



Commentary

Evidence-based regulations for bioinformatic prediction of allergen cross-reactivity are needed

Rod A. Herman^{a,*}, Ping Song^a, Henry P. Mirsky^b, Jason M. Roper^c^a Corteva Agriscience, 9330 Zionsville Road, Indianapolis, IN, 47968, USA^b Corteva Agriscience, 8325 NW 62nd Avenue, Johnston, IA, 50131, USA^c Corteva Agriscience, P.O. Box 30 Newark DE, 19714, USA

ARTICLE INFO

Keywords

Bioinformatics
Allergens
Cross-reactive
Genetically engineered crops
Regulation
Expert opinion
Empirical evidence

ABSTRACT

The bioinformatic criteria adopted by regulatory agencies to predict the potential cross reactivity between newly expressed proteins in genetically engineered crops and known allergens involves amino acid identity thresholds and was formulated nearly two decades ago based on the opinion of allergy experts. Over the subsequent years, empirical evidence has been developed indicating that better bioinformatic tools based on amino acid similarity are available to detect real allergen cross-reactive risk while substantially reducing false-positive detections. Although the formulation of safety regulations, in the absence of empirical evidence, may require reliance on expert opinion, such expert opinion should not trump empirical evidence once it becomes available. The failure of regulation to maintain consistency with the best available scientific evidence diminishes its value and creates arbitrary barriers to the use of beneficial technologies by society.

Amino acid identity allergen cross-reactivity thresholds: The bioinformatic criteria adopted by regulatory agencies to predict the potential cross reactivity between newly expressed proteins in genetically engineered crops and known allergens was developed nearly two decades ago based on the opinion of allergy experts (CODEX, 2009; Codex Alimentarius Commission, 2003; FAO/WHO, 2001). The presence of >35% amino acid identity (exact match) between a newly expressed protein and known allergens over an alignment of at least 80 amino acids was set as a threshold to account for the most disparate cross-reactive allergens, and the shared presence of a short identical stretch of amino acids (typically 8 or more) was chosen as an additional threshold based on the size of small IgE epitopes (Kleter and Peijnenburg, 2004; Taylor, 1997). Thus far, only the birch allergen Bet v 1 is known to cross-react with other allergens (e.g., Dac c 1 from carrot and Vig r 6 from mung bean) that share <35% identity across their entire sequences (although they do share >35% identity across alignments of >80 amino acids) (Guhs et al., 2014; Song et al., 2014). In fact, cross-reactivity is rare below 50% identity which may typically be required to produce the similar folding patterns that enable IgE cross reactivity (Aalberse, 2000).

Amino acid similarity is superior: While the threshold of >35% identity over at least an 80-amino-acid local alignment (e.g., as

generated by the FASTA computer program) has been shown to be sensitive to the detection of cross-reactive allergens, it has also been found to have poor specificity (high false-positive rates), and the use of sequence similarity search algorithms (weighting substitutions based on the similarity of physicochemical properties of the amino acids within local alignments with appropriate statistical thresholds such as E values that estimate the probability of a random alignment) have been found to be equally sensitive at detection of true allergen cross reactivity, but with substantially more selectivity against false detection of non-cross-reactive proteins (considerably lower false-positive rates) (Cressman and Ladics, 2009; Herman and Song, 2019, 2020; Ladics, 2019; Ladics et al., 2007; Silvanovich et al., 2009; Song et al., 2014, 2015). This is not surprising because the latter approach accounts for the physiochemical similarity or dissimilarity of amino acid substitutions and thus offers greater granularity in detecting the physiochemical relatedness or dissimilarity between proteins that govern IgE antibody cross reactivity (Ladics et al., 2011). For the aforementioned example of cross-reactive allergens, Bet v 1 vs. carrot Dac c 1 and mung-bean Vig r 6 with low overall identity, similarity searches result in highly significant E-values (8.7E-24 and 1.9E-12, respectively, using a 1:1 FASTA approach) demonstrating the sensitivity of similarity search algorithms even in the absence of high identity (Song et al., 2014). It is noteworthy that

* Corresponding author.

E-mail address: rod.herman@corteva.com (R.A. Herman).

multiple different applications of similarity searches have been investigated to detect allergen cross reactivity (each with different suggested E-value thresholds), and while all have shown equal sensitivity and improved specificity for detection of true cross reactivity compared with the existing >35% identity criteria, a comprehensive comparison among these proposed similarity algorithms and E-value criteria has not been completed (Herman and Song, 2019). It is not the purpose of the current discussion to review and compare these different similarity algorithms, but rather to note that each is superior to the current identity approach that is used by regulatory agencies, and thus, that this general approach deserves consideration.

Short identical amino acid sequences: Short identical amino acid match thresholds (e.g. 8-mer matches) have also been shown to generate many false detections of potential cross reactivity (between innocuous proteins and allergens) (Ladics et al., 2006; Silvanovich et al., 2006). In addition, the presence of such matches, in the absence of identity or similarity across much larger amino acid stretches, has not been found useful in identifying true cross-reactive allergens (Goodman et al., 2005; Herman et al., 2009; Ladics, 2019). We are aware of no case where a shared 8-mer identity in the absence of >35% identity over at least 80 amino acids is found in even a single pair of cross-reactive allergens. This is the likely reason that such short-mer amino acid matches are no longer incorporated in the European Food Safety Authority (EFSA) allergenicity guidance for newly expressed proteins in genetically engineered crops (EFSA, 2010). Thus, this criterion, in isolation from greater overall protein amino acid similarity or identity, only results in false detections and is not useful as a threshold for detecting cross-reactive risk between newly expressed proteins in genetically engineered crops and known allergens (Herman et al., 2009). Although an IgE epitope may be a short sequence, such sequences out of the context of the greater allergen structure appears insufficient to enable IgE cross reactivity (Klinglmayr et al., 2009; Ladics et al., 2011).

Expert opinion vs. empirical evidence: The formulation of safety regulations in the absence of empirical evidence necessarily relies solely on the best available expert opinion. However, safety regulations should be modified as empirical evidence is accumulated, with prior expert opinion being de-emphasized, especially expert opinion that fails to adapt to account for new empirical evidence (Herman and Raybould, 2014; Sackett and Rosenberg, 1995). The more definitive the empirical evidence, the less expert opinion should be relied upon. Such empirical evidence is now available to update the bioinformatic procedures used to identify potential cross reactivity between allergens, and this evidence should be explored and provide the basis for modifying the regulations that govern the safety assessment of newly expressed proteins in genetically engineered crops. The failure of regulation to maintain consistency with the best available scientific evidence diminishes its value and creates arbitrary barriers to the use of beneficial technologies by society.

Funding statement

This manuscript received no funding from external sources outside Corteva Agriscience.

Declaration of competing interest

The authors declare the following financial interests/personal

relationships which may be considered as potential competing interests: The authors are employed by Corteva Agriscience which develops and markets genetically engineered seed. No specific funding was provided to prepare this manuscript.

Acknowledgments

We thank Nicolas Storer of Corteva Agriscience for helpful edits to a draft of this paper.

References

- Aalberse, R.C., 2000. Structural biology of allergens. *J. Allergy Clin. Immunol.* 106, 228–238.
- CODEX, 2009. *Foods Derived from Biotechnology*, second ed. World Health Organization, Food and Agriculture Organization of the United Nations, Rome, pp. 7–34.
- Codex Alimentarius Commission, 2003. *Guideline for the Conduct of Food Safety Assessment of Foods Derived from Recombinant-DNA Plants*.
- Cressman, R.F., Ladics, G., 2009. Further evaluation of the utility of “Sliding Window” FASTA in predicting cross-reactivity with allergenic proteins. *Regul. Toxicol. Pharmacol.* 54, S20–S25.
- EFSA, 2010. Scientific opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. *EFSA Journal* 8, 1–168.
- FAO/WHO, 2001. Evaluation of allergenicity of genetically modified foods. report of Joint FAO/WHO Expert Consultation. Food and Agriculture Organization of the United Nations, Rome.
- Goodman, R.E., et al., 2005. Assessing genetically modified crops to minimize the risk of increased food allergy: a review. *Int. Arch. Allergy Immunol.* 137, 153–166.
- Guhs, E.E., et al., 2014. Vig r 6, the cytokinin-specific binding protein from mung bean (*Vigna radiata*) sprouts, cross-reacts with Bet v 1-related allergens and binds IgE from birch pollen allergic patients’ sera. *Mol. Nutr. Food Res.* 58, 625–634.
- Herman, R., et al., 2009. Value of eight-amino-acid matches in predicting the allergenicity status of proteins: an empirical bioinformatic investigation. *Clin. Mol. Allergy* 7, 1–7.
- Herman, R.A., Raybould, A., 2014. Expert opinion vs. empirical evidence: the precautionary principle applied to GM crops. *GM Crops Food* 5, 8–10.
- Herman, R.A., Song, P., 2019. Validation of bioinformatic approaches for predicting allergen cross reactivity. *Food Chem. Toxicol.* 132, 110656.
- Herman, R.A., Song, P., 2020. Allergen false-detection using official bioinformatic algorithms. *GM Crops Food* 11, 93–96.
- Kleter, G.A., Peijnenburg, A.A., 2004. Prediction of the potential allergenicity of novel proteins. *Allergy Matters: New Approaches to Allergy Prevention and Management: the International Conference on Allergy Prevention*. The Netherlands, Wageningen, pp. 85–93, 4–6 February 2004.
- Klinglmayr, E., et al., 2009. Identification of B-cell epitopes of Bet v 1 involved in cross-reactivity with food allergens. *Allergy* 64, 647–651.
- Ladics, G.S., 2019. Assessment of the potential allergenicity of genetically-engineered food crops. *J. Immunot.* 16, 43–53.
- Ladics, G.S., et al., 2007. Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens. *Mol. Nutr. Food Res.* 51, 985–998.
- Ladics, G.S., et al., 2006. Lack of cross-reactivity between the *Bacillus thuringiensis* derived protein Cry1F in maize grain and dust mite Der p7 protein with human sera positive for Der p7-IgE. *Regul. Toxicol. Pharmacol.* 44, 136–143.
- Ladics, G.S., et al., 2011. Bioinformatics and the allergy assessment of agricultural biotechnology products: industry practices and recommendations. *Regul. Toxicol. Pharmacol.* 60, 46–53.
- Sackett, D.L., Rosenberg, W.M., 1995. On the need for evidence-based medicine. *J. Publ. Health* 17, 330–334.
- Silvanovich, A., et al., 2009. The use of E-scores to determine the quality of protein alignments. *Regul. Toxicol. Pharmacol.* 54, S26–S31.
- Silvanovich, A., et al., 2006. The value of short amino acid sequence matches for prediction of protein allergenicity. *Toxicol. Sci.* 90, 252–258.
- Song, P., et al., 2015. 1:1 FASTA update: using the power of E-values in FASTA to detect potential allergen cross-reactivity. *Toxicology Reports* 2, 1145–1148.
- Song, P., et al., 2014. Evaluation of global sequence comparison and one-to-one FASTA local alignment in regulatory allergenicity assessment of transgenic proteins in food crops. *Food Chem. Toxicol.* 71, 142–148.
- Taylor, S.L., 1997. Food from genetically modified organisms and potential for food allergy. *Environ. Toxicol. Pharmacol.* 4, 121–126.