EVENT REPORT

APPROVED: 13 November 2018

doi:10.2903/sp.efsa.2018.EN-1512

EFSA Scientific Colloquium 24 – 'omics in risk assessment: state of the art and next steps

European Food Safety Authority

and Jaime Aguilera[1], Margarita Aguilera-Gomez[2], Federica Barrucci[1], Pier Sandro Cocconcelli[3], Howard Davies[4], Nancy Denslow[5], Jean Lou Dorne[1], Lutz Grohmann[6], Lieve Herman[7], Christer Hogstrand[8], George E. N. Kass[1], Peter Kille[9], Gijs Kleter[10], Fabien Nogué[11], Nick J. Plant[12], Matthew Ramon[1], Reinhilde Schoonjans[1], Elisabeth Waigmann[1] and Matthew C. Wright[13]

Abstract

In recent years, the development of innovative tools in genomics, transcriptomics, proteomics and metabolomics (designated collectively as 'omics technologies) has opened up new possibilities for applications in scientific research and led to the availability of vast amounts of analytical data. The interpretation and integration of 'omics data can provide valuable information on the functional status of an organism and on the effect of external factors such as stressors. The European Food Safety Authority's (EFSA) 24th Scientific Colloquium on 'omics in risk assessment: state of the art and next steps explored the opportunities for integration of datasets produced via specific 'omics tools within the remit of EFSA's risk assessment approaches and tried to build further towards concrete paths of implementation. Discussions focused on genomics in microbial strain characterisation, metabolomics for the comparative assessment of GM plants and the use of 'omics for toxicological and environmental risk assessment. From the Colloquium it became clear that 'omics technologies are a valuable addition in some aspects of risk assessment of food and feed products and the environment, especially now that this technology is almost mature and stable. However, a consistent reporting framework for data collection, processing, interpretation, storage and curation should be further drawn up together with national and international organisations before 'omics technologies can be routinely used in risk assessment. For 'omics datasets in chemical and environmental risk assessments, the use of 'omics technologies alongside current toxicological or environmental risk assessment approaches is needed to re-inforce confidence and expertise before implementation of these datasets as a standalone tool in risk assessment. Test cases could be worked out to enhance confidence in the use of 'omics datasets in risk assessment.

Key words: metabolomics, next-generation sequencing, 'omics, proteomics, risk assessment, transcriptomics

Question number: EFSA-Q-2018-00113

Correspondence: Any enquiries related to this output should be addressed to GMO@efsa.europa.eu

[1] European Food Safety Authority

[2] University of Granada

[3] Universitá Cattolica del Sacro Cuore

[4] Honorary Research Fellow, James Hutton Institute

[5] University of Florida

[6] Federal Office of Consumer Protection and Food Safety, Berlin
[7] Flanders Research Institute for Agriculture, Fisheries and Food, ILVO
[8] King's College London
[9] Cardiff University
[10] RIKILT, Wageningen University
[11] French National Institute for Agricultural Research, INRA
[12] University of Leeds
[13] Newcastle University

Acknowledgements: EFSA wishes to acknowledge all participants to the 24th EFSA Scientific Colloquium for their valuable contribution during the discussion. The organising committee would like to thank Francesco Amoretti and Justyna Slodek Wahlstrom for the administrative organisation of this Colloquium.

Suggested citation: EFSA (European Food Safety Authority) and Aguilera J, Aguilera-Gomez M, Barrucci F, Cocconcelli PS, Davies H, Denslow N, Dorne JL, Grohmann L, Herman L, Hogstrand C, Kass GEN, Kille P, Kleter G, Nogué F, Plant NJ, Ramon M, Schoonjans R, Waigmann E and Wright MC 2018. EFSA Scientific Colloquium 24 – 'omics in risk assessment: state of the art and next steps. EFSA supporting publication 2018:EN-1512. 30 pp. doi:10.2903/sp.efsa.2018.EN-1512

ISSN: 2397–8325

Table of Contents

1. Introduction

The 24th meeting in the EFSA Scientific Colloquium Series was held in Berlin, Germany on 24–25 April 2018 and addressed the state of the art on 'omics technologies and next steps for its integration in the risk assessments performed by EFSA.

'omics is a term that covers several technologies used to characterise and quantify the roles and relationships of large sets of different types of molecules in an organism. Genomics can facilitate the analysis of entire or component genome sequences of an organism. Transcriptomics and proteomics provide significant bodies of information on temporal and spatial expression of genes and gene products and their modifications, respectively, while metabolomics captures data for a large pool of metabolites. The interpretation and integration of 'omics data can provide valuable information on the functional status of an organism and on the effect of external factors such as stresses.

In 2014, EFSA started mapping the use of 'omics tools in risk assessment related to food and feed safety (EFSA, 2014). Building further towards a concrete path of implementation, the 24th EFSA Colloquium discussed diverse topics for which EFSA intends to exploit 'omics datasets to support scientific safety evaluation. The outcome of this Colloquium aimed at helping risk assessors in the process of incorporating 'omics tools in the risk assessment of products for the food and feed chain.

Following the opening plenary session with introductory keynote talks, the meeting focused the discussions on four specific topics:

- Discussion Group 1 discussed the use of genomics for the identification and characterisation of microbial strains used in food and feed products.
- Discussion Group 2 focused on the use of metabolomics in the comparative risk assessment of genetically modified plants.
- Discussion Group 3 discussed the use of 'omics in human risk assessment of chemicals.
- Discussion Group 4 focused on the use of 'omics in environmental risk assessment.

The 24th EFSA Scientific Colloquium brought together participants with scientific expertise in 'omics technologies and/or risk assessment related to the specific topics of the discussion groups from academic institutions, national and international research bodies, regulatory authorities, non-governmental organisations and industry. The Colloquium was attended by 116 experts.

This event report presents the abstracts of speakers in the opening plenary session and summarises the discussions and conclusions of the Colloquium.

2. Abstracts of speakers in the introductory plenary session

Taxonomic and toxicogenic potential derived from whole-genome sequencing (WGS) information

María José Figueras, *University Rovira i Virgili, Pere Virgili Health Research Institute, Reus, Spain*

New technologies enable whole-genome sequencing of a microorganism both relatively cheaply and easily. In parallel, many new bioinformatics tools have been developed for genome comparison such as the average nucleotide identity (ANI) and *in silico* DNA–DNA hybridisation (isDDH). ANI values of >96% showed a very good correlation with isDDH values of >70% for microbial genomes belonging to the same species. Therefore, the use of ANI and isDDH have been proposed as the 'gold standards' for species delineation using genomes, as both allow a fast comparison between bacterial genomes. Use of these tools will facilitate a more objective defining of new taxa using approaches that are straightforward and less prone to errors or misinterpretations. However, looking at the available genomes in the databases it is evident that old problems still remain, e.g. the selection of inadequate strains for comparison, the existence of mislabelled genomes that may lead to misinterpretations, the use of old names or incorrect taxonomy, etc. In addition, problems also arise due to the fact that many of the available genomes are drafts of varying quality. To tackle these problems, we use our experience with the analysis of the genomes of the genera *Aeromonas* and *Arcobacter* that can be extrapolated to other genera. In addition to taxonomic identification, ANI can be useful to determine if the compared genomes are epidemiologically related. In this sense, genomes belonging to the same strain or to strains derived from the same clone show ANI values of 99.99–100%. Furthermore,

genome analysis can provide toxicogenetic information by recognising potential antimicrobial resistances and the presence of characteristic virulence genes.

Mechanistic modelling of metabolism to understand and predict plant performance: a case study in growing tomatoes

Sophie Colombié, Bertrand Beauvoit, Sylvain Prigent, Pierre Pétriacq, Annick Moing, Christine Nazaret, Jean-Pierre Mazat and Yves Gibon, *UMR1332 Fruit Biology and Pathology*, *INRA*, *University of Bordeaux*

A key objective for plant sciences is to understand what influences the growth and quality of plant products, with a view to improving them. Metabolism is an obvious target for improvement and an understanding of the mechanisms that link it to phenotypes will help focus breeding strategies and/or improve agricultural practices. However, metabolism is extremely complex, with a number of highly intricate pathways that undergo extensive reprogramming throughout plant organ development, making it difficult to find the right targets for improvement. Systems biology, which can be defined as an iterative theory and experimentation approach, represents a great opportunity to deal with this complexity. On the one hand, newly emerging top-down approaches have been shown to provide highly predictive statistical models linking metabolome and plant performance. On the other hand, mechanistic modelling of metabolism has never been so easy thanks to significant advances in analytics and computing power. We have developed kinetic enzymatic models consisting of sets of ordinary differential equations to study central metabolism as well as stoichiometric models describing larger areas of metabolism. The models were used to study tomato fruit metabolism throughout growth and development. In particular, they have revealed and explained the importance of carbon management in growing fruits, which leads us to formulate strategies to improve yield and better anticipate the decline in yields linked to climate hazards. These models are currently being extended to other metabolic pathways, integrated into ecophysiological models and are being transferred to other fruit species.

The value of 'omics to chemical risk assessment?

Tim Gant, *Centre for Radiation, Chemical and Environmental Hazards*, *Public Health England*

'omics technologies have developed from metabolomics (that measures molecular composition changes) as the founder 'omic method based on developments in nuclear magnetic resonance in the early 1990s to methods now that allow measurement of most aspects of cellular biochemistry via appropriate technological developments. The development of capillary sequencing, and subsequent microarray development that established transcriptomics in the late 1990s has led to transcriptomics probably being the most applied 'omic method to date. However technological developments in all the 'omic platforms are now making all of these 'omic methods quite accessible to most laboratories. From these beginnings further technological advances such as high-throughput sequencing and mass spectrometry allied with developments in computing and the internet, allowing the processing and handling of large data, have all combined to ensure the firm establishment of the 'omics in biological science, including risk assessment. 'omics contributes to all levels of risk assessment but is perhaps best established in hazard identification. 'omics methods allow the evaluation of many endpoints simultaneously allowing the recognition of hazard potential in a wider biological context than has been previously possible. Furthermore such change can occur before the development of pathological change and thus is potentially more sensitive as an endpoint. As a result alterations in transcript levels have for example been suggested as points of departure for assessing toxicity. In doing this though a question has to be considered of the relationship of gene transcript levels to an apical endpoint and therefore how meaningful are changes in some gene expression levels. Does for example a change in a stress-response gene or metabolism gene represent toxicity, or is this the cell utilising its hormetic potential such that on removal of the insult reversion to the previous baseline phenotype occurs with no long lasting damage to the cell. Risk is a function of exposure, hazard and susceptibility. Advances made in understanding epigenomic methods and the association of some epigenetic change with exposure have allowed 'omics to find novel applications in assessing exposure both historical, through epigenetic biomarkers, and in real time via the advances in analytical methods that can now also be argued to be part of the 'omics landscape. Finally we must consider the profound effect that the 'omics have had on our understanding of genetic diversity and how this understanding will impact risk assessment. Whole genomes can now be sequenced more quickly and cheaply than has been

previously possible allowing a much greater understanding of genetic diversity. Genetic diversity impacts on risk by altering susceptibility. In order to factor this into risk assessment we need to understand more about the key xenobiotic molecular interactions and mechanisms of toxicity. The 'omics technologies have much to contribute in this arena and the data can be visualised in new ways such as adverse outcome pathways. Armed with these pathways, and an understanding of how genetic difference can impact at critical points in the pathway, risk assessment may address susceptibility in a more targeted way, at least to sub-population level. Important in all applications of 'omics in risk assessment is the use of bioinformatics. Bioinformatics methods have developed alongside the 'omics but the application of these methods is still far from established in a consistent manner. This can be challenging because the application of different bioinformatics methods to the same data set can lead to quite different conclusions. It is therefore essential that the 'omics community continue to work towards best practices in all areas of the application of 'omics in risk assessment to ensure continued utility in the assessment of risk. Can then 'omics contribute to chemical risks assessment? The answer is a clear yes, and indeed has done, but there is still quite a long way to go before the use of 'omics in risk assessment becomes routine.

'Omics Prospects in Ecological Risk Assessment (OPERA)

Christer Hogstrand[1], Peter Kille[2]

[1]*King's College London, UK, christer.hogstrand@kcl.ac.uk*
[2]*Cardiff University, UK, kille@cardiff.ac.uk*

'omics data provide value adjunct to current approaches to the environmental risk assessment (ERA) process. A range of approaches may be employed within tier decision making to support stratification of assessment delivering 3Rs imperatives while providing a mechanistic based evaluation of ecological risk. Furthermore, 'omics approaches have value both for prospective and retrospective (monitoring) ERA. However, the complexity and quantity of data inherent within 'omics approaches places the duty for appropriate implementation on the assessment team who should endeavour to exploit the data to answer specific ERA questions. We would like to consider four ERA challenge areas where 'omics currently delivers novel insight: 1) Chemical read-across to estimate differential species sensitivity; 2) Non-lethal 'effect pathways'; 3) Classification; and, 4) Community/ecosystem impact. Estimating chemical read-across to predict differential sensitivity and species at risk can exploit an Ecological-QSAR model (Quantitative Structure–Activity Relationship), as implemented by Eco-SAR for example. These approaches are applicable for chemicals with established biomolecular targets, however, it is important to consider toxicokinetics separately. Genomic resources provide the ability to assess species-conservation of targets and evaluation of possible comparative toxicodynamic responses of a rapidly increasing number of organisms, with increasing evidence that even simple estimates of conservation are informative. Dose/time course analysis of changes in mRNA, proteins and metabolites can be interrogated to investigate key sublethal 'effect pathways' with clear mechanistic links to adverse outcome, such as core genes involved in DNA repair. The same 'omics data can be used in combination with unsupervised clustering methods in a core suite of surrogate species to support chemical classification. This has successfully been employed using metabolic data to classify toxicological potential of pharmaceuticals, the limitation within the ecological sphere being the generation of 'omics profiles for a comprehensive database of hundreds of compounds with known modes of action. Holistic ecological community impact, either prospective and as a retrospective monitoring, is very straightforward and cost effective using metagenomic profiling with the only limitation being the availability of reference databases to support ecological interpretation. These aspects provide immediately available opportunities for 'omics to support the ERA process.

3. Summary of discussion groups

3.1. Discussion Group 1 – Genomics for the identification and characterisation of microbial strains used in food and feed products

3.1.1. Introduction

Microorganisms are used to obtain a variety of foods and food products. They are responsible for the fermenting process in foods such as bread or cheese and are also used to produce a variety of substances used for food production, such as food enzymes and food additives. Also, in animal nutrition, feed additives including enzymes, vitamins and amino acids are often obtained through microbial fermentation. Microorganisms can also be directly added to food and feed, as is the case with probiotics.

In the European Union, the current legislative framework establishes that all food enzymes, food additives, feed additives and, in some cases, microorganisms directly added to food, need to undergo a risk assessment before market authorisation, conducted by EFSA. The identification and the characterisation of the microorganism used as production strain or as active agent are two fundamental aspects of the assessment, as they affect further steps of the evaluation of the product. For example, if the strain is capable of producing toxic compounds, the product should be free of those compounds or if a production strain carries antimicrobial resistance genes, those genes should not be present in the product. In addition, genetically modified strains require full characterisation of the added modifications.

At this time, with the availability of next-generation sequencing technologies, it is possible to sequence the full genome of a microorganism in a fast and affordable way. Parallel with the development of WGS technologies, there has been an increase in the number of bioinformatics tools, platforms and databases for annotation and analysis. As a result, today WGS analysis is routinely performed by many laboratories and the number of microbial strains completely sequenced increases year after year. Having the complete genome sequence of a microorganism offers the possibility to study many of its characteristics that are relevant for safety, e.g. the presence of genes involved in antimicrobial resistance or in the synthesis of undesirable substances such as toxins.

In recent years, EFSA has been developing capabilities and discussing criteria for the use of WGS in the characterisation and tracing of foodborne pathogens (EFSA, 2015, 2018). Now EFSA is also in the process of implementing the use of WGS analysis for the characterisation and safety assessment of microbial strains intentionally inserted in the food chain or used as production strains. For example, the EFSA NDA Panel Guidance on novel foods (EFSA NDA Panel, 2016) requests that microorganisms not eligible for Qualified Presumption of Safety (QPS) status should be characterised by 'fully assembled and validated whole-genome sequence analysis to enable the detection of virulence-related genes, antibiotic resistances and their potential horizontal transfer, and other potentially adverse metabolic features'. Similarly, the new EFSA FEEDAP Panel Guidance on the characterisation of microorganisms used as feed additives or as production organisms provides more detailed guidance on the ways in which different possible traits of concern, as well as any genetic modification, should be investigated by WGS of the microbial strains used for animal feed. EFSA also intends to apply these criteria to food products, with the aim of harmonising assessments throughout similar products.

In this discussion group the participants discussed the potential use of WGS analysis for a better hazard identification of microbial strains introduced in the food/feed chain or used as production strains, the extent to which WGS data enable an accurate characterisation, the most suited bioinformatics tools and the criteria for interpreting the data.

3.1.2. Outcome of the discussion

Whole-Genome Sequencing (WGS) for the identification of microbial strains at species and strain levels and the challenges encountered

The group discussed the value of WGS analysis for an accurate taxonomical identification of a microorganism, both at the species and strain level, including the best practices and limitations in data handling and analysis. The following issues were addressed:

The choice between the use of assembled genome sequences, contigs or short DNA sequences as single genes.

In principle, the source of data for the identification could be the raw, unassembled set of contiguous sequence runs (contigs), the assembled full genome sequence, or certain target genetic sequences of taxonomical value. The most significant message derived from the discussions was that the data and method applied for the identification should be fit for purpose. This implies cost efficiency based on the available facilities, the current need for the identification and the complexity of the species/taxon to which the microorganism belongs.

For the identification of well-known species or microorganisms considered low risk, complete genome sequences would not always be necessary. In contrast, for not so well known species and species with pathogenic strains, as well as identification at the strain level, complete genome sequence data can be of higher value. It should also be taken into account that large variable regions and rearrangements in the chromosome could be missed by using only contigs and therefore comparison of the whole-genome sequence would be recommended.

For some species/taxa, there are limitations in the availability of reference genomes for comparison. For that reason it could be more suitable to use shorter sequences. Examples for this are the use of the 16S rRNA gene sequence, which is historically the most frequently used for prokaryotes and for which many reference sequences are available. For certain species/taxa for which the 16S rRNA gene offers too low a resolution, other sequences from housekeeping genes could be used for reliable identification. Although the taxonomy of yeast and fungi is much more complex than that of prokaryotes, shorter signature sequences are available for identification purposes and could be extracted from the WGS.

Finally, it is important to consider that procedures and technologies for applying WGS for microbial identification are evolving rapidly and will influence assessment substantially in the future.

Standardisation of the WGS approach for microbial identification: methodologies and tools.

Minimal quality standards for genome sequence data were recently proposed as guidelines on how to apply WGS for bacterial identification (Chun et al., 2018). For eukaryotes, no such standards exist at this moment. Parameters most important as quality standards suggested: completeness of the sequence, sequence coverage, number of contigs, genome size and absence of contamination. The weight and relevance of each parameter vary depending on the species/taxon and the data/method used for the identification (see point 1). With respect to contamination, it is recommended to use specific bioinformatics software for its detection.

When WGS data are available, species identification is possible based on digital DNA–DNA hybridisation (dDDH) and ANI. For bacterial identification at the species level, the boundary is set at a dDDH >70% and for ANI it is suggested to be set at >95–96%. Several platforms are available to calculate ANI, and orthoANI and one platform is available for defining dDDH.

Epidemiological investigations of pathogenic outbreak strains revealed a clonal relationship with ANI values of >99.9%, which indicates that WGS can be also used for identification purposes at the strain level. This faces the challenge that there is not a recognised definition of strain at the genomic level. In bacteria, individual single nucleotide polymorphisms (SNPs) may be of limited value for strain identification as they can accumulate due to genetic drift during bacterial growth of the same strain.

Main challenges for using WGS for bacterial identification.

It is essential that the WGS process is performed under high laboratory standards, starting from the growth of the strain and the DNA extraction. It is important to avoid contamination of data with

sequences from other organisms or, if present, to detect it with bioinformatics tools during the data analysis procedure. To provide reliable genome sequences for comparison, databases should be sufficiently curated, avoiding as much as possible mislabelled genomes and non-verified metadata (e.g. contaminated genomes, chimeras). For some species, properly annotated genomes to be used as references are missing. The link between the type strain of a given species and its whole-genome sequence in the database is made by the culture collection where the type strain is available.

At what point and under what circumstances can WGS enable an accurate characterisation of genetic modifications of strains used for the production of food/feed additives?

Through the use of a WGS analysis, it might be possible to identify and characterise genetic modifications inserted in a given microorganism without the need to describe all the steps of the modification and to design targeted experimental tests to demonstrate that such modifications are exactly the intended ones. This can be particularly useful in the assessment of strains with a long history of genetic modifications, in which different vectors and marker genes were used. The group discussed the best ways to undertake such characterisation and the limitations of the methodology.

1) The choice of the comparator

WGS can be useful for the characterisation of the genetic modifications in microbial strains. The way it is applied should be decided on a case-by-case basis and should be scientifically substantiated. A prerequisite is the availability of high-quality sequence data of the organism under characterisation and the existence of genome sequences for comparison. The WGS of the genetically modified microorganism should preferably be compared with the non-modified parent or recipient strain but also a reference (type strain) genome available in a database or another strain of the same species could be used, as long as it is genetically well known and described.

2) Genetic modifications versus point mutations

Larger introduced genetic modifications can easily be differentiated and characterised. However, small changes of few nucleotides inserted/deleted by genetic modification or gene editing cannot easily be discriminated from natural variation. This is particularly relevant for bacterial genomes, whose genetic plasticity and variability are much higher than that of eukaryotes and mutations may occur to a higher rate throughout the genome.

Several mining tools are available for data analysis and may be applied on a case-by-case basis.

3) Extrachromosomal elements and repeated sequences

In general, WGS analysis can identify plasmid sequences, although there are some limitations. Some small plasmids could be missed due to, for example, inappropriate DNA extraction methodology. In this sense, there is not a validated procedure for DNA extraction for WGS purposes in place, although this would be desirable. Conversely, large plasmids could be difficult to be discriminated from contigs. Databases are not comprehensive for extrachromosomal genetic elements, and this in some cases may limit data interpretation.

Long repeat sequences can sometimes be difficult to correctly assemble. In those cases, long read sequencing techniques are helpful.

Are existing analytical tools 'fit for purpose' for the identification of sequences involved in antimicrobial resistance, toxigenicity and virulence and therefore in supporting risk assessment? What do we require to obtain maximum predictability in terms of safety evaluation derived from WGS data?

One of the big advantages of obtaining the WGS data of a microorganism is that it can be compared, using suitable bioinformatics tools, against curated databases to look for genes potentially involved in antimicrobial resistance (AMR), toxigenicity or virulence. This might enable predicting safety concerns of the assessed microorganism and driving possible further tests to characterise the risk. The following points were addressed on the state of the art and predictive value of this approach.

1) Virulence and toxicity

The current level of curation of databases with annotations to sequences involved in virulence and toxicity is not yet sufficient to allow a WGS analysis to be used as the only tool for risk assessment. Several virulence genes are species/genus related and multifactorial. Many sequences currently annotated as virulence genes may be actually coding for structural components, such as flagella or stress-related proteins. The currently available information on sequences encoding bacterial toxins is more reliable than that for virulence. Nevertheless, there was a general agreement on the need to develop, curate and update comprehensive databases, especially for the correct annotation of virulence genes.

For bacteria, WGS analysis can add to the risk assessment of virulence on a case-by-case basis and the contribution should be scientifically substantiated. It is important to keep in mind that virulence is, in most cases, a multifactorial phenotype with several genes involved. There are also no set-out similarity/homology threshold values to identify hits as virulence factors or toxins, nor standardised procedures for analysing and interpreting the results of the searches. Therefore, a key point is that any hits obtained after screening the databases should be used as one element in a weight-of-evidence approach, which would include additional data from other sources, e.g. the scientific literature or confirmatory experimental tests.

For fungi, specific dedicated databases exist for some species, (e.g. *Saccharomyces cerevisiae*), which could be used for targeted searches, although sequences should be available in the general nucleotide databases. The WGS of the assessed microorganism can be searched against those databases to identify putative gene clusters that code for known or possibly toxic secondary metabolites. The interpretation of the results is currently still very challenging. The metabolic function and possible toxicity of many secondary metabolites is not clear yet. In addition, most secondary metabolites are synthesised by complex biosynthetic pathways encoded by different clusters and modular genes, for which regulation remains unknown. Therefore, the value of this approach for fungi is particularly limited to well known cases.

2) Antimicrobial resistance

WGS analysis for the detection of AMR genes is, in many cases, useful for risk assessment purposes, although interpretation of the results is still sometimes challenging. For virulence genes, there are no set-out similarity threshold values to identify a functional AMR gene and there is not a unique reference database to be used for searches. Therefore, WGS analysis is not yet totally reliable as a predictive tool for AMR and should be combined with other data, in particular phenotypic tests (determination of minimum inhibitory concentrations to antimicrobials). Apart from this potential predictive value, WGS analysis is especially useful to provide information on the genetic nature of AMRs found in previous phenotypic tests, and to check whether AMR genes used in genetic modifications remain in the genome of the strain.

3.1.3. Conclusions

WGS analysis is a powerful technology and useful approach for the comprehensive characterisation of a microorganism at the structural level, including the characterisation of introduced genetic modifications. It allows predicting potential functionalities and bioactivities present in the analysed microorganism. Therefore, WGS analysis is a useful tool for risk assessment and should be applied on a case-by-case basis, depending on the requirements. It offers valuable information to be used in a weight-of-evidence approach. Transparency in the data collection and evaluation is very important. Rapidly evolving technologies in the WGS field will strongly influence its potential for application in risk assessment.

At this point in its development, the use of this technology still faces challenges. The quality of the whole WGS procedure is essential and standards for quality control of the laboratory practice and the data analysis should be in place. The interpretation of the data requires expertise and input from other knowledge sources, such as literature and further experimental data. There is a need for standardisation of the process, reference datasets and curated databases. When using WGS analysis in risk assessment, the uncertainties should be identified and described to appropriately weigh the value of the results.

3.2. The use of metabolomics in the comparative risk assessment of GM plants

3.2.1. Introduction

Genetically modified organisms (GMOs) and derived food and feed products are subject to risk analysis and regulatory approval before entering the market in the EU. In this process, the role of EFSA is to independently assess and provide scientific advice to risk managers on any possible risks that consumption or cultivation of a GMO may pose to human and animal health and the environment. The main focus of EFSA in the GMO field lies in the evaluation of GMO market registration applications and in the development of risk assessment and monitoring guidelines.

The risk assessment of food and feed produced from GM plants and imported in the European Union covers not only the assessment of intended effects (e.g. protein(s) newly expressed in the GM plant) but also the identification and assessment of unintended effects linked to the genetic modification and their biological relevance. For the latter, a comparative approach is used, aiming at identifying biologically relevant differences by comparing the GM plant with a non-GM counterpart (EFSA GMO Panel, 2010, 2011). The underlying assumption of this approach is that non-GM crops obtained by conventional breeding have gained a history of safe use for consumers and/or animals. In the comparative analysis, a two-step approach is followed: first differences between the GM plant and its appropriately selected comparator are identified (difference test); second an assessment of whether the characteristics of the GM plant fall within the natural variability estimated from a set of non-GM commercial reference varieties with a history of safe use (equivalence test) (EFSA GMO Panel, 2010, 2011). The simultaneous use of these two tests allows an assessment of the relevance of observed differences, taking into account the natural variability of the measured endpoints. The experimental design for comparative analysis includes non-GM commercial reference varieties together with the GM plant and its comparator in a set of field trials carried out under representative receiving environments (EFSA, 2011). These field trials need to be carried out in a minimum number of locations reflecting the range of environmental conditions and agricultural practices under which the GM crop will be grown once commercialised. The lay-out of plots in each location should be designed to account for local environmental variability within a location and to have sufficient data for each variety per location. Plots with all varieties tested should be grouped in blocks, which are to be replicated at least four times per location. The lay-out should be a randomised block design.

In this break-out session participants discussed three different scenarios in which the use of metabolomics data for the comparative assessment of GM plants could be envisaged:

- Scenario 1: Comparative assessment would be based on the establishment of metabolomics profiles of the tested material: the GM plant, its non-GM comparators and the non-GM reference varieties. The metabolomics profiles of the GM plant and the non-GM plants would replace the current endpoint-by-endpoint comparison.
- Scenario 2: Metabolomics analyses would be used to measure more and/or different endpoints within the current comparative approach.
- Scenario 3: Metabolomics analyses would be used to complement the current comparative endpoint analysis on a case-by-case basis depending on the nature of the genetic modification.

3.2.2. Introductory talk by Dr Esther Kok: 'omics as a refined tool for hazard identification in new plant material

'omics analyses allow for a broad analytical screening of the metabolic and physiological status of new plant varieties and may comprise transcriptomics, proteomics and metabolomics analyses. In this way a detailed overview of a new plant variety can be obtained that can be compared with plant varieties that are considered to be safe. For risk assessment strategies, 'omics profiles of new plant varieties would be compared with their conventional counterparts. As there are many thousands of endpoints in each individual profile, this means that even when profiles are expected to be highly similar (e.g. in subsamples from the same plant) many endpoints will show significant differences as a result of the statistical assessment (normal distribution, confidence interval of 95%, 5% may be considered as

significantly different). To overcome this issue and to focus on relevant differences, the SIMCA-based one-class model has been developed (van Dijk et al., 2014).

To build this classification model, principal component analysis (PCA), a multivariate statistical approach, is used to reduce the dimensionality of the data. This model uses the profiles of a (training) set of varieties, considered to be 'safe', to build the one class of 'safe' profiles. In a second set, this one-class model is assessed by a set of profiles that are also considered as safe (cross-validation). The optimal numbers of components to be included into the model is then determined as the highest number still causing all profiles of the validation set to belong to the baseline class. Finally a third independent set is used to test the model accuracy. Depending on the tightness of the model, the external sample can fall inside or out. After cross-validation, the model is ready to be used for new (test) profiles.

Based on this model, the new profiles can be classified as inside or outside the one 'safe' class. In cases in which the new profile classifies as inside, it means that the profile is very similar to profiles that are in this group and no further assessment is required. If the profile is classified as outside, this does not mean that the variety is not safe, but it does mean that the profile will need to be further assessed for the basis of the 'out' classification.

The model has been developed in a way that it is conservative, in the sense that it sets the border tightly around the 'safe' varieties, leading to an earlier 'out' classification. In other words false-positive results are expected, but the focus is to avoid any false-negative results. For evaluating model accuracy, the behaviour of the non-GM comparator of the GM crop is checked; if it already falls outside or is located at the border of the model, then the model is considered to be of not sufficient quality.

An example was given for maize, in which the one-class model was tested using transcriptomics profiles. The example showed how a GM maize variety consistently fell within the one class based on a training set of 14 varieties designated as 'safe'.

The presenter concluded that unintended effects can be identified by 'omics. The question how these datasets can be integrated into the risk assessment was discussed. The proposed statistical 'one-class' approach for interpretation of 'omics data sets may be applied.

3.2.3. Outcome of the discussion

Scenario 1: The comparative assessment would be based on the establishment of metabolomics profiles of the tested material

1) What would be the minimal accepted metabolomics dataset in the context of a GMO dossier?

It is estimated that there are around 200,000 metabolites in the plant kingdom (Saito and Matsuda, 2010) and between 5,000 and 25,000 for a specific species (Hegeman, 2010 and contained references). In general, 200 to 500 metabolites may be observed by targeted metabolomics, whereas upwards of 1,500 metabolites have been detected in untargeted metabolomics studies (Markley et al., 2017). The number and nature of metabolites detected depends upon several parameters including extraction media (polar vs non-polar) and detection methodology [gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), liquid chromatography-tandem mass spectrometry (LC-MS/MS), etc.] used.

The idea to cover the largest number of metabolites possible (polar and non-polar) to produce the metabolomics profiles and to focus on metabolite patterns instead of individual endpoints, was discussed in the group. However, the group considered that with a completely untargeted metabolomics approach, in which mass spectral fragmentation patterns of unfractionated extracts are produced, it may be difficult, if the profile of the GM plant differs from those of the reference varieties, to identify the metabolites responsible the difference and assess them for safety concerns. In a more targeted approach, metabolite identification is more likely to be possible and outcomes are more quantifiable. Nevertheless, significant questions remain on the appropriate analytical approaches and parameters to use to capture compounds of interest.

The 50–60 endpoints provided by the OECD consensus documents do not necessarily provide guidance on which endpoints to include in an 'omics analysis as suggested in Scenario 1. Indeed,

when OECD documents advise screening for minerals, this would not be possible using conventional metabolomics approaches such as GC-MS and LC-MS. Metabolite pools will vary quantitatively and qualitatively in a crop-specific manner, so any standardised metabolite coverage should reflect this situation. For sufficient coverage, combined 'omics approaches could be used if needed.

The group indicated that standardisation of detection and extraction methodology should be developed. Standard operating procedures would need to be set up upfront, e.g. for extraction protocols, chromatography and mass spectrometry. Linear extraction methods are important for easy comparison. In addition, the possibility for creating standardised plant reference material was discussed.

2) Is the one-class model approach an appropriate approach for the use in GMO risk assessment?

The group discussed the possible use of the unsupervised one-class model approach in the risk assessment of GM plants and agreed that it might be a suitable tool to compare the 'omics profile of the GM plant variety with the one class defined by the non-GM reference varieties. No difference test between the GM variety and its conventional counterpart would be needed when using the one-class model. For 'omics analyses, it was recognised that multivariate approaches are more preferable than the current univariate approaches, basing themselves on a statistical endpoint-by-endpoint comparison. Multivariate approaches are able to capture the dependencies among variables and can better control global false discovery rate. The group further discussed the possibility to explore alternative approaches besides PCA, as a basis for the one-class model, and the need for robust statistical parameters to be used for analysis (e.g. confidence intervals, threshold values).

The samples to be included in the training set for the construction of the one-class model were also discussed. To be representative, a good coverage of different plant varieties, genetic backgrounds and environmental growth conditions should be included. To be in line with the current comparative approach used in the GM plant risk assessment, only plants with a history of safe use for food and feed should be included in the training set for the one-class model.

An alternative scenario for the one-class model was proposed by one of the participants, in which a set of GM plants with the same genetic modification but different genetic backgrounds is tested with 'omics against their respective conventional counterparts. This would facilitate the identification of changes in profiles (not endpoint by endpoint) that are truly linked to the genetic modification. Only if the GM plants fall consistently outside the one-class model, the implication of these differences on the safety of the genetic modification needs to be further checked.

3) What are the advantages and disadvantages when this scenario would be used in the GM risk assessment?

The use of metabolomics on crops for which no endpoints have been set up by OECD consensus documents and the possible discovery of new types of metabolites were considered to be an advantage of this approach. In addition, the use of state-of-the-art metabolomics tools could contribute to an increased level of information, and could be cost efficient if globally accepted. The annotation work, e.g. identification of peaks based on mass spectra, as well as identification of low-level metabolites, might have positive implications for the wider scientific community. Finally, the use of multivariate statistical approaches is perceived as being more robust, as it helps to take more factors into account than the current univariate approach.

The group considered that analysis of big datasets generated by metabolomics experiments, in particular data processing before the actual statistical analysis, adds a challenge and is therefore a potential disadvantage currently. In addition, standardised protocols (which would also benefit the scientific community at large) are currently lacking and the interpretation of bigger datasets might be more difficult than the endpoint-by-endpoint analysis.

In summary, the group was divided in its views, however approximately half of the group could see a potential benefit in the use of metabolomics profiles for the comparative assessment of GM plants. Before integration in the risk assessment frame, method standardisation and global agreement on harmonisation would be needed to be able to draw appropriate conclusions for the risk assessment. In addition, it may currently be challenging to bring metabolomics assays under good laboratory practice (GLP) or ISO protocols for quality assurance.

Scenario 2: Metabolomics analyses would be used to measure more and/or different endpoints within the current comparative approach.

A possible advantage of using metabolomics is the analysis of more and/or different endpoints and therefore provide a more holistic view (compared with endpoint-by-endpoint analysis) of pathways potentially affected by the introduction of a genetic modification. However, the selection of additional endpoints would probably be driven by the trait introduced through genetic modification. This is basically a case-by-case basis application of 'omics as described in Scenario 3 below. In addition, some participants indicated that the OECD consensus documents on the endpoints to measure per crop are already comprehensive enough for their purpose.

In summary, the group as a whole did not favour the use of metabolomics analyses as a routine addition to the current endpoint-by-endpoint analysis.

Scenario 3: Metabolomics analyses would be used to complement the current comparative endpoint analysis on a case-by-case basis depending on the nature of the genetic modification.

In Scenario 3, metabolomics tools would only be used to extend the targeted analysis of GM crops on a case-by-case, hypothesis-driven basis, hence not on a routine basis. The group discussed types of genetic modification that could possibly justify an extended metabolomics analysis, e.g. events expressing a transcription factor that may cause pleiotropic changes. One of the advantages of such an approach would be that it could also be applied to crops for which there are no set-up endpoints in OECD consensus documents. This would also allow flexibility when unexpected new findings for a particular GM crop would warrant a tailored approach. There is no need for specific risk assessment guidance considerations for global harmonisation as this is to be carried out on a case-by-case basis.

In summary, the use of 'omics tools to further enhance risk assessment on a case-by-case basis was seen as a valuable approach by the group.

3.2.4. Conclusions

The use of state-of-the-art metabolomics technologies in the risk assessment of GM plants was considered to be a potentially useful tool. However, the need for standardisation of wet laboratory methodology and analysis, together with global regulatory harmonisation were factors that need to be further explored and agreed upon. International activities towards standardisation are ongoing, e.g. at OECD. Even in light of these challenges, some advantages compared with the current endpoint-by-endpoint comparative assessment were indicated for both the sole use of metabolomics profiles in the comparative assessment of GM plants (Scenario 1) as well as the complementary use of metabolomics tools on a case-by-case basis depending on the trait introduced in the GM plant (Scenario 3). Further discussion will be needed to explore how any of these approaches could be integrated into the current GMO risk assessment frame.

3.3. The use of 'omics in human risk assessments of chemicals

3.3.1. Introduction

Over the last decade, several *in silico*, *in vitro* and *in vivo* approaches have been developed to investigate toxicokinetic (TK) and toxicodynamic (TD) processes in chemicals. Currently, a pivotal toxicological study in the appropriate test species, using endpoints from pathological or biochemical investigations, is being used to identify a reference point that is then used to derive either a health-based guidance value (e.g. ADI), a margin of safety (MoS) or a margin of exposure (MoE).

Over the last decades, 'omics technologies have increasingly matured, and are now often used to complement traditional toxicology studies. However, their translation to regulatory decision making has been much slower. 'omics technologies have been used in hazard characterisation through the application of *in vivo* toxicogenomics and toxicoproteomics to identify cancer-based and non-cancer-based reference points (RPs) (also known as points of departure) (Thomas et al., 2013; Buesen et al., 2017).

The aim of this discussion group was to critically discuss whether 'omics data can be integrated into chemical risk assessment to the extent that these can be used alone or in combination with traditional endpoints to identify a RP for hazard characterisation. Four major themes were discussed, and the consensus views of the group will be summarised below.

3.3.2. Outcome of the discussion

Discussion Point 1: Do 'omics data provide molecular readouts that are suitable and sufficiently robust to derive reference points (RP) (or points of departure) for the derivation of health-based guidance values?

The development of stable, accepted approaches for the generation, analysis and storage of 'omics data was considered important. However, the rate of development of appropriate analysis techniques, plus our understanding of the underpinning biological mechanisms and pathways, is currently lagging and represents a limiting factor to leveraging the maximal insight from 'omics datasets.

The group considered that 'omics data have been widely used for discovery and hypothesis generation, with much success. It was felt that such an approach has had three distinct effects on hazard classification:

1) The use of 'omics in investigative toxicology has generated a large amount of data on the biological effects of chemical exposure, which underpins our understanding of chemical MoA.
2) The generation of 'omics data has informed the development of adverse outcome pathways (AOPs), generally by filling mechanistic gaps in existing/putative AOPs.
3) The use of 'omics data to group chemicals (e.g. genotoxic vs non-genotoxic carcinogens) is an increasing practice in many areas.

In all cases, findings are generally supportive of discovery/development questions or regulatory submissions rather than an inherent part of the regulatory decision-making process.

One proposed advantage of 'omics approaches is their sensitivities in detecting responses to chemical exposure. This allows the detection of sublethal effect pathways that can be used to inform MoA and AOPs. However, it is important to note that this information has the potential to identify associations that may not be causative, may reflect adaptive mechanisms and/or may diverge under certain conditions. The group considered it important to incorporate phenotypic anchoring in 'omics experiments (i.e. using at least one dose/time that elicits a toxicological response that can be measured by standard methodologies).

While there is considerable evidence for the benefits of 'omics approaches for understanding toxicity in the non-regulatory arena, success within a regulatory framework is sparse. However, in comparison with previous groups that have met to discuss this topic, the group felt that progress had been made when 'omics approaches were being considered as tools that could be applied in the risk assessment process. One example provided was the transcriptomics initiative by Health Canada, in which they assessed health-based guidance values (HBGV) generated through 'omics data (see e.g. Farmahin et al., 2017). This project, if successful, could act as a watershed for a more general acceptance of 'omics data in the HBGV setting.

The group also noted that another significant outcome from 'omics investigations was biomarker prediction. The group noted that, at present, few biomarkers [e.g. kidney injury molecule-1 (KIM-1) (Yin and Wang, 2016)] had made the translation from research to regulatory acceptance. However, this may be a function of time rather than of effectiveness. Over the next few years, it will become clear if biomarkers that emerge from 'omics studies have a higher (or lower) success rate compared with traditionally used endpoints in toxicological studies.

The group felt that 'omics data could have significant benefit in understanding human variability and susceptibility, and is important to derive chemical-specific uncertainty factors.

Several gaps were identified in the current use of 'omics data:

1) Robust exploration of the 'normal' biological 'omics profile, including inter-individual variability, is required to leverage the maximum benefit from the sensitivity and data density afforded by

'omics data. This approach could be of particular value in providing confidence around negative responses to chemical challenges.

2) There is a need for OECD test guidelines to support and standardise 'omics technologies and data reporting and analyses, as well as allowing comparison with other approaches within OECD test guidelines. Recent 'omics approaches for skin sensitisation against current standards were noted in this regard (Zuang et al., 2015).

3) It is important that guidance exists to ensure robust, reproducible 'omics data analyses and setting of RPs from these. Recent European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) projects on metabolomics and transcriptomics reporting frameworks, plus the OECD reporting framework projects were all noted and significant progress has recently been achieved with the completion of metabolomics best practice and minimal reporting in the ECETOC MERIT project.[1]

4) It is important that 'omics data contain both time and dose–response data. This allows for an exploration of the development of response(s), which is an area in which 'omics can provide considerable insight. Many legacy datasets are not optimally designed due to cost/technical issues.

5) To leverage the maximum benefit from 'omics data, it is important that data continue to be made publicly available on curated, managed servers.

In summary, the group concluded that 'omics technologies should be viewed as a complementary technique, similar to *in vitro* or QSAR. As a set of approaches to provide additional data that can support the setting of HBGVs, 'omics data were seen as moving from promise to (near) reality. Over the next few years, several international initiatives should help to draw up the role of 'omics in regulatory frameworks.

Discussion Point 2: Which types of 'omics data are most suitable for use in risk assessment?

The group considered that the 'omics data easiest to assess (genomics and transcriptomics) are also those furthest from the biological phenotype. While genomics and transcriptomics can boast effective 100% coverage, proteomics and metabolomics coverage are more limited. As such, they may miss significant information that would alter the interpretation of the dataset. The balance between coverage and biological relevance is at the heart of 'omics analysis, and movement towards a multiomics approach, coupling high-coverage and high-relevance approaches (e.g. transcriptomics and metabolomics) is occurring (Huang et al., 2017). An additional advantage of metabolomic measurements was that (if the experiment was correctly set up) the chemical under study and its metabolite(s) could also be detected, providing additional information on toxicokinetics.

It was felt that two areas of research are currently receiving much attention, allowing for their potential use in the regulatory arena: whole-genome sequencing and epigenomic signatures. Together, these technologies have the potential to inform our understanding of susceptibility factors, and potentially signatures of prior chemical exposures.

The group spent some time discussing the role of 'omics data in biomarker generation, be it prognostic, effect or exposure. It was concluded that 'omics data had provided a large number of potential biomarkers, however very few of these have been validated at present. There was a general preference for a panel of biomarkers, which may or may not be related to a single pathway, but there are good examples of single biomarkers (e.g. miR122 or KIM-1).

In summary, all 'omics data sets are potentially suitable for use in risk assessment. However, as they each have their own limitations and advantages, a sensible strategy may be to use a multiomics approach.

[1] http://www.ecetoc.org/topics/standardisation-metabolomics-assays-regulatory-toxicology/

Discussion Point 3: Are 'omics data helpful to identify the human relevance of an adverse effect observed in animal bioassays?

The overwhelming view of the group was that 'omics data were helpful to identify human relevance of adverse effects observed in animal bioassays. The data richness delivered by 'omics approaches provides considerable information towards understanding mechanisms of action for chemicals. This data are useful in exploring the potential for conservation of MoAs between species, as well as understanding when this was unlikely. The identification of human-specific adverse events was more difficult, in general, but that 'omics data could add value in some instances here as well.

One potential issue with extrapolating from animal models to humans is a lack of understanding around genetic susceptibility; in standard bioassays, a single animal strain is used, and even if this strain is outbred, genetic variability is limited. The Jackson Laboratory for Genomic Medicine in the USA has recently introduced a diversity outbred mouse panel, produced through random outcross mating using a panel of 160 inbred mouse lines (Churchill et al., 2012). The resultant highly heterogeneous nature of diversity outbred mice genomes represents a new approach for studying human susceptibility in animal models.

The group considered if 'omics approaches currently added value for challenges in toxicology and risk assessment. The consensus view of the group is shown in the table below.

Table 1: Current potential of 'omics approaches to contribute towards toxicology and risk assessment

	Toxicity	**Risk assessment**
Species-specific toxicity	YES	YES
Mixture toxicity	YES	NO
Low-dose effects	YES[1]	NO
Endocrine effects	YES	NO
Nanotoxicology	YES	NO
Epidemiology	YES	YES
Grouping of chemicals	YES	YES

[1]'omics approaches were felt to be of use when there was an absence of overt phenotypic changes at low dose.

In summary, the group felt that 'omics data had already added considerable value to the understanding of chemical MoA (both desired and adverse), and the human relevance of adverse effects identified in animal bioassay. These contributions have not fully translated to the risk assessment arena at present, but the group concluded that as 'omics approaches continued to mature in the future they were likely to add value.

Discussion Point 4: What are the challenges for collecting, processing, interpretation, storing and curating large-scale 'omics data?

The group considered each of the major steps in an 'omics experiment, from data collection, through processing and interpretation to long-term storage and curation. It was felt that there are some drawn-up standards for the collection of 'omics data, and for the subsequent storage and curation of these data. However, both data processing and interpretation were now seen as challenging. This was due to a lack of standardisation in protocols for processing and interpretation, as well as the lag in biological information that is required to place the generated data in context. While it has been possible to progress towards generating standards around data collection, for example, this was much more challenging for processing and interpretation as there were many divergent views. One possible approach put forward by the group would be a dual track system: all 'omics data would be processed (and interpreted) using a reference protocol so that the generated data were directly comparable with other datasets. However, this would not prevent bespoke processing and interpretation of data to fit the specific requirements of the experiment.

The group noted that there were a number of activities being undertaken by OECD and ECETOC to support standardised curation and reporting processes. However, these were at different stages for different 'omics technologies with genomics, transcriptomics and metabolomics being further developed than proteomics. The group noted that a significant part of any standard or test guidelines was the need for appropriate performance and system performance checks.

In summary, while several aspects of 'omics experiments have agreed/developing standards, data processing and interpretation remain areas in which considerable inter-lab variability exists. This has the potential to slow translation of 'omics approaches into risk assessment, and ongoing work is attempting to address these issues.

3.3.3. Conclusions

The potential role of 'omics approaches in the human risk assessment of chemicals is a topic that has be considered on many occasions. In general, previous considerations have concluded that while 'omics approaches have contributed positively to the further understanding of chemical MoA, they were not mature enough to provide the robust, reproducible data required for their use in the regulatory decision-making processes. This group agreed with the first conclusion, noting that 'omics technologies are now a standard component of the researcher's toolkit for hypothesis generation and testing when exploring chemical MoA. However, the group noted that over the last few years there has been considerable efforts to generate consistent, stable 'omics technologies. In addition, ongoing work is to develop standards for collection, processing, interpretation, storage and curation of 'omics data (Russell et al., 2013; Buesen et al., 2017). These are significant steps towards the routine use of 'omics in human chemical risk assessment, and the group noted some examples for which this is already being applied. In conclusion, the group felt reasonable confidence that the use of 'omics approaches in human chemical risk assessment would become a significant complementary approach within the short to medium term.

3.4. The use of 'omics in environmental risk assessment

3.4.1. Introduction

From a general point of view, it can be argued that ERA employs three major groups of ERA disciplines:

1) Exposure driven approaches focusing on monitoring substances in the terrestrial or aquatic environment using analytical methods.
2) Ecosystem-driven ecology/ecotoxicology focusing on protecting the ecosystem(s) as a whole e.g. using mesocosm studies, field studies.
3) Laboratory-based ecotoxicology focusing on molecular and biochemical aspects to generate ecotoxicological and mechanistic data.

Environmental risk assessors may use one or more of these disciplines, but there is a need for cross-disciplinary learning between them. Moreover, there is a need to investigate how new technologies, like 'omics, can further support the application of these ERA disciplines used in EFSA. These aspects constitute the major goals of this break-out session.

To address how 'omics can support EFSA's environmental risk assessment, it is important to identify the relevance of ongoing environmental 'omics research for the substances/organisms and endpoints under assessment. EFSA performs ERA not only for substances (i.e. plant protection products (PPP) and feed additives (FAs)), but also for living organisms (i.e. GMOs and invasive alien species such as plant pests). For regulated products such as pesticides, feed additives and GMOs, premarket authorisation requires a detailed hazard assessment (identification and characterisation) using ecotoxicological information on the most frequent effects on mortality, reproduction and growth in relevant terrestrial and aquatic test species (e.g. bees, earth worms, *Daphnia*, fish etc.). For living organisms, the ERA also covers other types of endpoints such as spread or invasiveness.

Over the last 2 decades, several tools have been developed through research to further increase our mechanistic understanding of toxicity in both the TK and TD dimensions. In all areas of chemical risk assessment, these tools include biologically based models (e.g. physiologically based TK–TD models, dynamic energy budget models), *in silico* models (e.g. QSAR) and 'omics technologies. One of the foremost goal of 21st century ERA is to integrate these tools and relevant data in the risk assessment process for regulated products and contaminants. An historical mechanistic framework to perform such integration for both the TK and TD dimension is the MoA and the more recent AOPs and aggregate exposure pathway (AEP) (exposure and TK). At EFSA, these have received attention for

their applicability in hazard characterisation in humans (EFSA, 2014) and in ERA to depict and integrate species-specific traits when assessing endangered species (EFSA Scientific Committee, 2016). Practical tools and databases are currently available to make use of 'omics data for regulatory purposes:

- Databases of key events linked to an AOP, AOP knowledge base (https://aopkb.oecd.org), AOP wiki (https://aopwiki.org), and Effectopedia (https://www.effectopedia.org/).
- Manually curated information on chemical–gene and disease–gene interaction: Comparative Toxicogenomics Database (CTD) website.

Overall, 'omics technologies generate molecular endpoints based on RNA, protein and metabolite levels detected in tissues or in body fluids that represent sublethal effects (EFSA, 2014). In ERA, mortality, growth or reproduction data are most often used to derive environmental standards for species of ecological relevance (i.e. RPs/point of departure/reference values) or the whole ecosystem (e.g. species sensitivity distributions). We need to assess how to use 'omic sublethal effects to derive environmental standards and to evaluate their advantages and weaknesses over conventional data. Examples of application of 'omics in ERA include:

- Identification of potential AOP molecular initiating events and supportive evidence of key events at different levels of biological organisation and across taxonomic groups (Brockmeier et al., 2017).
- Transcriptomics and increasing numbers of available genomes for non-model species allow testing effects of environmental contaminants on species quite inexpensively (Marjan et al., 2017). The transcriptome of the model species *Daphnia magna* (by RNA-seq) was assessed upon exposure to a range of environmental perturbations (Orsini et al., 2016). Proteomic biomarkers were used to assess water pollutants in fish (Roland et al., 2016). Metabolomics were used, for example, for bio-monitoring the health status of environmental compartments (Davis et al., 2016).
- Monitoring biodiversity through environmental DNA (eDNA) that can be sampled and analysed using genomic technologies, including high-throughput sequencing, to monitor the genetic presence of species (Lodge et al., 2012). Research is moving from species identification (targeted genome sequencing) to full community profiles (metagenetics, using phylogenetic or functional marker genes) and understanding functional biodiversity and the dynamics of ecosystem functioning (metagenomics).
- Use of genomics data to support biological read-across – the web-based Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool was developed to use *in silico* molecular docking to identify the potential of contaminants to affect non-model sensitive species (http://actor.epa.gov/dashboard; Lalone et al., 2016).

The current situation at EFSA is that little or no 'omics data are being applied in ERA today for PPP, FA and invasive alien species and, to a limited extent, for GMOs. The aim of this discussion group is to critically discuss steps towards their application in such prospective ERAs.

3.4.2. Outcome of the discussion

'omics tools to deliver explicit protection goals

The group felt that it was important to articulate first that when considering the application of 'omics in ERA the discussion should focus explicitly on delivery of the explicit protection goals embodied in European legalisation. Fortuitously, the UK Department for Environment, Food & Rural Affairs (DEFRA) lead Centre of Excellence (CoE) for DNA-based applications convened in May 2018 a workshop involving the agencies involved in supporting UK and international environmental legislation, for example Water Framework Directive (WFD), Marine Strategy Framework Directive (MSFD), Natura 2000, Habitats Directive (HD), Environmental Impact Assessment (EIA) and Invasive Non-Native Species (INNS) requirements. The direct application of DNA-based tools was investigated further through a survey that engages a wider group of researchers and end-users involved in environmental protection (n = 23). These groups were asked to comment on when they were actively exploiting DNA and 'omics tools to environmental protection goals. Of the organisations surveyed, all were using DNA

tools in four broad headings: ecosystems and biodiversity; animal and plant health; food safety; and environmental pollution. The group clearly identified the priorities, benefits and barriers to exploiting DNA-based methods for environmental and ecological assessment (personal communication Dr Andrew Nesbit, Natural England). The largest use was in ecosystems and biodiversity, in which these technologies were being deployed in species detection, ecosystem/bio-assessment, non-native species monitoring as well as population and evolutionary genetics. In the area of animal and plant health, DNA-based approaches were being used for the detection of plant/aquatic pathogens, pest surveillance and animal health. While in food safety the major applications were in food authenticity and labelling and detection of human pathogens. The agencies involved in environmental pollution were mainly exploiting the approaches in the areas of microbial tracking and impact studies. The specific techniques used ranged from single molecular endpoint assays through to application metagenomic profiling and, although it showed that not all protection goals are exploiting DNA and 'omics tools, it does indicate and increasing prevalence of their application.

How to use molecular endpoints from 'omics data (transcriptomics, proteomics, metabolomics) in environmental hazard assessment

The ability to holistically measure the response of an entire suite of transcripts, proteins or metabolites in response to an environmental hazard and then select specific endpoint(s), which can be used robustly in the pursuit of environmental protection, is challenging. Despite the technical developments within the platforms used for 'omics analysis, which have improved measurement accuracy, the intrinsic intra-individual variation and interaction between normal animal physiology and stress-response pathways results in single endpoint 'biomarkers' having limited utility. However, the efficacy is substantially improved when suites of endpoints associated with functionally linked pathways are used. Pathway enrichment analysis (sometimes referred to as pathway bias analysis) is routinely integrated into 'omics pipelines to determine if there is a statistical enrichment in molecular objects associated with specific pathways, an approach that is more robust than individual markers. The logical progression of using biological pathways to improve the effectiveness of 'omics data are to develop specific pathways associated with toxicological responses, the AOPs associated with specific MoA. By anchoring endpoints within AOPs a direct quantitative link is set up between the molecular endpoint and the adverse outcome, measured at the level of the individual or population, while also providing the opportunity to gain additional assurance through considering multiple endpoints point along the AOP. Secondary advantages of exploiting AOPs are the ability to map 'shared biology' onto the schemes and therefore supporting predictive biology and driving species extrapolation. Experimental approaches that are designed to address specific hypothesis-driven analysis associated with biological functions such as reproduction or molecular function, such as DNA damage, support targeted investigations. Emergent analysis can reveal early warnings of longer term effects, for example current concerns associated with epigenetic, pollinator behaviour or innate immune effects, can be used to indicate where further analysis is required.

Use of metabarcoding and metagenomics in environmental hazard assessment

The group discussed 'omics approaches to characterising communities and microbial community function. The ability to generate quantitative community profiles of microbial or micro-/macro-invertebrates using high-throughput DNA 'barcodes' is being supported by the expansion of reference databases and the reduction in the cost of next-generation sequencing (NGS). More expansive studies quantifying microbial function, from nitrogen cycle to the presence of AMR, through to the application of metagenomic approaches provide the ability of these data to assist in the evaluation of critical ecosystem services and hazards. Metabarcoding techniques have two levels of utility; they can be used to generate diversity or richness metrics without direct species assignment by employing operations taxonomic units (OTUs) measures. The utility of diversity measures based on OTUs is mainly limited to comparative applications, ideally to assess temporal changes within a specific site (during treatment or when monitoring recovery after mitigation), when the control site is in close proximity or when the site extremely well matched apart from a specific factor, i.e. presence of a pollutant or agrochemical. As soon as species or taxa level assignment can be applied to the metabarcoding data then an ecologically informed evaluation can be performed to give a more informed assessment of ecological process and evaluate good environmental health. Recent deployment of DNA barcoding for diatom biodiversity analysis by the UK Environment Agency to support the water framework directly has delivered a tool that is consistent with a microscopic derived

diversity measure, but is also more cost effective and has a lower turn-around time. However, while reference databases for freshwater eukaryotic microbiota and invertebrates are growing rapidly, this is not repeated for the marine environment, especially marine algae and plankton, where the reference database is extremely poorly populated. The recently announced Earth Bio-genome Project (https://www.earthbiogenome.org/) aims to sequence 80% of all eukaryotes within 10 years and will make significantly increase the power of these analytical tools.

Metagenomic analysis followed by appropriate annotation can be used to evaluate the functional potential of soils, a process that has been significantly enhanced through academic projects such as the Earth Microbiome Project (http://www.earthmicrobiome.org/), which provides a comprehensive suite of standard operating procedures (SOPs) for extraction and data analysis. Although the information gained is pertinent and informative for assessing ecosystem function, this information will be substantially enhanced as basic research provides more robust annotations for key bio-geochemical processes and pathways of interest.

As with other 'omics tools these approaches benefit substantially from being anchored to mechanistic frameworks such as AOPs or MoAs, especially when addressing hypothesis-driven analysis, i.e. does this treatment reduce the nitrification potential of soil. Although emergent analysis is possible, this requires substantive expertise and only provides an indicator that further investigation is needed.

How would environmental standards derived from conventional mortality, growth and reproduction data or biochemical endpoints compare to those derived from 'omics data

While statistical association of 'omics profiling may be exploited as research tools to annotate orphan (unannotated) genes and increase biological understanding in the context of hazard assessment, endpoints derived from 'omics data should have a clear mechanistic link to the conventional life history metrics when used as a surrogate. Again, biological process ontological assignments, especially when supported by AOP or MoA annotation, provide ideal frameworks to support this association. Unfortunately, the number of AOPs that are comprehensively annotated is low (https://aopwiki.org/), especially for environmentally relevant organisms, invertebrates such as arthropods being especially neglected but ecologically significant.

The group also felt that the conventional reliance on mortality, growth and reproduction ignores emergent areas and that there is some justification for challenging this model, so that immune function, epigenetic effects and maternal and transgenerational hazards would be considered.

What are the advantages over conventional endpoints bearing in mind costs, complexity of experiments and 3Rs considerations

A significant benefit of 'omics data is that it provides a mechanistically based framework to link *in vitro* screening, fish embryonic tests and invertebrate screens within a tiered testing risk assessment, delivering the core 3Rs principles. The US Toxcast program (https://www.epa.gov/chemical-research/toxcast-dashboard) provides an excellent example of how these approaches can be used to evaluate specific areas, such as endocrine disruption chemicals (EDCs). However, as with the other parts of this discussion, it was emphasised that these benefits will only be fully implemented when more comprehensive AOP/MoA are developed that mirror currently existing EDC resources. Another area that has seen significant development is the TK and TD basis of species sensitivity that may support evidence based on species extrapolation and identification of specifically 'at risk' species. These techniques exploit comparative genomic data to assess the conservation of AOP, with emphasis on the set-up molecular initiating event (MIE) in sentinel species and predicting TD changes by considering conservation of known receptors that act as MIEs. These approaches may eventually deliver *in silico* models that would be useful to predict read-across and highlight specific mixtures that justify further experimental investigations.

How can we implement the use of 'omics in prospective ERA

Considering the current state of the art in 'omics technology together, combined with the limited availability of rigorously supported data, the group felt that at present the immediate application for functional 'omics (transcriptomics, proteomics and metabolomics) was within *in vitro* prospective screening for specific functional pathways. Given a drafted AOP or MoA for a compound/group of compounds, comparative genomics could be used to evaluate the validity an appropriate surrogate organism and be used to predict cross-species toxicology and identify sensitive species when

appropriate data were available. When explicit environmental scenarios need to be tested, explicit data on the chemical properties and phylo-toxicity of a compound can be used to design 'omics investigations to evaluate hazard. Non-targeted metabolomics analysis can be used to support exposure analysis to develop rapid biomarker profiles to assess resilience, while orthogonal 'omics datasets from complementary sentinel species can be used to identify conserved response pathways. These approaches will allow the gap to be closed between 'what we can measure and what we want to protect', increase confidence and translate 'omics from 'predictive' to 'indicative'.

Can 'omics help to identify sensitive species and or endangered species when standard ecotoxicology tests fail (for impact assessment of PPP, FAs, GMOs and of alien species)?

Exploiting 'omics to support *in silico* TK/TD prediction given drafted AOPs and comparative genomic information may allow a level of prediction for sensitive species, even if they are endangered and direct testing is prohibited. However, where 'omics can deliver immediately is the measurement of biodiversity using metabarcoding (see above), which can assist a more ecosystem centric approach to determining the effects of PPP, FAs, GMOs and alien species. The issue of predicting any secondary effects of RNAi embedded within GMOs was raised. Although the *in silico* prediction of potential binding sites has improved substantively, the prediction would only be valid when combined with experimental transcriptomic data to substantiate the predicted interactions.

What are the current challenges for collecting, processing and curating 'omics data? What are the data gaps and future needs?

Challenges and needs: There is still substantial need to acquire much larger qualitative data sets to support the classification or hazard and effective dose modelling to derive environmental standards. Pharmaceutical toxicology, in which 'omics data are used extensively by industry to remove compounds with potential harmful toxic side-effects, has exploited training sets of thousands of compounds in an endeavour that has not been replicated for ERA. Essential to the deployment of 'omics with ERA is the development of high-quality reference datasets to describe natural and physiologically based variation.

Requirement: It was felt that improvements are needed to enhance the accuracy of data generation and reproducibility of the analytical pipelines. Key to this is the development of data standards and repositories for 'ecotoxicological' data linked to 'omics responses. Data packages that more easily navigate between chemical structures, biological pathways and life history outcomes would greatly assist the transparency of applying these tools. Ensuring data address current issues (prospective) rather than focusing on historical challenges is also essential.

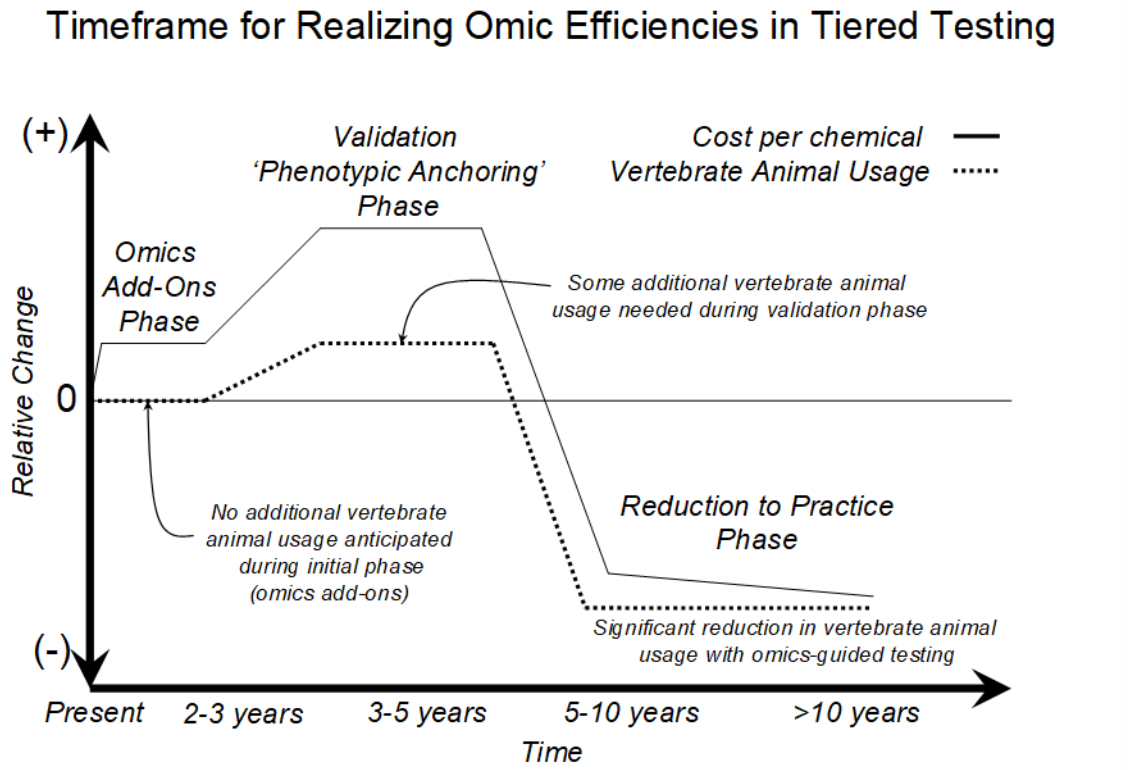
Key to delivery is resolving the regulatory demand loop – investment to deliver the need and requirements will only arise given a regulatory requirement / acceptance of the data.

3.4.3. Conclusion

What are the scientific and regulatory opportunities and barriers to implement 'omics in prospective ERA? This discussion point provided an excellent vehicle to conclude these discussion group deliberations. It was agreed that 'omics could assist in moving ERA towards Precision Environmental Health in which an evidence-based risk assessment is based on mechanistic understanding informing prediction leading to stratified tiered testing. This would require regulatory drive, new expertise in risk assessment and the comparative implementation alongside current approaches to build confidence when delivering structural information to support a weight of evidence of the validity of the approach and the value added that stemmed from 'omics.

The barriers reside in the complexity of the data and its analysis combined with the fact that 'omics focus on sublethal endpoints rather than lethality. It may therefore superficially be perceived a threat given that the lowest observed effect concentration (LOEC) and NOEC 'omics levels could be substantially lower than those used for conventional methods. Increased confidence will allow the precautionary factors to be reduced. Much of the concern about 'omics will be relieved given the transparent availability of associated (eco)toxicology data. However, by far the greatest effect in adaption will stem from the development of robust and well supported AOPs and MoAs that link 'omics changes through functional pathways to specific hazard. This requires a concerted effort to develop a

confidence framework for the use of 'omics data. A timeframe for this development is described previously (see Figure 1).



Timeframe for implementing 'omics efficiencies in tiered testing (from Tyler et al., 2007 © TAYLOR & FRANCIS GROUP LLC)

Figure 1: Final plenary discussion

The final discussion focused on issues arising from the reporting back on the outcome of the four discussion groups. Some general questions were asked about the current status of validation and standardisation of the 'omics methodology and the need of it before integration in the risk assessment. The Chairs and Rapporteurs of the different discussion groups acknowledged the need of method validation and standardisation. Participants flagged up several international initiatives currently running for method validation and standardisation and integration of 'omics in different parts of the risk assessment.

3.5. Discussion on the outcome of Discussion Group 1

The Chair and Rapporteur of Discussion Group 1 reported back the outcome of their discussion. A question was asked on whether WGS could be useful to detect genetic modifications that are not insertions (such as deletions, duplications, etc.). The discussion group clarified that detecting unreported genetic modifications by WGS has its limitations. Rather than to detect unreported genetic modifications, the group considered WGS as a tool to better characterise reported genetic modifications in regulatory risk assessments, with particular usefulness in the detection of possible AMR marker genes left out in the genome, and in the assessment of the current modifications made. (including deletions and duplications). At present, this assessment must be carried out by a step-by-step reconstruction of all the modification steps performed, with the need for experimental data to evaluate the outcome of each of these. A WGS approach would simplify and add accuracy to the assessment.

3.6. Discussion on the outcome of Discussion Group 2

The Chair of Discussion Group 2 reported back the outcome of the discussion. A question was asked on whether allergens would be measured in the targeted approach. The discussion group clarified that the measurement of allergens could fit under Scenario 3, for which case-by-case basis additional components could be measured (e.g. for soybean containing several well-known food allergens). Another participant asked whether the proposed scenarios will be implemented and whether this can be carried out in a cost-effective way. The discussion group answered that it depends on the way the trials would be set up for these experiments and on the global acceptance of the method. An alternative field trial approach was presented as well, in which a possible unintended effect would be linked to genetic modification by comparing the genetic modification in different genetic backgrounds with their non-GM comparators. Another questioner asked if the regulatory community had missed something in risk assessment by using current OECD endpoints and what the motivation was to have different mandatory requirements. The Chair of the discussion group replied that metabolomics might add to the current risk assessment by providing information on multiple components in a pathway. If a metabolomics profile of a GM plant would be created, the individual measurement would only be further investigated if the profile did not fall within the range of those of the reference varieties.

3.7. Discussion on the outcome of Discussion Group 3

The Rapporteur of Discussion Group 3 reported back the outcome of the discussion. A question was asked on what was meant when saying that all metabolites can be measured. Indeed, metabolomics has the power not only to measure changes in cellular metabolism but also to identify the metabolites formed from the substance tested. Several questions on the validation aspects of 'omics were also raised. The Rapporteur of the discussion group replied that progress is rapidly being made towards generating standards around data collection. However, processing and interpretation of data remained challenging, as there were many divergent views.

3.8. Discussion on the outcome of Discussion Group 4

The Rapporteur of Discussion Group 4 reported back the outcome of the discussion. A participant said that political drive is needed for 'omics to be integrated in the risk assessment. Indeed, the discussion group acknowledged that regulatory requirements would stimulate the use of 'omics in the ERA and lead to acceptance of the use of such data.

4. Overall conclusions

Over the last decade, further advancement in technologies to obtain large molecular datasets ('omics datasets) from an organism has aided obtaining more detailed information on the functional status of an organism and the effect of external factors on that organism. The 24th EFSA Scientific Colloquium aimed at exploring the current status of the use of 'omics technologies in the risk assessment of food and feed products and the environment and the future steps to take to integrate 'omics in that risk assessment.

From the Colloquium it became clear that 'omics technologies are a valuable addition to some aspects of risk assessment of food and feed products and the environment, especially now that the technology is almost mature and stable. Even though several national and international organisations have already started with different activities to standardise 'omics methodology, analysis and reporting, a more concerted effort is needed before these technologies can be stably integrated into risk assessment. Consistent frameworks for data collection, processing, interpretation, storage and curation should be further set up before 'omics technologies can be routinely used in risk assessment.

To start using 'omics datasets for human risk assessment of chemicals and ERA, a confidence framework for the use of 'omics data needs to be built. The regulatory implementation would require regulatory drive, new risk assessment expertise and the initial use of comparative approaches in which 'omics datasets would be used alongside current toxicological/ERA approaches to build confidence in 'omics datasets used for the risk assessment. Therefore, currently, 'omics approaches should generally be seen as a complementary approach rather than a replacement for current methodologies until enough confidence has been built. The group concluded that test cases could be worked out to enhance confidence in the use of 'omics datasets in risk assessment.

References

Brockmeier EK, Hodges G, Hutchinson TH, Butler E, Hecker M, Tollefsen KE, García-Reyero N, Kille P, Becker D, Chipman K, Colbourne J, Collette TW, Cossins A, Cronin M, Graystock P, Gutsell S, Knapen D, Katsiadaki I, Lange A, Marshall S, Owen SF, Perkins EJ, Plaistow S, Schroeder A, Taylor D, Viant M, Ankley G and Falciani F, 2017. The Role of omics in the application of adverse outcome pathways for chemical risk assessment. Toxicological Sciences, 158(2), 252–262.

Buesen R, Chorley BN, Lima BD, Daston G, Deferme L, Ebbels T, Gant TW, Goetz A, Greally J, Gribaldo L, Hackermuller J, Hubesch B, Jennen D, Johnson K, Kanno J, Kauffmann HM, Laffont M, McMullen P, Meehan R, Pemberton M, Perdichizzi S, Piersma AH, Sauer UG, Schmidt K, Seitz H, Sumida K, Tollefsen KE, Tong WD, Tralau T, van Ravenzwaay B, Weber RJM, Worth A, Yauk C and Poole A, 2017. Applying 'omics technologies in chemicals risk assessment: report of an ECETOC workshop. Regulatory Toxicology and Pharmacology, 91, S3–S13.

Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S and Trujillo ME, 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. International Journal of Systematic and Evolutionary Microbiology, 68, 461–466.

Churchill GA, Gatti DM, Munger SC and Svenson KL, 2012. The diversity outbred mouse population. Mammalian Genome, 23, 713–718.

Davis JM, Ekman DR, Teng Q, Ankley GT, Berninger JP, Cavallin JE, Jensen KM, Kahl MD, Schroeder AL, Villeneuve DL, Jorgenson ZG, Lee KE and Collette TW, 2016. Use of Linking field-based metabolomics and chemical analyses to prioritize contaminants of emerging concern in the Great Lakes basin. Environmental Toxicology and Chemistry, 35, 2493–2502. doi: 10.1002/etc.3409

EFSA (European Food Safety Authority), 2014. Modern methodologies and tools for human hazard assessment of chemicals. EFSA Journal 2014;12(4):3638, 87 pp. doi:10.2903/j.efsa.2014.3638

EFSA (European Food Safety Authority), 2015. EFSA's 20th Scientific Colloquium on whole-genome sequencing of food-borne pathogens for public health protection. EFSA Supporting Publication 2015;12(2):EN-743, 63 pp. doi: 10.2903/sp.efsa.2015.EN-743

EFSA (European Food Safety Authority), García Fierro R, Thomas-Lopez D, Deserio D, Liebana E, Rizzi V and Guerra B, 2018. Outcome of EC/EFSA questionnaire (2016) on use of whole-genome sequencing (WGS) for food- and waterborne pathogens isolated from animals, food, feed and related environmental samples in EU/EFTA countries. EFSA supporting publication 2018:EN-1432, 49 pp. doi:10.2903/sp.efsa.2018.EN-1432

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2010. Scientific opinion on statistical considerations for the safety evaluation of GMOs. EFSA Journal 2010;8(1):1250, 59 pp. doi:10.2903/j.efsa.2010.1250

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2011. Guidance document on selection of comparators for the risk assessment of GM plants. EFSA Journal 9(5):2149, 21 pp. doi: 10.2903/j.efsa.2011.2149

EFSA NDA Panel (EFSA Panel on Dietetic Products, Nutrition and Allergies), Turck D, Bresson J-L, Burlingame B, Dean T, Fairweather-Tait S, Heinonen M, Hirsch-Ernst KI, Mangelsdorf I, McArdle H, Naska A, Neuhauser-Berthold M, Nowicka G, Pentieva K, Sanz Y, Siani A, Sjödin A, Stern M, Tom_e D, Vinceti M, Willatts P, Engel K-H, Marchelli R, Poeting A, Poulsen M, Salminen S, Schlatter J, Arcella D, Gelbmann W, de Sesmaisons-Lecarr_e A, Verhagen H and van Loveren H, 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 Journal 2016;14(11):4594, 24 pp. doi:10.2903/j.efsa.2016.4594

EFSA Scientific Committee, 2016. Scientific opinion on coverage of endangered species in environmental risk assessments at EFSA. EFSA Journal 2016;14(2):4312, 124 pp. doi:10.2903/j.efsa.2016.4312

Farmahin R, Williams A, Kuo B, Chepelev NL, Thomas RS, Barton-Maclaren TS, Curran IH, Nong A, Wade MG and Yauk CL, 2017. Recommended approaches in the application of toxicogenomics to derive points of departure for chemical risk assessment. Archives of Toxicology, 91, 2045–2065.

Hegeman AD, 2010. Plant metabolomics—meeting the analytical challenges of comprehensive metabolite analysis. Briefings in Functional Genomics, 9(2), 139–148.

Huang S, Chaudhary K and Garmire LX, 2017. More is better: recent progress in multi-omics data integration methods. Frontiers in Genetics, 8, 84,

LaLone CA, Villeneuve D L, Lyons D,, Helgen H W, Robinson S L, Swintek J A, Saari TW and Ankley GT, 2016. Editor's Highlight: Sequence alignment to predict across species susceptibility (SeqAPASS): a web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. Toxicological Sciences, 153, 228–245.

Lodge DM., Turner CR, Jerde CL, Barnes MA, Chadderton L, Egan SP, Feder JL, Mahon AR and Pfrender ME, 2012. Conservation in a cup of water: estimating biodiversity and population abundance from environmental DNA. Molecular Ecology, 21, 2555–2558.

Marjan P, Bragg LM, MacLatchy DL, Servos MR and Martyniuk CJ, 2017. How does reference site selection influence interpretation of omics data? Evaluating liver transcriptome responses in male rainbow darter (*Etheostoma caeruleum*) across an urban environment. Environmental Science and Technology, 51, 6470–6479.

Markley JL, Brüschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D and Wishart DS, 2017. The future of NMR-based metabolomics. Current Opinion in Biotechnology, 43, 34–40.

Orsini L, Gilbert D, Podicheti R, Jansen M, Brown JB, Solari OS, Spanier KI, Colbourne JK, Rusch DB, Decaestecker E, Asselman J, De Schamphelaere KA, Ebert D, Haag CR, Kvist J, Laforsch C, Petrusek A, Beckerman AP, Little TJ, Chaturvedi A, Pfrender ME, De Meester L and Frilander MJ, 2016. *Daphnia magna* transcriptome by RNA-seq across 12 environmental stressors. Scientific Data, 4, 170006.

Roland K, Kestemont P, Dieu M, Raes M and Silvestre F, 2016. Using a novel "Integrated Biomarker Proteomic" index to assess the effects of freshwater pollutants in European eel peripheral blood mononuclear cells. Journal of Proteomics, 137, 83–96.

Russell C, Rahman A and Mohammed AR, 2013. Application of genomics, proteomics and metabolomics in drug discovery, development and clinic. Therapeutic Delivery, 4, 395–413.

Saito K and Matsuda F, 2010. Metabolomics for functional genomics, systems biology, and biotechnology. Annual Review of Plant Biology, 61, 463–489.

Tyler C, Filby A, Iguchi T, Kramer V, Larsson J, van Aggelen G, van Leeuwen K, Viant M and Tillet D, 2007. Application of genomics to tiered testing. In: Ankley G and Daston G (eds). Genomics in Regulatory Ecotoxicology – Applications and Challenges. CRC Press, Taylor and Francis. pp. 33–62.

Thomas RS, Wesselkamper SC, Wang NC, Zhao QJ, Petersen DD, Lambert JC, Cote I, Yang L, Healy E, Black MB, Clewell HJ 3rd, Allen BC and Andersen ME, 2013. Temporal concordance between apical and transcriptional points of departure for chemical risk assessment. Toxicological Sciences, 134, 180–194.

van Dijk JP, de Mello CS, Voorhuijzen MM, Hutten RCB, Maisonnave Arisi AC, Jansen JJ, Buydens LMC, van der Voet H and Kok EJ, 2014. Safety assessment of plant varieties using transcriptomics profiling and a one-class classifier. Regulatory Toxicology and Pharmacology 70, 297–303.

Yin C and Wang N, 2016. Kidney injury molecule-1 in kidney disease. Renal Failure, 38, 1567–1573.

Zuang V, Desprez B, Barroso J, Belz S, Berggren E, Bernasconi C, Bessems J, Bopp S, Casati S, Coecke S, Corvi R, Dumont C, Gouliarmou V, Griesinger C, Halder M, Janusch-Roi A, Kienzler A, Landesmann B, Madia F, Milcamps A, Munn S, Price A, Prieto P, Schäffer M, Triebe J, Wittwehr C, Worth A and Whelan M, 2015. EURL ECVAM status report on the development, validation and regulatory acceptance of alternative methods and approaches (2015). Joint Research Centre (JRC) Science and Policy Report. JRC97811-EUR27474 EN. https://ec.europa.eu/jrc/en/eurl/ecvam

Abbreviations

AEP	Aggregate exposure pathway
AMR	Antimicrobial resistance
ANI	Average nucleotide identity
AOP	Adverse Outcome pathway
CoE	Centre of Excellence
CTD	Comparative Toxicogenomics Database
DEFRA	Department for Environment, Food & Rural Affairs
ECETOC	European Centre for Ecotoxicology and Toxicology of Chemicals
EFSA	European Food Safety Authority
EIA	Environmental Impact Assessment
ERA	Environmental Risk Assessment
GLP	Good laboratory practice
GMO	Genetically modified organisms
HBGV	Health-based guidance values
HD	Habitats Directive
INNS	Invasive Non-Native Species
JRC	Joint Research Centre
KM	Kahl MD
LOEC	Lowest observed effect concentration
MoA	Mode of action
MoE	Margin of exposure
MoS	Margin of safety
MSFD	Marine Strategy Framework Directive
NGS	Next-generation sequencing
PCA	Principal component analysis
PPP	Plant protection products
QPS	Qualified Presumption of Safety
QSAR	Quantitative Structure–Activity Relationship
RP	Reference Points
SeqAPASS	Sequence Alignment to Predict Across Species Susceptibility
WFD	Water Framework Directive
WGS	Whole-genome sequencing

Annex A – Programme of the Colloquium

Overall Chair

Howard Davies, *Honorary Research Fellow*, *James Hutton Institute*, *Scotland*

Overall Rapporteur

Matthew Ramon, *European Food Safety Authority*, *Italy*

Tuesday, 24 April 2018

08:30–09:00 Registration participants

09:00–13:00 INTRODUCTORY PLENARY SESSION

09:00 Welcome and introduction to the Colloquium

Howard Davies, *Honorary Research Fellow*, *James Hutton Institute*, *Scotland*

09:15 Objectives of the Colloquium

Matthew Ramon*, European Food Safety Authority*, *Italy*

09:30 Taxonomic and toxicogenic potential derived from whole-genome sequencing (WGS) information

María José Figueras*, University Rovira i Virgili*, *Pere Virgili Health Research Institute*, *Spain*

10:10 Mechanistic modelling of metabolism to understand and predict plant performance: a case study in growing tomato fruits

Yves Gibon*, INRA Bordeaux*, *France*

10:50 COFFEE BREAK

11:20 The value of 'omics to chemical risk assessment

Tim Gant*, Public Health England*, *UK*

12:00 Views on the use of genomic and proteomic information in environmental risk assessment

Christer Hogstrand*, King's College London*, *UK*

12:40 Introduction to Discussion Groups

13:00–14:30 LUNCH

14:30–18:30 DISCUSSION GROUPS (DG)

DG 1 – Genomics for the identification and characterisation of microbial strains used in food and feed products

Introduction to the EFSA activities and initiatives on next-generation sequencing

Jaime Aguilera*, European Food Safety Authority*, *Italy*

Chair: Lutz Grohmann*, BVL*, *Germany*

Co-chair: Jaime Aguilera*, European Food Safety Authority*, *Italy*

Rapporteurs: Lieve Herman*, ILVO*, *Belgium*

Pier Sandro Cocconcelli*, Universitá Cattolica del Sacro Cuore*, *Italy*

DG 2 – The use of metabolomics in the comparative risk assessment of GM plants

Introductory talk: Introduction to the one-class model to classify 'omics profiles

Esther Kok, *RIKILT*, *The Netherlands*

Chair: Elisabeth Waigmann*, European Food Safety Authority*, *Italy*

Co-chair: Fabien Nogué, *INRA*, *France*

Rapporteurs: Gijs Kleter, *RIKILT*, *The Netherlands*

Federica Barrucci*, European Food Safety Authority*, *Italy*

DG 3 – The use of 'omics in human risk assessment of chemicals

Introductory talk

Georges Kass, *European Food Safety Authority*, *Italy*

Chair: Matt Wright*, Newcastle University*, *UK*

Co-chair: Georges Kass*, European Food Safety Authority*, *Italy*

Rapporteur: Nick Plant*, University of Leeds*, *UK*

DG 4 – The use of 'omics in environmental risk assessment

Introductory talk,

Reinhilde Schoonjans, *European Food Safety Authority*, *Italy*

Chair: Christer Hogstrand*, King's College London*, *UK*

Co-chair: Reinhilde Schoonjans*, European Food Safety Authority*, *Italy*

Rapporteurs: Nancy Denslow*, University of Florida*, *USA*

Peter Kille*, Cardiff University*, *UK*

16:30 COFFEE BREAK

19:30 Networking cocktail

Wednesday, 25 April 2018

09:00–10:00 CONTINUATION OF DISCUSSION GROUPS

Focus on summarising the main outcome, possible challenges, need for guidance and production of reports to the plenary session

10:00 COFFEE BREAK

10:30–13:20 FINAL PLENARY SESSION

10:30 Report back from DG 1 and discussion

Lieve Herman*, ILVO*, *Belgium*

Pier Sandro Cocconcelli, *Universitá Cattolica del Sacro Cuore*, *Italy*

11:05 Report back from DG 2 and discussion

Elisabeth Waigmann*, European Food Safety Authority*. *Italy*

Fabien Nogué*, INRA Versailles*, *France*

11:40 Report back from DG 3 and discussion

Nick Plant*, University of Leeds*, *UK*

Georges Kass*, European Food Safety Authority*, *Italy*

12:15 Report back from DG 4 and discussion

Nancy Denslow*, University of Florida*, *USA*

Peter Kille*, Cardiff University*, *UK*

12:50 Take-home messages

Howard Davies*, Honorary Research Fellow*, *James Hutton Institute*, *Scotland*

13:20 COLLOQUIUM ADJOURNS