prominent examples that provide incentives for companies to adopt sustainable practices instead of forest conversion. In this setting, the use of smart technologies and blockchain allows more transparent and effective monitoring of the entire supply chain [13]. Biotechnology can be a useful complement to enhance the traceability of food and bio-based materials, such as the use of biomarkers [14]. A synergistic combination of technologies can reduce transaction costs, allowing more uptake of various certifications by smaller players in developing regions, who are often blamed for their inefficient and unsustainable land-use practices. Interestingly, the advancement of connectivity between suppliers and potential buyers through digital marketing also permits marketing new, novel biomaterials that are difficult to access due to logistical constraints. A noteworthy example is the digital documentation of indigenous medicine and perfumes through screening of bioactive compounds extracted from tropical forests in Borneo [15]. Digital market platforms can further contribute to the creation of new markets for such products, creating new income sources from conserving tropical rainforests.

These five strategies are important in addressing the critical question of how to support growth, not only without causing further environmental impacts, but also repairing damage done in the past. The digital revolution, coupled with biotechnology, opens up more possibilities to reconcile economic development and conservation. The *Six Transformations* framework marks the importance of crafting and adopting technologies for improving people's lives, prosperity, and wellbeing in the context of sustainable development. Nevertheless, it is crucial to recognize that technological innovation in the land-use system requires much deeper thoughts on implementation and business models that fit well in specific local contexts. While interdisciplinary collaboration between different scientific communities is imperative, working closely with stakeholders on the ground to effectively design, execute, and manage the strategies is key for realizing the *Six Transformations*. The examples shown in Box 1 may shed some light on such on-ground efforts. This paper offers initial implications of synergies between biotechnologies, land-use systems, and the digital revolution, which may hold important prospects both in academia and among practitioners in driving sustainable development.

[1]Harvard University Asia Center, Cambridge, MA, USA
[2]Jeffrey Sachs Center on Sustainable Development, Sunway University, Bandar Sunway, Selangor Darul Ehsan, Malaysia
[3]Department of Innovation Science, School of Environment and Society, Tokyo Institute of Technology, Tokyo, Japan
[4]Department of Economics, University of California, Davis, CA, USA

*Correspondence:
chunshenggoh@fas.harvard.edu (C.S. Goh).

https://doi.org/10.1016/j.tibtech.2020.05.010

References

1. Sachs, J.D. *et al.* (2019) Six Transformations to achieve the sustainable development goals. *Nat. Sustain.* 2, 805–814
2. Taranto, F. *et al.* (2018) Biotechnological and digital revolution for climate-smart plant breeding. *Agronomy* 8, 277
3. Sandino, J. *et al.* (2018) Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. *Sensors* 18, 944
4. Jiménez, D. *et al.* (2019) A scalable scheme to implement data-driven agriculture for small-scale farmers. *Glob. Food Sec.* 23, 256–266
5. Grobelak, A. *et al.* (2018) Plant growth-promoting rhizobacteria as an alternative to mineral fertilizers in assisted bioremediation –sustainable land and waste management. *J.Environ. Manag.* 227, 1–9
6. Salam, A. *et al.* (2019) Di-Sense: in situ real-time permittivity estimation and soil moisture sensing using wireless underground communications. *Comput. Netw.* 151, 31–41
7. Sayer, J.A. *et al.* (2016) Measuring the effectiveness of landscape approaches to conservation and development. *Sustain. Sci.* 12, 465–476
8. Heredia, I. (2017) Large-scale plant classification with deep neural networks. In *Proceedings of the Computing Frontiers Conference*, ACM
9. Liu, J. *et al.* (2019) Positioning of coordinates and precision analysis of sample trees using the intelligent forest survey calculator. *Comput. Electron. Agric.* 159, 157–164
10. Boles, K.S. *et al.* (2017) Digital-to-biological converter for on-demand production of biologics. *Nat. Biotechnol.* 35, 672–675
11. Ahl, A. *et al.* (2019) Smart technology applications in the woody biomass supply chain: interview insights and potential in Japan. *Sustain. Sci.* Published online September 4, 2019. https://doi.org/10.1007/s11625-019-00728-2
12. Scheitz, C.J.F. *et al.* (2018) Biotechnology software in the digital age: are you winning? *J.Ind.Microbiol.Biotechnol.* 45, 529–534
13. Hirbli, T. (2018) *Palm Oil Traceability: Blockchain Meets Supply Chain*, Massachusetts Institute of Technology
14. Vlam, M. *et al.* (2018) Developing forensic tools for an African timber: regional origin is revealed by genetic characteristics, but not by isotopic signature. *Biol. Conserv.* 220, 262–271
15. Yeo, T.C. *et al.* (2014) Building a discovery partnership with Sarawak biodiversity centre: a gateway to access natural products from the rainforests. *Comb. Chem.High Throughput Screen.* 17, 192–200

Science & Society

Allergenicity Assessment of Novel Food Proteins: What Should Be Improved?

Antonio Fernandez,[1,*] E.N. Clare Mills,[2] Frits Koning,[3] and F. Javier Moreno[4]

Allergenicity prediction is one of the most challenging aspects in the safety assessment of foods derived from either biotechnology or novel food proteins. Here we present a bottom-up strategy that defines *a priori* the specific risk assessment (RA) needs based on a database appropriately built for such purposes.

Background, Challenges, and Future Needs

Obtaining alternative proteins from genetic engineering approaches and novel food sources is a priority for Food 2030[i], the research and innovation policy framework from the European Commission to future-proof our food system. This is because there is an urgent need for plants with improved resistance to abiotic stress and pathogens and for dietary proteins from novel sources that will support reductions

in greenhouse gas emissions from a more sustainable food system. One of the most difficult aspects of assessing the safety of foods derived from biotechnology or novel food proteins is allergenicity RA. Central to the current weight-of-evidence approach applied to assess potential allergenicity is the use of bioinformatics to compare the sequence of a novel protein with those of allergen proteins, which cause allergic reactions [1–3][ii]. Traditionally, the sequence of a novel protein is compared with those of known allergens using a local alignment algorithm such as FASTA and a threshold value against amino acid sequence alignments (i.e., 35% sequence similarity over a sliding window of 80 amino acids) is used to indicate potential allergenic risks requiring further assessment[ii]. The scientific community has also used this approach to compare allergens and non-allergens and to develop more advanced *in silico* tools (e.g., using machine learning) for improved allergenicity prediction for novel food proteins [4–6]. However, poor understanding of the specific characteristics of a protein that confer potential allergenicity limits the usefulness of bioinformatics for RA. Here we present a bottom-up strategy, which, contrary to the current paradigm, defines *a priori* the specific RA needs for the investigation of any given novel protein's cross-reactive allergenic potential.

Bottom-Up Food Safety Evaluation in Allergenicity Assessment

A bottom-up strategy would place greater emphasis on the development of allergen sequence databases, where curation allows additional criteria to be applied to rank the clinical relevance of the allergens including, for example, their proven ability to act as triggers of allergic disease. Such a robust, reliable, and verifiable database will allow risk assessors to calibrate and frame the RA around the defined public health objectives. For example, similarity of a novel protein to a potent allergen affecting many individuals will be of greater concern and of higher regulatory burden than if similarity is shown to an allergen that affects only a few allergic individuals and has low potency to elicit a reaction.

Currently, the allergen sequence databases used for allergenicity RA do not provide systematic data about their allergenic potential and often employ different inclusion criteria. For example, an allergen from peanuts called Ara h 5

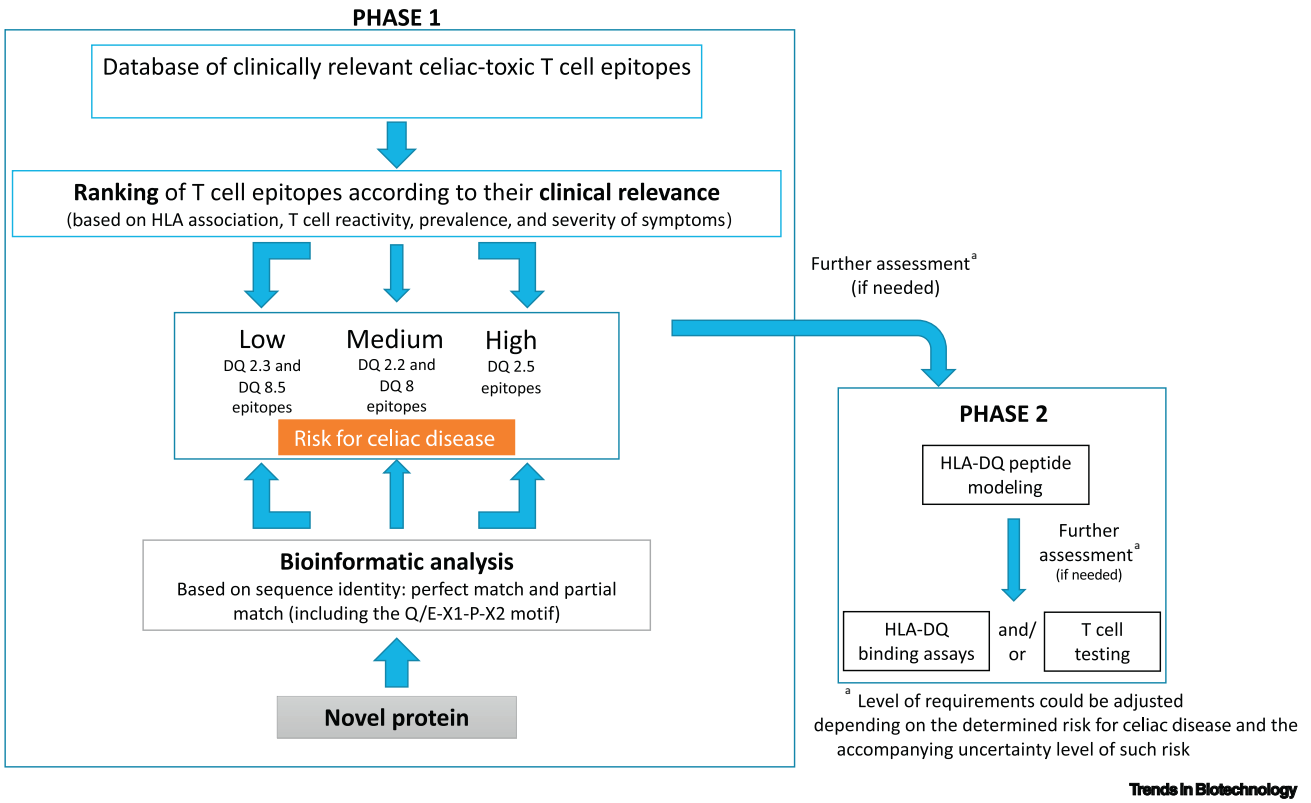

Figure 1. Strategy for the Risk Assessment of Novel Food Proteins Potentially Causing Celiac Disease. Clinical relevance and related features can be used to allocate any specific epitope to a defined 'category of risk'. In the case of celiac disease, we propose to rank relevant gluten T cell epitopes into three categories (i.e., high, medium, and low). When the ranking phase is completed, a subsequent targeted bioinformatic tool can be developed together with dedicated follow-up risk assessment steps depending on the risk level and accompanying uncertainties.

was identified using human sera by screening an expression library [7]. However, this protein is of very low abundance in peanut seed but is more likely to be expressed in pollen and is thought to be a problem for inhalant allergen cross-reactivity [8]. This allergen therefore poses a low risk to individuals with allergies to peanut seed but might be important in considering pollen allergy. Currently, such metadata are missing from allergen sequence databases, reducing the usefulness of bioinformatic analysis and making interpretation of the sequence similarities identified more difficult. Their lack also undermines efforts to develop advanced bioinformatics tools that might better predict the risk of a novel protein triggering allergic reactions by providing higher sensitivity, specificity, and accuracy than the classical FASTA algorithm. Such a bottom-up approach could benefit RA strategies for novel proteins and their potential capacity to trigger adverse immune reactions related to both celiac disease (CD) and IgE cross-reactivity.

Strategy for the RA of Novel Proteins and CD

In the case of CD, there have been several attempts to build a database containing many celiac epitopes[iii,iv]. However, no agreement has been reached on the inclusion criteria – based on precise evidence requirements – necessary to identify CD-relevant gluten epitopes [9]. While Propepper[iii] has more than 400 celiac epitopes, AllergenOnline[iv] comprises more than 1000 celiac epitopes, and Sollid and coauthors [9] identified only around 40 epitopes, showing large inconsistencies between databases. These discrepancies are the documented evidence of the lack of consensus on the inclusion criteria for building an appropriate/reliable database. Thus, while Sollid and coauthors [9] have compiled a reviewed list of the most important and immunodominant epitopes mainly based on the recognition of gluten peptides by $CD4^+$ T cells from one or more CD patients, the Propepper[iii] and AllergenOnline[iv] databases comprehensively cover the full scale of currently available

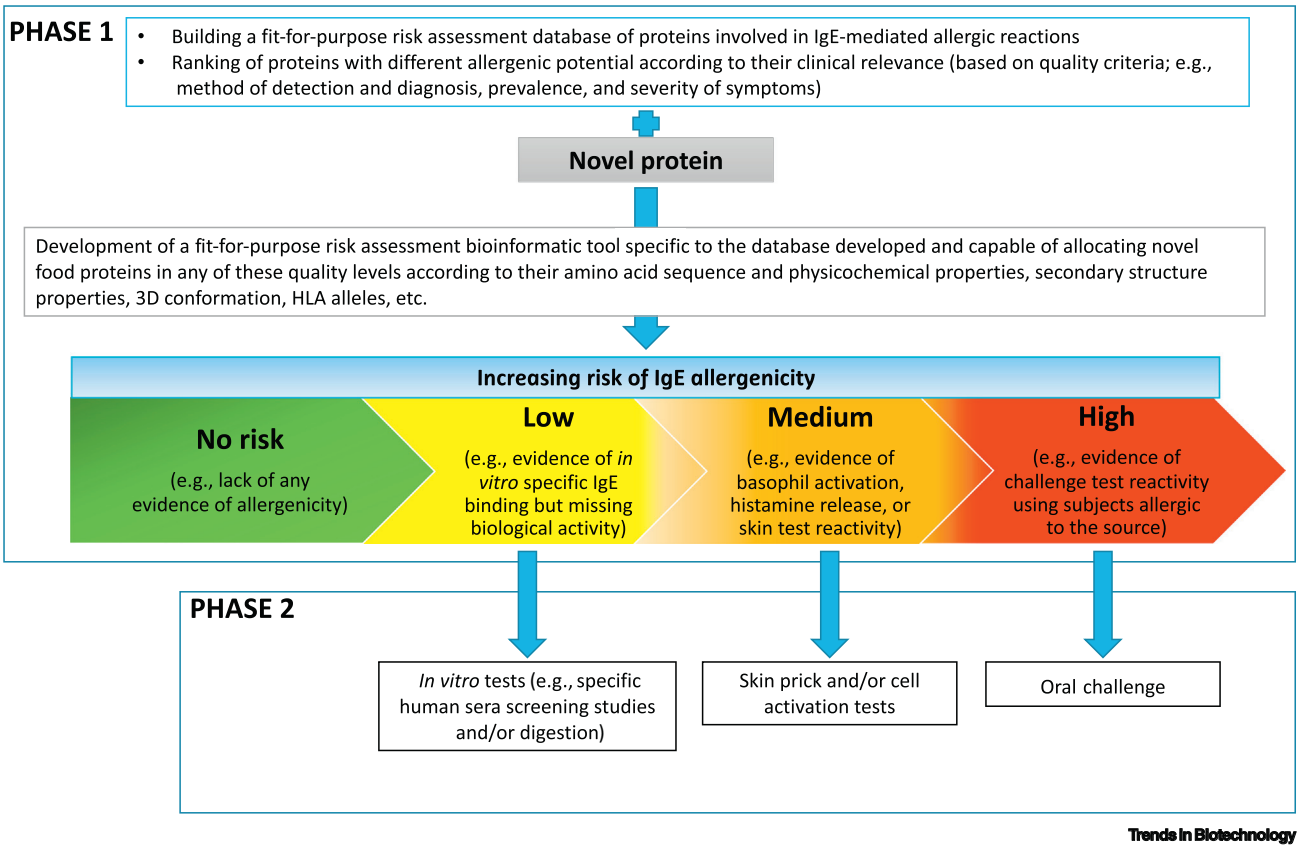

Figure 2. Strategy for the Risk Assessment of Novel Food Proteins Potentially Causing IgE-Mediated Adverse Immune Reactions. The principles described in Figure 1 can also be applied here. However, this is more challenging for IgE-mediated food allergy than celiac disease since many more molecules and foods are involved and our knowledge of what constitutes a clinically relevant allergen is more limited. Our proposal is to rank 'allergens' depending on the clinical relevance and quality of the scientific evidence, as described in this figure, and to attribute different follow-up actions depending on the risks allocated to the allergen 'hit'. When the ranking phase is completed, a targeted bioinformatic tool can be developed together with dedicated follow-up risk assessment steps depending on the risk level and accompanying uncertainties.

peptides that have been tested for T cell activation potential, even if the induction of celiac enteropathy has not been clinically demonstrated for the vast majority of these peptides. Likewise, it is important to stress that several additional relevant epitopes are as yet likely to be identified [9].

Our proposal can resolve the issue. Practically, we propose to define clear inclusion criteria that can serve to launch an initial screening phase designed to identify any potential protein fulfilling such *a priori* established criteria. The criteria defined by Sollid and coauthors [9] could serve such purpose, although other criteria may also be fit for purpose. Since no univocal list of criteria exists, an in-depth discussion and possibly consensus within the scientific community is indispensable. In a following phase, we propose to rank T cell epitopes according to their clinical relevance and related features to further boost the delivery of more transparent and robust RA to the public (Figure 1). Attempts in this direction have already been initiated [3,10], and additional efforts are imperative since new findings show that proteins from origins other than cereals might inherit hazardous potential for individuals with CD [11].

Strategy for the RA of Novel Proteins and IgE-Mediated Adverse Reactions

In the case of IgE food allergy, several databases have been developed and are currently in use for RA purposes[iv–vii], offering diverse possibilities. The criteria for inclusion of allergens and the number of entries varies between databases [12], while those relying on the reviewed UniProtKB[viii] may be affected by the curation strategies used in that database, which relies on the identification of canonical sequences to reduce redundancy [13]. These differences can lead to misunderstandings and different RA outcomes; for example, bioinformatic analysis results may vary depending on the database used to search for relevant hits.

Under the current paradigm, whenever a relevant hit is identified, the follow-up RA strategy analyses the quality of the pairwise sequence alignment and the specific similarity regions between the novel protein and the allergen. Clinical relevance is usually considered only as an additional element in the overall picture. Preliminary attempts to differentiate allergens (allergens and putative allergens versus proteins with insufficient evidence of allergenicity) within a database were initiated by AllergenOnline. However, there are no clear common views on how such a grading of allergenic potency should be assigned/interpreted and the weight-of-evidence that a risk assessor should attribute to such a classification. Attempts in such directions have been proposed for allergenic tree nuts [14]. Nevertheless, the current approach heavily relies on expert judgement to interpret *a posteriori* the outcome of the bioinformatic analysis of the assessment, which can lead to a lack of harmonization, reproducibility, and transparency of the RAs.

An alternative approach would be to define *a priori* the characteristic and clinical relevance that any potential allergen has and the specific follow-up actions to be undertaken if 'hits' are identified (Figure 2).

Concluding Remarks

Building fit-for-purpose RA databases for food allergy is an urgent priority. To this end, the ranking of CD-relevant gluten T cell epitopes and IgE allergens within any given database is of major importance as it provides more precise information on the clinical relevance of any given allergen to the RA process. This will translate into a sounder RA capturing specific needs considering the input from the scientific community and stakeholders. These two relevant actors should interact and eventually collaborate following principles previously described [15]. This approach will also allow refinement of the current, oversimplistic, RA view where proteins are categorized as allergens or non-allergens according to their inclusion or exclusion in a specific allergen database. Resources devoted to the development of sophisticated and refined bioinformatics tools will be better used once additional relevant protein features are defined and considered in the assessment.

Ranking allergens according to allergenic potential and the subsequent development of targeted bioinformatics tools founded on enhanced algorithms will streamline RA approaches, fostering more transparent and credible information for the public. Improving the quality of the bioinformatic assessment will also make it more straightforward to identify when further testing is warranted using more expensive *in vitro* and *in vivo* tests, which also have their shortcomings and are often not validated. They can also be difficult to execute, especially when relying on the availability of serum panels from food-allergic subjects. Developing an international consensus on a more robust approach to allergen-sequence database curation will be essential to improve the quality of allergenicity RA of foods produced by biotechnology and novel foods, which will be urgently needed in an era of climate change and transition towards more sustainable food systems.

Acknowledgments

The authors thank Claudia Paoletti and Elisabeth Waigmann for the inspiring comments.

Disclaimer Statement

A.F. is employed by the European Food Safety Authority (EFSA). The positions and opinions presented in this article are those of the authors alone and do not necessarily represent the views or scientific works of the EFSA.

Resources

[i]https://ec.europa.eu/knowledge4policy/publication/food-2030-innovative-eu-research-ensures-food-system-future-ready_en
[ii]www.fao.org/3/a-a1554e.pdf
[iii]https://propepper.net
[iv]www.allergenonline.org/
[v]https://comparedatabase.org/
[vi]www.allergome.org/
[vii]www.allergen.org/
[viii]www.uniprot.org/

[1]European Food Safety Authority (EFSA), via Carlo Magno 1A, 43021 Parma, Italy
[2]Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, Manchester Institute of Biotechnology, Manchester Academic Health Sciences Centre, The University of Manchester, Manchester, UK
[3]Department of Immunohematology and Blood Transfusion, Leiden University Medical Centre, Leiden 2333, ZA, The Netherlands
[4]Instituto de Investigación en Ciencias de la Alimentación, CIAL (CSIC-UAM), C/Nicolás Cabrera, 9, Campus de la Universidad Autónoma de Madrid, 28049 Madrid, Spain

*Correspondence:
antonio.fernandezdumont@efsa.europa.eu (A. Fernandez).

https://doi.org/10.1016/j.tibtech.2020.05.011

References

1. EFSA Panel on Genetically Modified Organisms (2011) Scientific opinion on guidance for risk assessment of food and feed from genetically modified plants. *EFSA J.* 9, 2150
2. EFSA Panel on Dietetic Products, Nutrition, and Allergies (2016) Guidance on the preparation and presentation of an application for authorisation of a novel food in the context of Regulation (EU) 2015/2283. *EFSA J.* 14, 4594
3. EFSA Panel on Genetically Modified Organisms (2017) Guidance on allergenicity assessment of genetically modified plants. *EFSA J.* 15, 4862
4. Dimitrov, I. *et al.* (2014) AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics* 30, 846–851
5. Maurer-Stroh, S. *et al.* (2019) AllerCatPro – prediction of protein allergenicity potential from the protein sequence. *Bioinformatics* 35, 3020–3027
6. Westerhout, J. *et al.* (2019) Allergenicity prediction of novel and modified proteins: not a mission impossible! Development of a random forest allergenicity prediction model. *Regul. Toxicol. Pharmacol.* 107, 104422
7. Kleber-Janke, T. *et al.* (1999) Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. *Int. Arch. Allergy Immunol.* 119, 265–274
8. Becker, W.M. *et al.* (2018) Peanut allergens: new consolidated findings on structure, characteristics, and allergome. *Allergol. Select* 2, 67–79
9. Sollid, L.M. *et al.* (2020) Update 2020: nomenclature and listing of celiac disease-relevant gluten epitopes recognized by CD4$^+$ T cells. *Immunogenetics* 72, 85–88
10. Fernandez, A. *et al.* (2019) Safety assessment of immune-mediated adverse reactions to novel food proteins. *Trends Biotechnol.* 37, 796–800
11. Petersen, J. *et al.* (2020) T cell receptor cross-reactivity between gliadin and bacterial peptides in celiac disease. *Nat. Struct. Mol. Biol.* 27, 49–61
12. Radauer, C. and Breiteneder, H. (2019) Allergen databases – a critical evaluation. *Allergy* 74, 2057–2060
13. The UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515
14. Javed, B. *et al.* (2017) A protocol for a systematic review to identify allergenic tree nuts and the molecules responsible for their allergenic properties. *Food Chem. Toxicol.* 106, 411–416
15. Fernandez, A. and Paoletti, C. (2018) Unintended effects in genetically modified food/feed safety: a way forward. *Trends Biotechnol.* 36, 872–875

Forum

Enhanced Strategies for Antibiotic Removal from Swine Wastewater in Anaerobic Digestion

Qi Zhou,[1,2,4] Xiang Li,[1,2,4] Shaohua Wu,[1,2,4] Yuanyuan Zhong,[1] and Chunping Yang [1,2,3,*]

There is a need for techniques that ensure antibiotic removal in anaerobic digesters for robust methane production. In this article, we discuss recent strategies for enhanced antibiotic removal from swine wastewater and offer insights on anaerobic digestion (AD) process design for improved antibiotic removal.

Antibiotic Removal in Aqueous and Sludge Phases

In the past two decades, the quantity of veterinary antibiotics used as disease prevention agents in animal feed has considerably increased. For example, veterinary antibiotic usage grew at a rate of 6000 tons per annum in China, surging from 97 000 t in 2010 to 132 000 t in 2016 [1]. However only 10–30% of consumed veterinary antibiotics are metabolized by livestock, causing a substantial amount of antibiotics to be excreted into swine wastewater as metabolites or in their original form, at concentrations of up to hundreds of micrograms per liter [1]. Swine wastewater has become a major pollution source of antibiotics. Thus, treatment of antibiotics in swine wastewater has also become a hot topic in research.

Biosorption and biodegradation are the two dominant antibiotic removal pathways during anaerobic digestion (AD) [2]. Biosorption processes involve bridging hydrophobic partitioning, cation exchange, electrostatic interactions, surface complexation, and electron donor–acceptor interactions (i.e., hydrogen bonding), where the extracellular polymeric substance (EPS) plays an important role due to the abundant functional groups on its surface. However, biosorption is only a phase-transfer phenomenon and cannot fully exclude the risk of antibiotic release into the environment [2]. Thus, biodegradation is typically needed to further transform (e.g., intermediates) or remove (e.g., complete mineralization) the remaining antibiotics from swine wastewater. Three principal degradation mechanisms for antibiotics have been reported: antibiotics as a growth substrate, organic matter as an electron acceptor, and co-metabolism. AD is generally a process using a sludge-dominated system for which the absolute mass of antibiotics in the sludge phases is expected to be higher than that in the aqueous phases [3].

Thus, removal of antibiotics in sludge typically occurs in the following order: rapid sludge sorption, followed by rapid sludge desorption, and then biodegradation (Figure 1). However, numerous experimental studies have demonstrated that anaerobic digesters are only moderately effective (40–77%) for antibiotic removal during AD treatment [2]. Accumulation of antibiotics remains a