCl, Tri-HCl, phosphatase and protease inhibitor) and their protein concentration calculated using a Nanodrop protein assay. Samples in triplicate were run through an SDS-PAGE 4–20% polyacrylamide gradient gel at 150 V. Excised gel bands were reduced with dithiothreitol (Sigma-Aldrich Ltd, Gillingham, Dorset, UK), alkylated with Iodoacetamide (Sigma-Aldrich Ltd) and digested with bovine sequencing grade trypsin (Roche, Penzberg, Germany; ref. 11418475001) at 37 °C for 18 hours. Subsequently extracted peptides were labelled with 60 mM TMT10plex Isobaric Label Reagents (ThermoFisher Scientific, Waltham, MA, USA; ref 90406) and the respective samples combined. Labelled peptides were then purified and extracted using Waters Sep-Pak Vac 3cc 200 mg tC18 cartridges, before being separated into 10 fractions by strong cation exchange (SCX) across an increasing salt concentration. The eluted peptide fractions were purified and extracted once again before being lyophilised for direct analysis by liquid chromatography-tandem mass spectrometry (LC-MS/MS).

Liquid chromatography-tandem mass spectrometry. Fractionated samples were resuspended in 100 μl of 50 mM ammonium bicarbonate and 10 μl of each of the 10 fractions was loaded onto a 50 cm EASY-spray column (ThermoFisher Scientific). Quantitative analysis was performed using the Orbitrap Velos-Pro mass spectrometer (ThermoFisher Scientific) in positive ion mode. The peptides were separated by gradient elution, from 5–80% 0.1% trifluoroacetic acid in acetonitrile (5–40% from 0–100 minutes, 40–80% from 100–110 minutes), at a flow rate of 300 nl/min. Mass spectra (m/z) ranging from 400–1600 Daltons was acquired at a resolution of 60,000 and the 10 most intense ions were subjected to MS/MS by HCD fragmentation with 35% collision energy.

Data processing. Protein identification was performed with Proteome Discoverer 1.4. Raw files were imported and searched against the UniProtKB/Swiss-Prot Database using Sequest for Proteome Discoverer. Raw files for all fractions were merged together in a single file search for each of the two TMT10plex sets. Precursor mass tolerance for the searches was set at 20ppm and fragment mass tolerance at 0.8ppm. The taxonomy selected was *Zea mays* and three enzymatic mis-cleavages were allowed. Dynamic modifications selected on the search were Oxidation/+15.995 Da (M) and Deamidated/+0.984 (N, Q) and static modifications were Carbamidomethyl/+57.021 Da (C), TMT10plex/229.163 Da (K), TMT10plex/229.163 Da (Any N-terminus). Only peptides with TMT reporter ion signal intensities for all ten samples were used for further bioinformatics analysis.

Metabolome analysis. The metabolome analysis was performed by Metabolon Inc. (Durham, NC, USA) as previously described[64]. Ground maize kernel samples were prepared using the automated MicroLab STAR® system from Hamilton Company (Reno, NV, USA). Several recovery standards were added prior to the first step in the extraction process for QC purposes. In order to remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI, and one sample was reserved for backup. Samples were placed briefly on a TurboVap® (SOTAX Corp, Westborough, MA, USA) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis.

Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS) for metabolome analysis. All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (Waters Corp, Milford, MA, USA) and a ThermoFisher Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyser operated at 35,000 mass resolution. The sample extract was dried then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic

consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18-2.1 × 100 mm, 1.7 μm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analysed using acidic positive ion conditions, however it was chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same afore mentioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analysed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5 mM ammonium bicarbonate at pH 8. The fourth aliquot was analysed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1 × 150 mm, 1.7 μm) using a gradient consisting of water and acetonitrile with 10 mM ammonium formate, pH 10.8. The MS analysis alternated between MS and data-dependent MSn scans using dynamic exclusion. The scan range varied slighted between methods but covered 70–1000 m/z.

Metabolome data processing. A quality control value assessment was undertaken to determine instrument variability by calculating the median relative standard deviation (RSD) for the internal standards that were pre-mixed into each sample prior to injection into the mass spectrometer. This yielded a value of 3% for instrument variability. Overall process variability as determined by calculating the median RSD for all endogenous metabolites (that is, non-instrument standards) present in 100% of the samples gave a value of 7%. Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software as previously described[65]. Metabolites were identified by automated comparison and curated by visual inspection for quality control using software developed at Metabolon[66]. Peaks were quantified using area-under-the-curve.

Integrative bioinformatics analysis. For plotting of results, a Principal Component Analysis (PCA) was first performed. The language and statistical environment R[67] together with the ade4 package[68] method was employed in order to explore the relationship between GM and non-GM varieties. Second, we performed a Multiple Co-Inertia Analysis (MCIA), using the language and statistical environment R together with the omicade4 package[69], in order to integrate multiple *omics* datasets where the same tissue have been assayed multiple times (in this case, proteomics and metabolomics).

Pairwise Welch's t-tests were performed, for bothproteomics and metabolomics datasets, for Isogenic *vs* NK603, Isogenic *vs* NK603+Roundup and NK603 *vs* NK603+Roundup comparisons. The resulting p-values were adjusted by the Benjamini-Hochberg multi-test adjustment method for the high number of comparisons. Volcano plots were also constructed in order to visualize the differences in metabolite and protein expression for each of the comparisons. The aforementioned tests and plots were performed using in-house R scripts. Pathway enrichment analysis of the proteomics dataset was conducted using the web tool STRING v10.0[70]. For the metabolomics data, due to a lack of well-annotated metabolome databases for maize, the pathway enrichment analysis was conducted as follows. First, enrichment scores (ES) for each pathways were determined using the following formula: ES = (k/m)/(n/N) where (# of significant metabolites in pathway(k)/total # of detected metabolites in pathway(m))/(total # of significant metabolites(n)/total # of detected metabolites(N)). Then, the statistical significance was assessed using a Fisher one-sided exact test. The STITCH v5.0 beta web tool[71] was used to investigate metabolite-protein interactions on maize endogenous pathways. The list of disturbed proteins and metabolites, including the protein EPSPS, was uploaded and the metabolic networks was studied using STITCH v5.0 initial parameters.

References

1. Van Montagu, M. It is a long way to GM agriculture. *Annu Rev Plant Biol* **62,** 1–23 (2011).
2. Parisi, C., Tillie, P. & Rodriguez-Cerezo, E. The global pipeline of GM crops out to 2020. *Nat Biotechnol* **34,** 31–36 (2016).
3. James, C. Global Status of Commercialized Biotech/GM Crops: 2015. *ISAAA Brief* **51** (2015).
4. AHTEG. Guidance Document on Risk Assessment of Living Modified Organisms. United Nations Environment Programme Convention for Biodiversity. http://www.cbd.int/doc/meetings/bs/mop-05/official/mop-05-12-en.pdf (Date of access:24/10/2016) (2010).
5. Benbrook, C. Enhancements Needed in GE Crop and Food Regulation in the U.S. *Frontiers in public health* **4,** 59 (2016).
6. Millstone, E., Brunner, E. & Mayer, S. Beyond substantial equivalence. *Nature* **401,** 525–526 (1999).
7. Cuhra, M. Review of GMO safety assessment studies: glyphosate residues in Roundup Ready crops is an ignored issue. *Environmental Sciences Europe* **27,** 1–14 (2015).
8. Heinemann, J. A., Kurenbach, B. & Quist, D. Molecular profiling–a tool for addressing emerging gaps in the comparative risk assessment of GMOs. *Environ Int* **37,** 1285–1293 (2011).
9. Agapito-Tenfen, S. Z., Guerra, M. P., Wikmark, O. G. & Nodari, R. O. Comparative proteomic analysis of genetically modified maize grown under different agroecosystems conditions in Brazil. *Proteome science* **11,** 46 (2013).
10. Agapito-Tenfen, S. Z. *et al.* Effect of stacking insecticidal cry and herbicide tolerance epsps transgenes on transgenic maize proteome. *BMC Plant Biology* **14,** 1–19 (2014).
11. Barros, E. *et al.* Comparison of two GM maize varieties with a near-isogenic non-GM variety using transcriptomics, proteomics and metabolomics. *Plant biotechnology journal* **8,** 436–451 (2010).
12. Zolla, L., Rinalducci, S., Antonioli, P. & Righetti, P. Proteomics as a complementary tool for identifying unintended side effects occurring in transgenic maize seeds as a result of genetic modifications. *J Proteome Res* **7,** 1850–1861 (2008).
13. Brandao, A. R., Barbosa, H. S. & Arruda, M. A. Image analysis of two-dimensional gel electrophoresis for comparative proteomics of transgenic and non-transgenic soybean seeds. *J Proteomics* **73,** 1433–1440 (2010).
14. Barbosa, H. S., Arruda, S. C., Azevedo, R. A. & Arruda, M. A. New insights on proteomics of transgenic soybean seeds: evaluation of differential expressions of enzymes and proteins. *Anal Bioanal Chem* **402,** 299–314 (2012).
15. Arruda, S. C., Barbosa, H. S., Azevedo, R. A. & Arruda, M. A. Comparative studies focusing on transgenic through cp4EPSPS gene and non-transgenic soybean plants: an analysis of protein species and enzymes. *J Proteomics* **93,** 107–116 (2013).

16. Lehesranta, S. J. *et al.* Comparison of tuber proteomes of potato varieties, landraces, and genetically modified lines. *Plant Physiol* **138,** 1690–1699 (2005).
17. Wang, L. *et al.* Comparative proteomics of Bt-transgenic and non-transgenic cotton leaves. *Proteome science* **13,** 15 (2015).
18. Gong, C. Y., Li, Q., Yu, H. T., Wang, Z. & Wang, T. Proteomics insight into the biological safety of transgenic modification of rice as compared with conventional genetic breeding and spontaneous genotypic variation. *J Proteome Res* **11,** 3019–3029 (2012).
19. Ricroch, A. E., Bergé, J. B. & Kuntz, M. Evaluation of Genetically Engineered Crops Using Transcriptomic, Proteomic, and Metabolomic Profiling Techniques. *Plant Physiol* **155,** 1752–1761 (2011).
20. Manetti, C. *et al.* A metabonomic study of transgenic maize (Zea mays) seeds revealed variations in osmolytes and branched amino acids. *J Exp Bot* **57,** 2613–2625 (2006).
21. EFSA. Guidance on selection of comparators for the risk assessment of genetically modified plants and derived food and feed. *EFSA J* **2149,** doi: 10.2903/j.efsa.2011.2150 (2011).
22. Losey, J. E., Rayor, L. S. & Carter, M. E. Transgenic pollen harms monarch larvae. *Nature* **399,** 214, doi: 10.1038/20338 (1999).
23. Hilbeck, A. & Schmidt, J. E. U. Another view on Bt proteins – how specific are they and what else might they do? *Biopesti Int* **2,** 1–50 (2006).
24. Hilbeck, A., Meier, M. & Trtikova, M., Underlying reasons of the controversy over adverse effects of Bt toxins on lady beetle and lacewing larvae. *Environmental Sciences Europe* **24,** doi: 10.1186/2190-4715-24-9 (2012).
25. Hammond, B., Dudek, R., Lemen, J. & Nemeth, M. Results of a 13 week safety assurance study with rats fed grain from glyphosate tolerant corn. *Food Chem Toxicol* **42,** 1003–1014 (2004).
26. Hammond, B. *et al.* Results of a 90-day safety assurance study with rats fed grain from corn rootworm-protected corn. *Food Chem Toxicol* **44,** 147–160 (2006).
27. Spiroux de Vendomois, J. *et al.* Debate on GMOs health risks after statistical findings in regulatory tests. *Int J Biol Sci* **6,** 590–598 (2010).
28. Seralini, G.-E. *et al.* Genetically modified crops safety assessments: present limits and possible improvements. *Environ Sci Eur* **23,** 10 (2011).
29. Doull, J. *et al.* Report of an Expert Panel on the reanalysis by of a 90-day study conducted by Monsanto in support of the safety of a genetically modified corn variety (MON 863). *Food Chem Toxicol* **45,** 2073–2085 (2007).
30. Mesnage, R. & Séralini, G.-É. *The Need for a Closer Look at Pesticide Toxicity during GMO Assessment, in Practical Food Safety: Contemporary Issues and Future Directions* (eds Bhat, R. & Gómez-López, V. M.) (John Wiley & Sons, Ltd, Chichester, UK, doi: 10.1002/9781118474563.ch10. 2014).
31. Mesnage, R. *et al.* Transcriptome profile analysis reflects rat liver and kidney damage following chronic ultra-low dose Roundup exposure. *Environ Health* **14,** 70 (2015).
32. Monsanto. Safety Assessment of Roundup Ready Corn Event NK603 http://www.monsanto.com/products/documents/safety-summaries/corn_pss_nk603.pdf (Date of access: 24/10/2016) (2002).
33. Ocana, M. F., Fraser, P. D., Patel, R. K., Halket, J. M. & Bramley, P. M. Mass spectrometric detection of CP4 EPSPS in genetically modified soya and maize. *Rapid Commun Mass Spectrom* **21,** 319–328 (2007).
34. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43,** D222–226 (2015).
35. National Academies of Sciences, E., and Medicine. *Genetically Engineered Crops: Experiences and Prospects* (Washington, DC: The National Academies Press doi: 10.17226/23395 2016).
36. OECD. Safety Considerations for Biotechnology: Scale-up of Crop Plants. Paris: OECD. https://www.oecd.org/env/ehs/biotrack/1958527.pdf (Date of access: 24/10/2016) (1993).
37. Codex Alimentarius Commission Guideline for the Conduct of Food Safety Assessment of Foods Using Recombinant DNA Plants. Doc CAC/GL 45-2003. Rome: World Health Organization and Food and Agriculture Organization. http://www.fao.org/input/download/standards/10021/CXG_045e.pdf (Date of access: 24/10/2016) (2003).
38. AHTEG Final Report of the Ad Hoc Technical Expert Group on Risk Assessment and Risk Management under the Cartagena Protocol on Biosafety. UNEP/CBD/BS/AHTEG-RA&RM/4/6. http://www.cbd.int/doc/meetings/bs/bsrarm-04/official/bsrarm-04-06-en.pdf (Date of access: 24/10/2016) (2012).
39. Latham, J. R., Wilson, A. K. & Steinbrecher, R. A. The mutational consequences of plant transformation. *Journal of biomedicine & biotechnology* **2006,** 25376 (2006).
40. Fonseca, C. *et al. In vitro* culture may be the major contributing factor for transgenic versus nontransgenic proteomic plant differences. *Proteomics* **15,** 124–134 (2015).
41. Garcia-Canas, V., Simo, C., Leon, C., Ibanez, E. & Cifuentes, A. MS-based analytical methodologies to characterize genetically modified crops. *Mass spectrometry reviews* **30,** 396–416 (2011).
42. Simó, C., Ibáez, C., Valdés, A., Cifuentes, A. & García-Cañas, V. Metabolomics of Genetically Modified Crops. *International Journal of Molecular Sciences* **15,** 18941–18966 (2014).
43. Heinemann, J. A. & El-Kawy, O. A. Observational science in the environmental risk assessment and management of GMOs. *Environ Int* **45,** 68–71 (2012).
44. D'Alessandro, A. & Zolla, L. We are what we eat: food safety and proteomics. *J Proteome Res* **11,** 26–36 (2012).
45. Trtikova, M., Wikmark, O. G., Zemp, N., Widmer, A. & Hilbeck, A.Transgene Expression and Bt Protein Content in Transgenic Bt Maize (MON810) under Optimal and Stressful Environmental Conditions. *PLoS ONE* **10,** e0123011 (2015).
46. Vidal, N., Barbosa, H., Jacob, S. & Arruda, M. Comparative study of transgenic and non-transgenic maize (Zea mays) flours commercialized in Brazil, focussing on proteomic analyses. *Food chemistry* **180,** 288–294 (2015).
47. Frank, T., Rohlig, R. M., Davies, H. V., Barros, E. & Engel, K. H. Metabolite profiling of maize kernels–genetic modification versus environmental influence. *J Agric Food Chem* **60,** 3005–3012 (2012).
48. Coll, A. *et al.* Natural variation explains most transcriptomic changes among maize plants of MON810 and comparable non-GM varieties subjected to two N-fertilization farming practices. *Plant Mol Biol* **73,** 349–362 (2010).
49. Li, Z., Wang, Z. & Li, S. Gene chip analysis of Arabidopsis thaliana genomic DNA methylation and gene expression in response to carbendazim. *Biotechnology letters* **37,** 1297–1307 (2015).
50. Hauser, M.-T., Aufsatz, W., Jonak, C. & Luschnig, C. Transgenerational epigenetic inheritance in plants. *Biochimica et Biophysica Acta* (*BBA*) - *Gene Regulatory Mechanisms* **1809,** 459–468 (2011).
51. Seralini, G.-E. *et al.* Republished study: long-term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. *Environmental Sciences Europe* **26,** 14 (2014).
52. Esco Working Group, M. EFSA Compendium of botanicals that have been reported to contain toxic, addictive, psychotropic or other substances of concern. *EFSA Supporting Publications* **6,** doi: 10.2903/j.efsa.2012.2663 (2009).
53. Tomar, P. C., Lakra, N. & Mishra, S. N. Cadaverine: a lysine catabolite involved in plant growth and development. *Plant signaling & behavior* **8,** doi: 10 4161/psb 25850 (2013).
54. Simon-Sarkadi, L., Ludidi, N. & Kocsy, G. Modification of cadaverine content by NO in salt-stressed maize. *Plant signaling & behavior* **9,** e27598 (2014).
55. Minois, N., Carmona-Gutierrez, D. & Madeo, F. Polyamines in aging and disease. *Aging* (*Albany NY*) **3,** 716–732 (2011).
56. Álvarez González, M. Á., Calles-Enríquez, M., Fernández García, M. & Ladero Losada, V. M. Toxicological Effects of Dietary Biogenic Amines. Curr Nutr *Food Sci* **6,** 145–156 (2010).

57. Nebelin, E., Pillai, S., Lund, E. & Thomsen, J. On the formation of N-nitrosopyrrolidine from potential precursors and nitrite. *IARC scientific publications* 183–193 (1980).
58. Soda, K., Dobashi, Y., Kano, Y., Tsujinaka, S. & Konishi, F. Polyamine-rich food decreases age-associated pathology and mortality in aged mice. *Exp Gerontol* **44,** 727–732 (2009).
59. Couto, N., Wood, J. & Barber, J. The role of glutathione reductase and related enzymes on cellular redox homoeostasis network. *Free Radical Biology and Medicine* **95,** 27–42 (2016).
60. Milligan, A. S., Daly, A., Parry, M. A. J., Lazzeri, P. A. & Jepson, I. The expression of a maize glutathione S-transferase gene in transgenic wheat confers herbicide tolerance, both in planta and *in vitro*. *Molecular Breeding* **7,** 301–315 (2001).
61. Minocha, R., Majumdar, R. & Minocha, S. C. Polyamines and abiotic stress in plants: A complex relationship. *Frontiers in Plant Science* **5** (2014).
62. Boocock, M. R. & Coggins, J. R. Kinetics of 5-enolpyruvylshikimate-3-phosphate synthase inhibition by glyphosate. *FEBS Letters* **154,** 127–133 (1983).
63. Armendariz, O., Gil-Monreal, M., Zulet, A., Zabalza, A. & Royuela, M. Both foliar and residual applications of herbicides that inhibit amino acid biosynthesis induce alternative respiration and aerobic fermentation in pea roots. *Plant Biol* (*Stuttg*) **18,** 382–390 (2016).
64. Evans, A. M. *et al.* High Resolution Mass Spectrometry Improves Data Quantity and Quality as Compared to Unit Mass Resolution Mass Spectrometry in High-Throughput Profiling Metabolomics. *Metabolomics* **4,** 132 (2014).
65. DeHaven, C. D., E., A., Dai, H. & Lawton, K. A. Software Techniques for Enabling High-Throughput Analysis of Metabolomic Datasets. *Metabolomics, Dr Ute Roessner (Ed.), ISBN: 978-953-51-0046-1, InTech*, doi: 10.5772/31277 (2012).
66. Dehaven, C. D., Evans, A. M., Dai, H. & Lawton, K. A. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *Journal of cheminformatics* **2,** 9 (2010).
67. Team, R. C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL http://www.R-project.org/ (Date of access: 24/10/2016) (2015).
68. Dray, S. & Dufour, A.-B. *The ade4 Package: Implementing the Duality Diagram for Ecologists. 2007* **22,** 20 (2007).
69. Meng, C., Kuster, B., Culhane, A. C. & Gholami, A. M. A multivariate approach to the integration of multi-omics datasets. *BMC bioinformatics* **15,** 1–13 (2014).
70. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43,** D447–452 (2015).
71. Szklarczyk, D. *et al.* STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res* **44,** D380–384 (2016).

Acknowledgements

We thank Pr. Nigel Crawford (USCD) for careful reading of the manuscript and helpful comments and suggestions. This work was funded by the Sustainable Food Alliance (USA), whose support is gratefully acknowledged.

Author Contributions

R.M. and S.Z.A. interpreted the data, and drafted the manuscript. V.V. performed the statistical analysis. G.R. and M.W. conducted the proteome experiment. R.O.N. assisted with data interpretation. G.E.S. conceived the animal feeding trial and provided maize samples for analysis. M.N.A. and G.E.S. conceived the study. M.N.A. coordinated the investigation and drafted the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Mesnage, R. *et al.* An integrated multi-omics analysis of the NK603 Roundup-tolerant GM maize reveals metabolism disturbances caused by the transformation process. *Sci. Rep.* **6**, 37855; doi: 10.1038/srep37855 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.