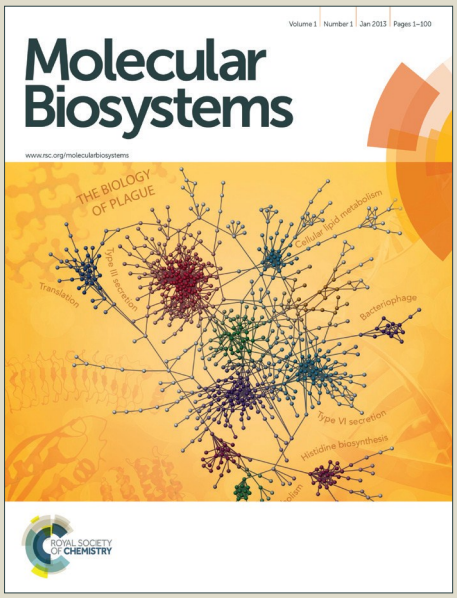
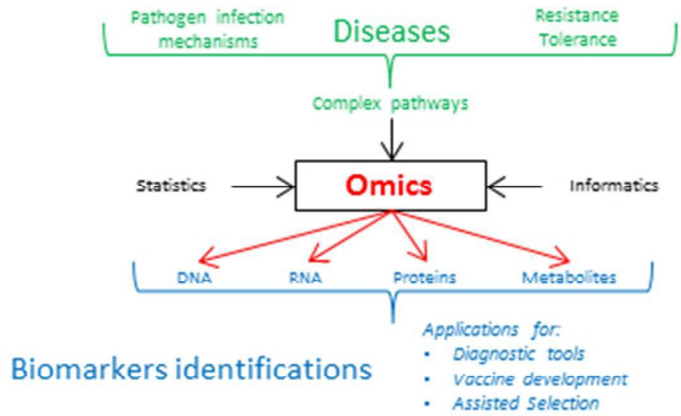
Molecular BioSystems

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: N. Guillemin, A. Horvati, J. Kuleš, A. Galan, V. Mrljak and M. Bhide, *Mol. BioSyst.*, 2016, DOI: 10.1039/C6MB00220J.



View Article Online
DOI: 10.1039/C6MB00220J



Diseases pathways can be explained into a list of biomarkers at different scales to develop applications

79x40mm (150 x 150 DPI)

View Article Online
DOI: 10.1039/C6MB00220J

Molecular BioSystems

REVIEW

Omics approaches to probe markers of disease resistance in animal sciences

Received 24th March 2016,
Accepted 27th April 2016

DOI: 10.1039/x0xx00000x

www.rsc.org/molecularbiosystems

N. Guillemin[a], A. Horvatić[a], J. Kuleš[a], A. Galan[a], Vladimir Mrljak[a], M. Bhide[a,b]

Omics technologies have been developed since decades and used in different thematics. More advancements were done in human and plants thematics. Omics is the conjugation of different techniques, studing all biological molecules (DNA, RNA, proteins, metabolites, *etc*.). Omics is then able to study entire pathways, elucidating phenotypes and their control. Thus, thanks to Omics, it is possible to have a broad overview of linkage between genotype and phenotype. Disease phenotypes (tolerance or resistance) are important to understand in both in production and health. Nowadays plethora of research articles are presenting results in the field of natural disease resistance of animals using Omics technologies. Moreover, thanks to advanced highthroughput technologies novel mode of infections (infection pathways) are coming to surface. Such pathways are complex (hundreds to thousands of molecules implied, with complicated control mechanisms), and Omics can generate useful knowledge to understand those pathways. Here we aim to review several angles of Omics used to probe markers of disease resistance with recent publications and data on the field, and presents perspectives and its utilization for a better understanding of diseases.

Introduction

Omics is a collection of biomolecular exploration techniques and methodologies. The term of "Omics", derived from Greek, has been a suffix added to a type a studied molecule, and means the study of all those molecules. Presently, Omics is a stand-alone term and define a collection of techniques, protocols and methods to study all the molecular content of a cell, organ and organisms. Omics is organized into different levels (Fig. 1): genomics (study of DNA), epigenomics (study of DNA non genetic modifications), transcriptomics (study of RNA content), proteomics (study of protein content), metabolomics (study of metabolites content) and other Omics derivative like lipidomics (study of lipids contents). All those levels are in hierarchical order from genetic information to final phenotype (Fig. 1). During the last decade, those different levels of Omics have been developed and optimized. Their utilization is presently growing in a lot of different thematic, as their great importance has been demonstrated in molecular biology, notably to better understand complex pathways and phenotypes. The utilization of different Omics techniques allows researchers to study complex pathways at different molecular level, and have a broad overview of the molecular

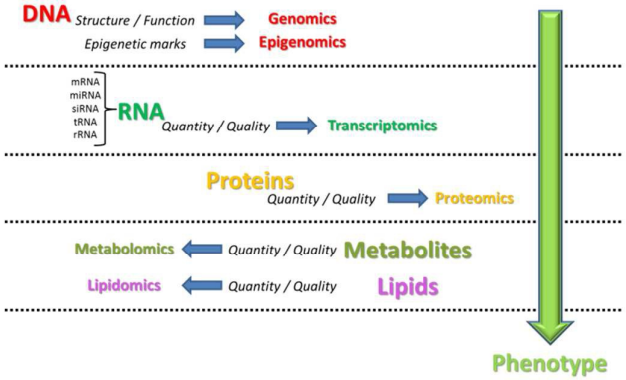


Figure 1. Different levels of Omics

mechanisms, from gene expression to the final phenotype production[1]. Biomarkers, defined as biological molecules which are quantitatively and/or qualitatively related with an observable and studied characteristic or trait, can be identified by Omics at different cellular levels. Omics allows researchers to analyses hundreds and thousands of samples at the same time. As a consequence, researchers are able to detect small but accurate differences in complex pathways mechanism and regulation.

[a] ERA Chair FP7, Internal diseases, Faculty of Veterinary Medicine, Heizelova 55, 10000 Zagreb, CROATIA.
[b] Laboratory of Biomedical Microbiology and Immunology, Department of Microbiology and Immunology, University of Veterinary Medicine and Pharmacy, Komenskeho 73, 04181, Kosice, Slovakia
[c] Electronic Supplementary Information (ESI) available: []. See DOI: 10.1039/x0xx00000x

View Article Online
DOI: 10.1039/C6MB00220J

Omics can be defined by 7 words:

- Biomolecules. Omics studies DNA, RNA, proteins, and metabolites.
- High-throughput. Omics technologies are characterized by their abilities to analyze simultaneously tens or hundreds of samples at the same time.
- Phenotype. Omics aim to make relationship between phenotype and biomolecules.
- Quantitative. Omics aim to study a quantitative status of a whole set of biomolecules.
- Qualitative. Omics aim to study a qualitative status of a whole set of biomolecules.
- Bioinformatics. Omics generate and manage a huge amount of information provided by different databases. To this purposes, Omics require dedicated bioinformatics tools, able to store and find any kind of information. High capacity physical hardware is also important.
- Biostatistics. Omics generate a tremendous amount of raw results which must be statistically analyzed, to discard false-positive and false-negative results, and find statistically significant differences. There is an immense need of biostatistics coupled with bioinformatics.

To be able to perform Omics analyses, it is necessary to have precise phenotype profiling. However this aspect is often neglected. Although the profiling of phenotype seems to be simple, in reality it is a collection of data from different traits, and can be tedious to collect whole set of phenotypic data. It can be at supra-cellular level (intramuscular fat, milk yield production, disease resistance, *etc*.), cellular level (cell count, cell morphology), and intra-cellular (mRNA or protein abundance). If the phenotyping is not well established, it can introduce technical biases, which produce statistical errors, and decrease the statistical power of Omics analysis.

Another element of importance in Omics is the studied population itself. As Omics are a collection of high-throughput, tedious and costly methods, a precise biological sampling is necessary, that can give enough statistical power. This number is highly dependent on the trait to study and can be estimated with a power calculation[2].

In this review, we will present different Omics principles available for the identification of disease resistance markers for animal sciences, highlighted by latest publications and data in the field.

Different levels of Omics

Genomics

Genomics is defined as the study of DNA structure (sequence and variations) and functions (information carried). Genomics aims to establish how DNA drives traits and phenotypes expression in organisms. By studying DNA differences between species and within a species, it is possible to understand how DNA regions encode some features and traits[3].

A DNA variation which can be easily identified is called a genetic marker. Knowledge of those markers and their possible association with traits and phenotype variation is the core for genomics studies. Two main methodologies used to identify genetic markers related with animal traits variation, like disease resistance, are: QTL (Quantitative Trait Loci) analysis and GWAS (Genome Wide Association Study). QTL is a DNA region associated with phenotype variation[4]. To identify a QTL, two lines of animals are required, with difference in their phenotype, for example resilience score for a disease. Those lines could be obtained by divergent selection. Animals in those lines must be characterized for a collection of markers among their chromosomes, to have a more complete map of DNA variations as possible. Then those animals are used to produce a first generation, F1. This generation will produce a second generation F2. Theoretically, markers will be randomly mixed among those generations. After the assessment of markers status in the F2 generation and quantification of their phenotype (in this example, disease resilience score), the QTL analysis will determine what markers are not randomly distributed according phenotype variation, i.e. markers physically associated with genes involved in phenotype control. Such markers define the QTL region associated with the phenotype (here disease resilience). As a consequence, QTL region length can be more or less important. It can be from a few nucleotides (even a single SNP) up to entire chromosome segments (up to hundreds of kilobases) containing different genes. QTL approach has been extensively used, particularly in commercial animals (cow, chicken, horse, pig, trout, and sheep) to successfully identify markers of interests for diseases related traits (Fig. 2), published in public repository http://www.animalgenome.org/cgi-bin/QTLdb/index. QTLs have been widely used for Marker Assisted Selection (MAS), consisting to increase the frequencies of favorable alleles in the population. Cows have been extensively studied, particularly for mastitis, with 1 297 QTLs only for this disease (due to the impact of cow mastitis on milk production). Further investigations for QTL identification of other disease than mastitis is required. Other species were not studied with the same intense level than cow. It is particularly the case for trout, with only 7 QTLs related with diseases. In the study of Wiens *et al.* in 2013 in trout, 3 significant QTLs related with bacterial cold water disease resistance were identified between resistance selected and divergent families[5]. Despite its importance in agriculture, sheep is poorly studied for diseases QTL, compared to cow (242 versus 1 724). The example of mastitis is particularly demonstrative (8 QTLs for sheep versus 1 297 for cow). There is so a huge gap of knowledge in diseases QTL between cows and other species, and needed to be completed. Moreover, for all species, an important trait is missing: disease tolerance (i.e. ability for an infected animal to maintain its production level, in the category "Disease susceptibility" in QTL databases). Only 6 QTLs are identified in cow. Also, there is a huge gap in QTL knowledge for this important trait.

View Article Online
DOI: 10.1039/C6MB00220J

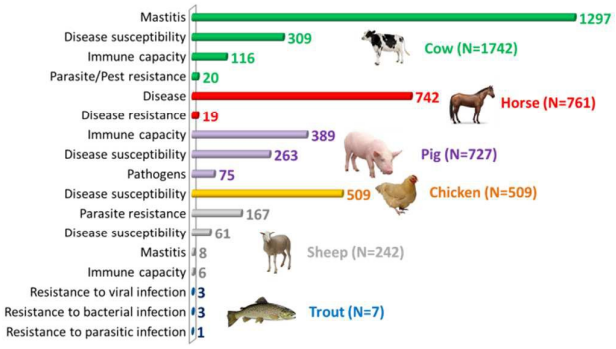


Figure 2. QTLs by disease traits and species of agronomical interest

However, QTL approach suffers from some issues. It is an expensive approach, both in terms of time and money. Due to the complex animal scheme required for QTL analysis (cross lines and 2 generations downstream), QTL experiments can be difficult to set for many laboratories with few animal husbandry abilities. Moreover, to identify more precise QTL regions, it is necessary to increase the number of animals in the QTL study design. Another issue with QTL markers is that it relies on the physical relationship between markers and genes involved in phenotype variation, called linkage disequilibrium. As meiosis mix parental DNA chromosomes, relationship between markers and QTL can be completely modified after few generations. Closest the markers are to the QTL, less probable is the recombinant event. That is why a fine QTL mapping is essential, *i.e.* identify the markers closest to the causatives genes. QTL can then be targeted by other Omics approaches to finely dissect information contained in this region. Another issue is that markers identified by QTL approaches are less effective when applied to an unrelated population.

GWAS is another tool to analyze the correlation between genetic variations all among the genome and a trait of interest in different unrelated individuals, to identify trait genetic markers[6]. GWAS typically compare two groups in a qualitative way, one with a specific property (infected cohort for example), and the other as a control (healthy cohort for example). A variation of GWAS has been introduced: instead to use two groups (1 positive, 1 control) in a qualitative way, the trait to study is considered as a quantitative trait (like molecule quantification, a gene expression profile, or an evaluation score). GWAS consist to perform a regression analysis by multi-testing all markers with the trait to study. GWAS identify the relevant and significant markers related with this trait.

GWAS is able to identify causative variations without any linkage effect (i.e. markers which directly impact on phenotype variation like SNPs in responsible genes). GWAS have a greater resolution power than linkage analysis[7]. As a consequence, GWAS is able to detect small effect markers, which are difficult to identify with linkage analysis. Considering that most of the traits variance is resulting from small effect markers, GWAS is very useful to have a broad overview of complex trait architecture[8]. GWAS have been successfully applied to find markers for diseases, both in humans and animals[9]. In humans, thanks to the huge amount of data collected in medical facilities, SNPs identification by GWAS produced a lot of data about different diseases. A recent GWAS study on 74 000 humans identified 11 new SNPs associated with Alzheimer's disease risk[10]. Advancement in human GWAS leaded to the creation of dedicated databases and tools, to collect and display GWAS results according a studied disease, and their associated articles, like GWAS catalog (http://www.ebi.ac.uk/gwas/home). This catalog highlight SNPs implied in a variety of diseases: cancer, cardiovascular, digestive, infections. Of course this catalog depends on results registration on its databases.

In animal sciences, GWAS studies have also been done on different diseases. In the study of Zare *et al.* in 2014, a GWAS has been done between 250 cows infected by *Mycobacterium paratuberculosis* and 249 control cows, with blood and fecal samples collected in different farms[9]. 2 different analytical approaches (single-marker and multiple marker regressions) have been used, to identify commonly 9 SNPs of paratuberculosis susceptibility. Many GWAS has been used for different diseases in different animal species, to identify genetic markers which enrich the QTL databases. Despite that, fewer diseases SNP markers are known in animals compared to humans, and fewer databases are available. It demonstrates the need to generate more data for diseases in animal and fill the gap with human knowledge.

The collection of available markers lead to the conception and utilization of genomic selection, instead to markers assisted selection. Genomic selection consists to use thousands of markers (with little effect) across the genome (identified by GWAS experiments), contrary to marker assisted selection (using few markers with high effect). Genomic selection is considered as more powerful than marker assisted selection[11].

The increase of genomics studies, and so markers identification, has been possible thanks to the improvements of two important techniques to identify genome variations: sequencing (determination of DNA sequences) and genotyping (determination of allelic variants). Thanks to the substantial decrease in sequencing/genotyping cost, GWAS become more and more affordable, notably in veterinary sciences. Moreover, some strategies or adaptations exist to target particular regions of the genome (candidates approach), instead to sequence all of it, which decrease considerably the experimental cost. With all those last technical improvements, genomics studies identify more and more DNA markers for diverse traits and phenotypes, notably in disease. Their application in veterinary researches will bring more elements to better understand animal diseases, resistance and resilience.

Transcriptomics

Transcriptome (the RNA cell content) is highly dynamic and variable, contrary to the genome. In order to transcript a gene into mRNA, and translate mRNA into a protein, different types of RNAs are involved (miRNA, siRNA, rRNA, tRNA, snRNA). All those RNAs form a complex web of interactions inside the transcriptome

View Article Online
DOI: 10.1039/C6MB00220J

(Fig. 3). Its quantitative and qualitative status depends on the cell type, functions, environment, *etc*. Transcriptome strongly influences the status and performance of the cell, and so tissues, organs and organism.

Transcriptomic aims to quantitatively and qualitatively define the status of different cell transcriptomes. Transcriptomics is used to identify RNA markers related with trait of interest. Such RNA markers bring valuable information and unfold series of pathways involved in the process: *from gene to phenotypes*.

Different techniques exist to access the quantitative and qualitative status of cells transcriptome, like qRT-PCR (quantitative Real-Time Polymerase Chain Reaction), which quantify a specific mRNA in a sample[12]. As this method has a better sensitivity and specificity than other transcriptomics technique, it is often used as validation technique. One limitation of qRT-PCR is the number of genes and samples to analyze, and the availability of primers to study targeted genes. qPCR is a method of reference for identification of RNA markers in veterinary and animal sciences, but also in diagnostic test[13]. DNA micro-array is the simultaneous quantitative analysis of differential expression of thousands of genes between two samples or group of samples, for example infected versus healthy[14]. With this quantitative and qualitative characterization of the transcriptome, it is possible to identify which gene expressions are different between the two samples, and so identify gene related with a disease. With the hundreds or thousands of data generated, it is possible to build entire pathways and explore the molecular biology of the disease. Micro-arrays require dedicated hardware and software, which can be a limitation for laboratories. As micro-arrays analyze much more genes than it require samples, micro-array are highly susceptible to false-positive discovery. It is why very good statistical analyses are important but false-positive discovery is the biggest problem of micro-arrays. The results of micro-array should be validated, usually with qRT-PCR. By using microarrays to compare transcriptome of infected versus control chickens, the study of Smith *et al.* in 2015 establishes pathways of the early immune response of infection by bursal disease[15]. This study enhance the knowledge of pathways involved in infection, and then prospect for markers of animals with a better disease resistance, and diagnostic tools.

In the study of Johansen *et al.* in 2015, pancreas disease and heart skeletal inflammation, associated with 2 different viral infections, were studied in Atlantic salmons[16]. Transcriptomes of infected salmons were compared with controls, but also between the 2 different diseases, by qRT-PCR and microarray. Data from those 2 kinds of experiments generated concordant data. This study is the first to establish a direct comparison of transcriptomes changes caused by viral diseases in salmon. Results allow a better understanding of pathways modified by virus infection, and so a better view on detrimental effect on salmon. This study also identified gene changes specific to each virus diseases, which could be used for diagnostic tools. In species like salmon, which have not been extensively studied like cows for example, such studies are

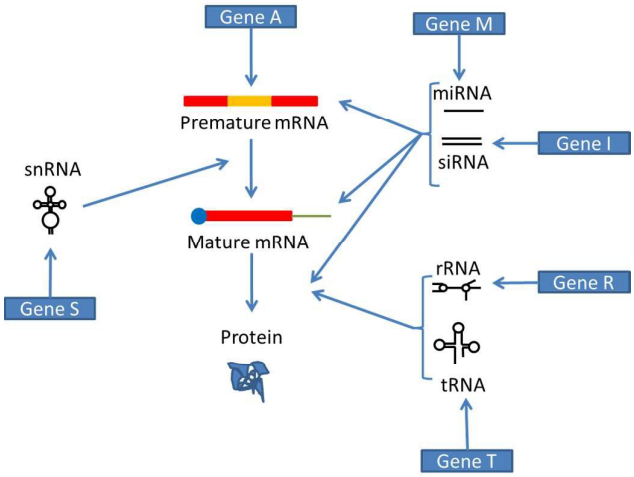


Figure 3. RNA world and complexity of gene expression

essential to produce first available datasets and markers. SAGE (Serial Analysis of Gene Expression) aim to produce an overview of mRNA content with the utilization of molecular cloning tools and sequencing to generate qualitative (identification) and quantitative data (gene expression). SAGE has been used since decades, and like microarray continues to produce results. SAGE can be used on non-sequenced species, which is a great advantage compared to microarray. The study of Mackintosh *et al.* in 2016 compared gene expression of 6 deers, some resistant and others sensible to paratuberculosis infection, with SAGE[17]. They used biopsy samples from lymph nodes taken during 3 different weeks on each individual. They produced 373 million of transcripts tags, and identified 36 632 unique transcripts. 81 genes were upregulated in resistant deers, 234 in sensible deers. They determined that pathways like inflammation, adaptive immune response, host-defense, apoptosis regulation or mitochondrial functions are involved in the resistance against paratuberculosis.

Classical transcriptomics (qRT-PCR, microarrays and SAGE) is still a method of choice to screen rapidly and efficiently changes in gene expression. Messenger RNA transcriptome is still unexplored and lot of works need to be done in animal species, for a lot of different traits related with diseases (infection, resistance, tolerance). Since some years, new or modified techniques were used, notably to produce results on other types of RNA. qRT-PCR is continuously improved to analyze mRNAs related with disease, like cytokines in cow, goat and sheep in the study of Puech *et al.* in 2015[18]. Moreover, qRT-PCR is often used as diagnostic tool to monitor outbreaks, both in human and animal diseases. For example, the study of Gwida *et al.* in 2016 showed that qRT-PCR is a method of reference to monitor outbreak of brucellosis in cattle[19].

RNAseq is a relatively new technique which consists to sequence all RNAs present in the cell with universal primers, and identify them[20]. This qualitative step (identification) is completed by a quantitative step, as RNAseq quantify each sequenced DNA fragments. So researchers have quantitative and qualitative information about gene expression. Contrary to micro-array, RNAseq is able to collect

View Article Online
DOI: 10.1039/C6MB00220J

data from small RNA, thanks to the utilization of universal RNA primers. The microRNA (miRNA) and silencing RNA (siRNA) have been identified recently[21]. They are small (20-25 nucleotides) and their function is to control gene expression. They do not produce any proteins. For that reason, DNA sequences implied in the synthesis of those RNAs has been considered for long time as junk DNA (i.e., DNA which do not finally code for protein). But those RNAs, which are related with trait variation, are the new challenging area to understand complex pathways, and so their genes and DNA targets. It means that a genetic marker could exist outside a DNA protein coding region (strictly speaking), as a potential site for small RNA silencing. This opens an exciting way to the identification of new genetic markers. This field is relatively new and unexplored, and possibility to explore and discover new markers in new regions is a promising tool to understand what has been incomprehensible before. New miRNA markers are published for diseases or conditions, from cancer to parasites infections, in humans and animals[22-24].

As miRNA is a new field of investigation, continuous efforts are required to catalog miRNA in different species. The study of Farell *et al.* in 2015 established a list of 30 new miRNAs in calves (infected with paratuberculosis versus control) but was not able to detect differentially expressed miRNAs and suggested more studies in that way[24]. In trout, the study of Juanchich *et al.* in 2016 established with RNA sequencing a catalog of 2 946 miRNAs (445 were already known) from 38 samples of 16 different tissues. Some tissues exhibited specific expression patterns, confirmed by qRT-PCR. Database about miRNAs is available on miRBase (http://www.mirbase.org/). With 2 588 mature miRNAs identified, human is the most studied species concerning miRNAs (Fig. 4). Of course all the miRNA content of human is far to be completed. Other animal species are far from the present status of human miRNA knowledge. Cow has 793 mature miRNAs identified, and sheep only 153. It is evident that a huge challenge and work is in front of the researchers to catalog miRNAs in animal species, and identify those which are related with diseases.

Proteomics

Protein family is very diverse, in length, physical and biochemical properties and function (Fig. 5). Proteome, like transcriptome, is very dynamic and can change under the effects of a lot of different factors like environment and cell signaling. Different tissues and cell types exhibit strong differences in proteome, quantitatively and qualitatively. Proteomics is the quantitative and qualitative study of the proteome content of cells, organs and organisms and has been widely used in animal sciences[25]. Information about lot of proteins and their isoforms is published in databases. A clear gap exists between the amount of available protein information between

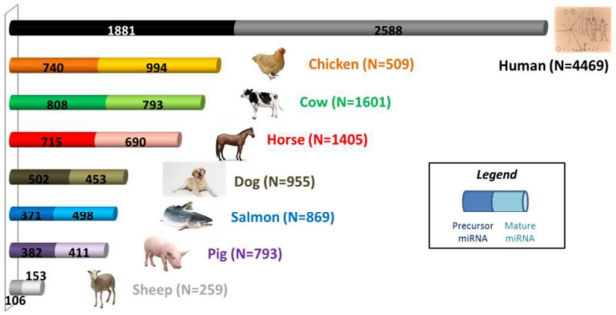


Figure 4. miRNAs identified in different species

humans (more than 1 million) and domestic animals (126 000 for cows), which indicate that a lot of proteomic exploration is necessary in the field of animal sciences (Fig. 6).

Like proteome diversity, investigation techniques in proteomics are versatile, and many were improved with years. Protein electrophoresis is a common and basic technique, and improvements are continuously done, to make the technique easier, cheaper or more accurate. Liquid chromatography coupled with mass spectrometry, a highthroughput approach, is used routinely to qualitatively and quantitatively characterize proteome[26]. This technique is becoming gold standard for the identification of proteins related with diseases. In the study of Valdenegro-Vega in 2014, a list of 52 proteins differentially expressed in salmon mucus after infection by *Neoparamoeba perurans* (parasite amoeba) has been identified with mass spectrometry[27]. Those results highlight the response pathways against parasite infection in salmon, and could bring key elements to set diagnostics tools for example.

In the work of Mansor *et al.* in 2013, milk samples were analyzed by capillarity electrophoresis, liquid chromatography and mass spectrometry, to identify protein biomarkers of mastitis infection[28]. They were able to discriminate the source of mastitis infection, between *Escherichia coli* or *Staphylococcus aureus*, using 47 peptides. This is important to design rapid and accurate diagnostic tests for mastitis. In the study of You *et al.* in 2012, paratuberculosis infected cows were compared with control using 2D-DIGE[29]. Proteins differentially expressed were analyzed with a mass spectrometry to identify them. Improvement of proteomics implies also to adapt techniques to other type of samples. For example, the study of Tanca *et al.* in 2013 accessed the biological relevance of 2D-DIGE on formalin-fixed tissues from hospitals bio-depositories[30]. This allows researchers to work with animal samples preserved with formalin, instead to design a new and costly sampling process.

Other Omics

View Article Online
DOI: 10.1039/C6MB00220J

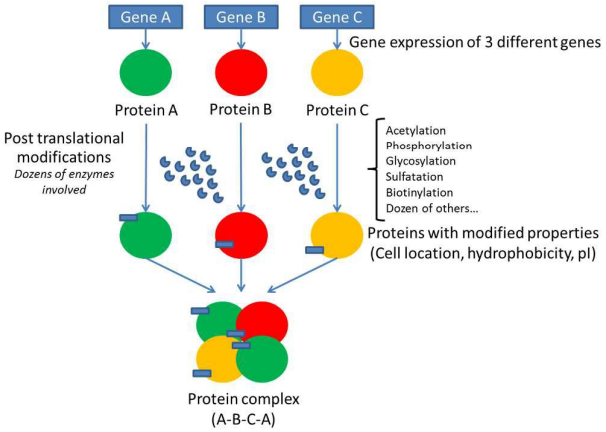


Figure 5. Protein world complexity

Metabolomics aims to study metabolome of the cells, *e.g.* the entire pool of metabolite[31]. Metabolome is highly variable among physiological status of cells. As metabolome is the result of gene expression and proteome activities, it is considered close to the phenotype. This is a promising field to identify metabolite markers related with different kind of diseases or physiological states for cells, both in cytoplasm or secretome. However, metabolomics suffers from limitations due to the chemical nature of metabolites themselves. Contrary to DNA or RNA molecules, metabolites cannot be amplified. Metabolites are not as stable as proteins and their manipulation can be hazardous. Moreover, metabolites are chemically very different among them and require tedious steps to differentiate them. Such experiments are still expensive and their application still remains bottleneck in the animal sciences. Some techniques used to study metabolomes are for example mass spectrometry derived techniques, chromatography, or nuclear magnetic resonance. Using the same techniques, lipidomics aims to study fatty acids and lipids metabolism in the cell[1]. In the study of Minamoto *et al.* in 2015, metabolites in serum of healthy dogs were compared with dogs suffering idiopathic inflammatory bowel disease, by mass spectrometry and chromatography[32]. They identified metabolites like 3-hydroxybutyrate, ribose or hexuronic acid more abundant in disease dogs, underlying a modification of oxidative stress mechanism of in infected dogs. The study of Imhasly *et al.* in 2014 identified 29 metabolites (amino acids and lipids) able to distinguished healthy and hepatic lipidosis suffering cows[33].

Epigenomics, another Omics field, studies the whole pattern of DNA non genetic modification, i.e. DNA methylation or histones modifications[34]. Such epigenetic marks can change DNA expression, and finally phenotype. Moreover, those marks can be inherited, or can be reprogrammed during the early stages of development. Epigenomics use different kind of techniques to identify and quantify those marks on DNA: chromatin immunoprecipitation (ChIP), nuclease susceptibility methods or chromatin sedimentation. The study of Sinclair *et al.* in 2015, a recent example of epigenomic approach, demonstrated that infection of *Anaplasma*

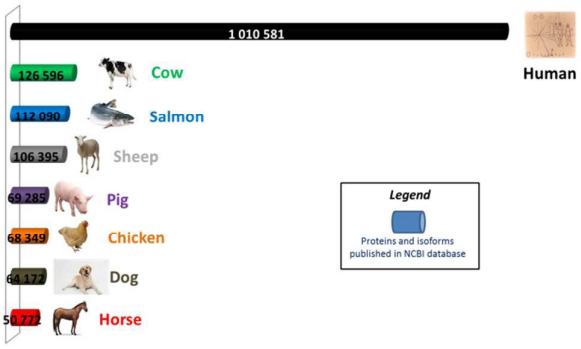


Figure 6. Proteins and isoforms identified in different species

phagocytophilum induces changes in DNA methylation patterns in neutrophils, to promote survival and replication[35].

Recently introduced Omics are still in development, notably to improve sensitivity, and more cost and ease of access. Those techniques target another layer in the molecular content of cells and enrich our knowledge of events between DNA and phenotype. New markers could be defined with those Omics and then used in diagnostic, to evaluate for example some physiological troubles and identify precisely what cellular pathways are in trouble.

Integrative/Systems biology

For the last decade, different Omics have generated a huge quantity of data (quantitatively and qualitatively), in relation with diseases. Omics have identified thousands of markers at different molecular scales, from DNA to chemicals, which are keys to understand complex disease pathways. Despite this collection of knowledge, gaps remain in the understanding of diseases. This is due to the lack of understanding about interactions of those different molecular scales, as a phenotype cannot be summarized in a list of implied molecules[36]. In cell, those levels are not separated and organized in a hierarchical order. Instead, information often goes and back between different scales, which interact each other. Also, different identified markers have not the same importance in the phenotype, and this importance can be modified by different conditions (environment, individual, *etc.*). As a consequence, datasets must be viewed more than a dynamic fluid among all of them, rather than an unchanged table clearly separated in clusters (DNA, RNA, proteins, *etc.*). Those dynamics movements are essential to better understand complex phenotypes, and fill the gap in our understanding.

This map of molecules producing phenotype is called an interactome. Systematic or integrative biology aim to make such interactomes from data generated by Omics[37], to study a complete disease as a whole set of Omics data. This approach constitutes the "post-genomics" era (Fig. 7). Systematic biology do not investigate lonely gene or a candidate protein, its aims is to define a list of every actors related with a disease at all molecular scales (from DNA to chemicals), establish their relationships (control, feedback), calculate a biological model to represent this dynamics, and

View Article Online
DOI: 10.1039/C6MB00220J

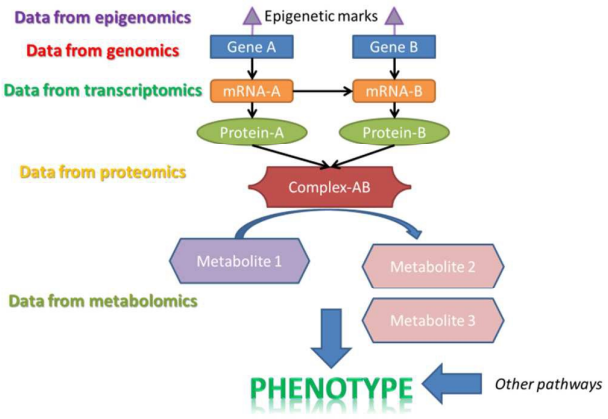


Figure 7. Integrative biology

integrate new data generated by further researches[38]. For example, the study of Kogelman *et al.* in 2015 enhanced the knowledge of obesity in pigs, by conjugating data obtained by transcriptomics and genomics[39]. In the same way, Low *et al.* in 2013 conjugated genomics, transcriptomics and proteomics to better understand hypertension mechanisms in rats[40]. The study of Pineda *et al.* in 2015 proposed a pipeline to integrate epigenomics, genomics and transcriptomics data to finely analyzed complex diseases[41]. Developing interactomes requires a mix of different skills, not only in biology (to bring knowledge and data) but also in mathematics and informatics, to determine accuracy of biomarkers and use tools for interactome construction. Developed interactomes need to be updated with new researches, which is a great challenge and require coordination between different researchers.

Systematic/integrative biology has been possible thanks to recent development of Omics, but also in bioinformatics tools. Researchers have access to software (both free and commercials) able to analyze data and generate pathways, combining sets of different data levels (from DNA to metabolites). Moreover, access to database is becoming easier, and generate a list of biomarkers to complete an interactome derived from experiments for a phenotype is now a common task[42]. Building interactomes is now "user-friendly" for biologists and they can be shared and published without particular skills in mathematics and informatics. Thanks to integrative biology, biologists can have a global overview of a studied phenotype, rather than to ignore all other molecular scales outside his range of usual experiments. Thanks to improvement in bioinformatics and biostatistics tools, biologists can conjugate datasets from different Omics to have a global overview of complex diseases and identify markers[43]. In the case of natural disease, this approach allows for example targeting key elements for vaccine production. Main challenges for integrative biology are to be able to integrate more and more different datasets, generated by a broad kind of experiments, and discrimination of accurate interactomes from false positives.

Exploitation of Omics in animal health sciences

Increase knowledge from QTL and DNA markers

For decades, quantitative genetics have explained phenotypes as the consequence of the effect of QTL, and generated an important list of markers for different diseases in animals. For many phenotypes and diseases, despite a huge list of QTL, full explanation of trait variance is still missing, particularly during the application of identified markers from studied population to on the field. Since few years, a new concept emerged, to open the black box between QTL and phenotypes, i.e. defined every unknown molecules implied in the pathway, from DNA to the final phenotype[3]. Opening this black box imply to study animal disease traits with different Omics and generate a complete interactome, to identify every actors in phenotype establishment.

Thanks to the application of Omics, complex pathways of animal diseases are becoming more and more explained by the conjugation of every kind of Omics[44]. And by consequence, more accurate markers are available to monitor animals and their characteristics concerning health.

Selection and breeding management

By the way, increased knowledge about markers related with disease brings new possibilities in term of applications. New markers identified by Omics enrich the collection of existing means to select animals for specific traits (like disease resistance or tolerance). They allow a more accurate selection scheme, using genomic selection, as they explain the missing variability of previous models. The idea of "personalized medicine" emerged since years, in human medicine[45]. This concept aims to evaluate individual health status, and adapt a care depending on the results. Adapted to animal sciences, animals with different potential in health (resistance against a disease for example) will be managed differently (food, practice) to set their disease resistance at their optimal level, depending on the environment (season), genotype or physiological status (stress, nutrition). That way, it could be possible to optimize breeding conditions depending to decrease losses[46].

Vaccine development

In animal sciences, development of vaccine is a crucial interest for economic performance, as infected animals represent a great loss of money (in term of veterinarian costs but also for production costs). Protected animals represent also a lower threat to human health, as they will no spread zoonosis. With the combination of the studies at different cellular levels (from DNA to metabolites), Omics is a well-adapted methodology to identify key factors for vaccine development. Bioinformatics and *in silico* biology are essential for the analysis and identification of potential targets for vaccine development. Different projects in such way has been started and provided results to identify markers of infection (and so potential vaccine targets) for different pathogens in animals, like *Brucella*, *Salmonella*, or *Streptococcus*[47]. As pathogens evolve to counter animal resistance, vaccine design is a perpetual race, where Omics knowledge, both of pathogens pathways and animal resistance pathways, is the key of success.

View Article Online
DOI: 10.1039/C6MB00220J

Diagnostic development

Diagnostic tools are of a great importance in animal sciences. Ability to detect infected animals at earlier stages, before they can infect other animals and/or humans, is crucial to avoid deleterious consequences (culling, veterinarian costs, and production losses). Such diagnostic tools must be able to detect infected animals with accuracy, i.e. without false positives and negatives. They must be cheap, fast and easy to set up, to not increase the production cost. Omics is also a key for designing such tests, based on microarray, qRT-PCR or ELISA principles for the most part. For examples, using transcriptomics data analysis, study of Rue-Albrecht *et al.*[48] identified macrophage gene responses specific to different Mycobacterium strains. Such genes are targets of interest for diagnostic tool, able to differentiate infectious strains.

Emerging diseases

Another challenge and future application for Omics in animal diseases sciences concern the emergent and re-emergent diseases. Due to the global warming, some infections are more susceptible to occur in European countries, like vector-born parasites[49]. Moreover, some diseases will be re-emerging like malaria or dengue fever. All those diseases are zoonotic and represent a major threat to human health. To be able to counter those diseases, it is essential to have the required tools and knowledge to analyze and identify potential targets for vaccine and diagnostic development, thanks to Omics. For example, a novel orthobunyavirus has been identified and monitored in cattle since its appearance in 2011 in Europe, thanks to Omics technologies which demonstrated their essential role for emerging animal diseases[50].

Conclusions

This review demonstrates that Omics can be deployed in different studies to generate more and more pertinent data. Moreover, different datasets, from genomics, transcriptomics, proteomics, metabolomics and epigenetics, can be conjugated to unravel architecture of complex traits. Omics techniques are so a promising tool, and future challenges in their utilization are the treatment and management of generated data (from a statistical and informatical point of view). Ironically, after simplification of different methods and protocols, allowing every molecular biologist to use every kind of protocols which are more user-friendly, Omics ask more and more skills in statistics and informatics. Data generated by Omics are more complicated to understand due to their important quantity, but also to be able to conjugate different kinds of datasets. Omics massively uses informatics to manage data and perform *in silico* biology, and thus a good understanding of this tool is essential. This constitutes a challenge for molecular biologists. In that sense, a molecular biologist should be multi-tasks, in biology of course, but also in statistics and informatics.

Knowledge about a type of molecule cannot be isolated from other types of molecules, as pathways encompass interrelationships between DNA to metabolites, to produce a final phenotype. That is why molecular biologists should not work only according to their specialization in a kind of Omics, but rather work and think at different levels, to reflect the essence of Omics studies. Team work, and association with other laboratories with their own specialty, is a motor for good Omics projects.

Despite much advancement, Omics are still a huge field of investigation in animal disease sciences. Processes like resistance and tolerance for particular disease, and host-pathogen interactions, remain largely unknown, especially in animals. Utilization of Omics will bring more light and comprehension in disease mechanisms, to be able to design more accurate diagnostic tools and more efficient vaccine. Those tools will then be able to enhance animal welfare and decrease management costs. Another point is the speed of Omics studies. With their high-throughput and big data analysis abilities, complete Omics projects can spend less time than their separated counter parts, and so produce useful knowledge quickly (within 2 to 4 years for example). Considering that new diseases, like Zika, are emerging from tropical and subtropical areas and spread to temperate areas, Omics projects are crucial for a quick development of diagnostics tools and vaccines. Although it is clear that some molecules like miRNAs remain an unknown space to explore, as well as some species are poorly understood, like trout, for which commercial interest is growing (fish farms). Omics are a good mean to "*boldly go where no man has never gone before*".

Acknowledgements

Authors acknowledge the European commission for funding the ERA chair team VetMedZg (ERA Chair Initiative). We also acknowledge HRZZ 4135 to support VM, and APVV-14-218; VEGA1/0258/15 and VEGA 1/0261/15 to support MB. We thank Sandra Dobranić for scientific English correction.

Figures and table legends

Figure 1. Different levels of Omics

DNA structure and functions are studied by genomics. DNA epigenetic marks are studied by epigenomics. Together, those fields bring knowledge about DNA dynamics and status. All RNA types (messenger, silencing, transfer, *etc.*) are studied by transcriptomics, to determine relationship between gene expression and traits. Protein levels and status are studied by proteomics. Like transcriptomics, it aims to establish relationships with traits. Metabolomics and lipidomics study metabolites and lipids. All those molecular levels are implied in phenotype production, and their study produce knowledge useful for phenotype understanding. Omics aim to conjugate all those fields rather to consider each level separately.

Figure 2. QTLs by disease traits and species of agronomical interest

Data come from http://www.animalgenome.org/. Number of QTL published by trait type related with diseases is shown, for

View Article Online
DOI: 10.1039/C6MB00220J

cow, chicken, horse, pig, trout and sheep. Total QTLs traits for diseases are written between brackets for each animal. Cow is the most studied specie for disease QTL (N=1 763), with a huge gravity point on mastitis (73.5%) rather than other diseases. Trout has been particularly ignored in this approach (N=7); sheep are also not so well studied for diseases QTL (N=242).

Figure 3. RNA world and complexity of gene expression

Messenger RNAs (mRNAs) transcript and export the information carried by DNA (Gene A) from nucleus to cytoplasm, to be translated into proteins. Ribosomal RNAs (rRNAs) composed with some proteins the ribosomal complex, which translate mRNAs to produce proteins. Transport RNAs (tRNAs) transport amino acids to the ribosomal complexes. Silencing RNAs (siRNAs) and micro RNAs (miRNAs) are implied in a fine tuning gene expression control. Small nuclear RNAs (snRNAs) modify others RNAs. RNAs are so essential actor of gene expression. All those RNAs, produced by their own genes (M, I, R, T, S), are also regulated by pathways. Single protein chain is rarely enough to produce a functional protein. RNA world and gene expression is then very complex. Transcriptomics aims to improve knowledge and identify actors of those steps, for different kind of RNAs.

Figure 4. miRNAs identified in different species

Studies identified thousands of miRNAs related with different phenotypes, notably diseases. Total numbers of miRNAs is indicated into brackets for each species; precursor and mature miRNAs are indicated on bars directly. With 2588 mature miRNAs identified, human is the best studied species before chicken (994) and cow (793). Sheep miRNAs are poorly identified (153). This graph demonstrated that a lot of researches on miRNAs are necessary to complete knowledge in animal diseases. Data from http://www.mirbase.org/

Figure 5. Protein world complexity

To produce a protein complex, protein chains are produced from their relative genes, with the complexity of gene expression controls. Produced proteins are then processed by post-translational modifications. There are dozens of such modifications, like acetylation or biotinylation. Different enzymes, with their own pathways of regulation, are the effectors of those modifications. Those processes can be influenced by environmental factors (nutrition, stress…). The resulting protein complex will have specific properties (hydrophobicity, cell location). Due to those processes, proteome is very diverse and dynamics.

Figure 6. Proteins and isoforms identified in different species

During decades, proteomics identified thousands of proteins and their isoforms in different species for different phenotypes, including diseases. Total numbers of proteins and isoforms are indicated on bars for each species. With more than 1 million identified proteins and isoforms, human is far ahead from domestic animals. This graph demonstrated that a lot of work remains to do in proteomics to complete knowledge, notably for in animal diseases. Data from http://www.ncbi.nlm.nih.gov/

Figure 7. Integrative biology

In this theoretical example, each level of Omics is studied with a phenotype. Relationships between each level are determined, to have a broad overview of regulation events between each level. Genes A and B, under control of their respective epigenomic marks, produce mRNAs A and B, respectively. mRNA A is implied in fine tuning of gene B expression via silencing. mRNAs A and B are translated into proteins A and B, respectively. Both proteins are only functional when associated in a complex AB. This complex catalyzes the degradation of metabolite 1 into metabolites 2 and 3. This pathway is implied in the studied phenotype control. Other pathways, more or less complexed, are also implied in phenotype control. Omics aimed to define all of those.

Notes and references

1. J. F. Hocquette, I. Cassar-Malek, A. Scalbert and F. Guillou, *Journal of Physiology and Pharmacology*, 2009, **60**, 5-16.
2. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences (Revised Edition)*, Academic Press, 1977.
3. D. P. Berry, K. G. Meade, M. P. Mullen, S. Butler, M. G. Diskin, D. Morris and C. J. Creevey, *Animal*, 2011, **5**, 493-505.
4. O. Abiola, J. M. Angel, P. Avner, A. A. Bachmanov, J. K. Belknap, B. Bennett, E. P. Blankenhorn, D. A. Blizard, V. Bolivar, G. A. Brockmann, K. J. Buck, J. F. Bureau, W. L. Casley, E. J. Chesler, J. M. Cheverud, G. A. Churchill, M. Cook, J. C. Crabbe, W. E. Crusio, A. Darvasi, G. de Haan, P. Demant, R. W. Doerge, R. W. Elliott, C. R. Farber, L. Flaherty, J. Flint, H. Gershenfeld, J. Gu, W. K. Gu, H. Himmelbauer, R. Hitzemann, H. C. Hsu, K. Hunter, F. A. Iraqi, R. C. Jansen, T. E. Johnson, B. C. Jones, G. Kempermann, F. Lammert, L. Lu, K. F. Manly, D. B. Matthews, J. F. Medrano, M. Mehrabian, G. Mittleman, B. A. Mock, J. S. Mogil, X. Montagutelli, G. Morahan, J. D. Mountz, H. Nagase, R. S. Nowakowski, B. R. O'Hara, A. V. Osadchuk, B. Paigen, A. A. Palmer, J. L. Peirce, D. Pomp, M. Rosemann, G. D. Rosen, L. C. Schalkwyk, Z. Seltzer, S. Settle, K. Shimomura, S. M. Shou, J. M. Sikela, L. D. Siracusa, J. L. Spearow, C. Teuscher, D. W. Threadgill, L. A. Toth, A. A. Toye, C. Vadasz, G. Van Zant, E. Wakeland, R. W. Williams, H. G. Zhang, F. Zou and C. Complex Trait, *Nature Reviews Genetics*, 2003, **4**, 911-916.
5. G. D. Wiens, R. L. Vallejo, T. D. Leeds, Y. Palti, S. Hadidi, S. X. Liu, J. P. Evenhuis, T. J. Welch and C. E. Rexroad, *Plos One*, 2013, **8**.
6. H. N. Kadarmideen, *Livestock Science*, 2014, **166**, 232-248.
7. J. S. Witte, in *Annual Review of Public Health, Vol 31*, eds. J. E. Fielding, R. C. Brownson and L. W. Green, 2010, vol. 31, pp. 9-20.
8. T. A. Manolio, F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, J. H. Cho, A. E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin, D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E.

View Article Online
DOI: 10.1039/C6MB00220J

Eichler, G. Gibson, J. L. Haines, T. F. C. Mackay, S. A. McCarroll and P. M. Visscher, *Nature*, 2009, **461**, 747-753.
9. Y. Zare, G. E. Shook, M. T. Collins and B. W. Kirkpatrick, *Plos One*, 2014, **9**.
10. J. C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, G. Jun, A. L. DeStefano, J. C. Bis, G. W. Beecham, B. Grenier-Boley, G. Russo, T. A. Thornton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C. F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M. T. Bihoreau, S. H. Choi, C. Reitz, F. Pasquier, P. Hollingworth, A. Ramirez, O. Hanon, A. L. Fitzpatrick, J. D. Buxbaum, D. Campion, P. K. Crane, C. Baldwin, T. Becker, V. Gudnason, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Moron, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fievet, M. J. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuinness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lleo, A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossu, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. D. Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannfelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. G. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O'Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley, D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A. Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. K. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltunen, E. R. Martin, R. Schmidt, D. Rujescu, L. S. Wang, J. F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Nothen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, G. D. Schellenberg, P. Amouyel, Eadi, Gerad, Adgc and Charge, *Nature Genetics*, 2013, **45**, 1452-U1206.
11. D. P. Berry, M. L. Bermingham, M. Good and S. J. More, *Irish Veterinary Journal*, 2011, **64**.
12. S. A. Bustin, *Journal of Molecular Endocrinology*, 2000, **25**, 169-193.
13. I. Cassar-Malek, N. Guillemin, J. F. Hocquette, D. Micol, D. Bauchart, B. Picard and C. Jurie, *Animal*, 2011, **5**, 867-874.
14. M. Schena, D. Shalon, R. W. Davis and P. O. Brown, *Science*, 1995, **270**, 467-470.
15. J. Smith, J. R. Sadeyen, C. Butter, P. Kaiser and D. W. Burt, *Journal of Virology*, 2015, **89**, 2469-2482.
16. L. H. Johansen, H. L. Thim, S. M. Jorgensen, S. Afanasyev, G. Strandskog, T. Taksdal, K. Fremmerlid, M. McLoughlin, J. B. Jorgensen and A. Krasnov, *Fish & Shellfish Immunology*, 2015, **46**, 612-623.
17. C. G. Mackintosh, J. F. T. Griffin, I. C. Scott, R. O'Brien, J. L. Stanton, P. MacLean and R. Brauning, *Veterinary Immunology and Immunopathology*, 2016, **169**, 102-110.
18. C. Puech, L. Dedieu, I. Chantal and V. Rodrigues, *Bmc Veterinary Research*, 2015, **11**.
19. M. Gwida, M. El-Ashker, F. Melzer, M. El-Diasty, M. El-Beskawy and H. Neubauer, *Irish Veterinary Journal*, 2016, **69**.
20. Z. Wang, M. Gerstein and M. Snyder, *Nature Reviews Genetics*, 2009, **10**, 57-63.
21. D. P. Bartel, *Cell*, 2004, **116**, 281-297.
22. C. Steudemann, S. Bauersachs, K. Weber and G. Wess, *Bmc Veterinary Research*, 2013, **9**.
23. L. Panach, D. Mifsut, J. J. Tarin, A. Cano and M. A. Garcia-Perez, *Calcified Tissue International*, 2015, **97**, 495-505.
24. D. Farrell, R. G. Shaughnessy, L. Britton, D. E. MacHugh, B. Markey and S. V. Gordon, *Plos One*, 2015, **10**.
25. A. M. Almeida, A. Bassols, E. Bendixen, M. Bhide, F. Ceciliani, S. Cristobal, P. D. Eckersall, K. Hollung, F. Lisacek, G. Mazzucchelli, M. McLaughlin, I. Miller, J. E. Nally, J. Plowman, J. Renaut, P. Rodrigues, P. Roncada, J. Staric and R. Turk, *Animal*, 2015, **9**, 1-17.
26. A. Bensimon, A. J. R. Heck and R. Aebersold, in *Annual Review of Biochemistry, Vol 81*, ed. R. D. Kornberg, 2012, vol. 81, pp. 379-405.
27. V. A. Valdenegro-Vega, P. Crosbie, A. Bridle, M. Leef, R. Wilson and B. E. Nowak, *Fish & Shellfish Immunology*, 2014, **40**, 69-77.
28. R. Mansor, W. Mullen, A. Albalat, P. Zerefos, H. Mischak, D. C. Barrett, A. Biggs and P. D. Eckersall, *Journal of Proteomics*, 2013, **85**, 89-98.
29. Q. M. You, C. P. Verschoor, S. D. Pant, J. Macri, G. M. Kirby and N. A. Karrow, *Veterinary Immunology and Immunopathology*, 2012, **148**, 243-251.
30. A. Tanca, S. Pisanu, G. Biosa, D. Pagnozzi, E. Antuofermo, G. P. Burrai, V. Canzonieri, P. Cossu-Rocca, V. De Re, A. Eccher, G. Fanciulli, S. Rocca, S. Uzzau and M. F. Addis, *Proteomics Clinical Applications*, 2013, **7**, 252-263.
31. K. Hollywood, D. R. Brison and R. Goodacre, *Proteomics*, 2006, **6**, 4716-4723.
32. Y. Minamoto, C. C. Otoni, S. M. Steelman, O. Buyukleblebici, J. M. Steiner, A. E. Jergens and J. S. Suchodolski, *Gut microbes*, 2015, **6**, 33-47.
33. S. Imhasly, H. Naegeli, S. Baumann, M. von Bergen, A. Luch, H. Jungnickel, S. Potratz and C. Gerspach, *Bmc Veterinary Research*, 2014, **10**.
34. N. Friedman and O. J. Rando, *Genome Research*, 2015, **25**, 1482-1490.
35. S. H. G. Sinclair, S. Yegnasubramanian and J. S. Dumler, *Clinical Epigenetics*, 2015, **7**.
36. H. Kitano, *Science*, 2002, **295**, 1662-1664.
37. T. Ideker, T. Galitski and L. Hood, *Annual Review of Genomics and Human Genetics*, 2001, **2**, 343-372.
38. H. Y. Zhang, H. Hu, C. Deng, Y. N. Chun, S. T. Zhou, F. Q. Huang and Q. Zhou, *Combinatorial Chemistry & High Throughput Screening*, 2012, **15**, 286-298.
39. L. Kogelman, D. Zhernakova, H.-J. Westra, S. Cirera, M. Fredholm, L. Franke and H. Kadarmideen, *Genome Medicine*, 2015, **7**, 105.
40. T. Y. Low, S. van Heesch, H. van den Toorn, P. Giansanti, A. Cristobal, P. Toonen, S. Schafer, N. Hubner, B. van

View Article Online
DOI: 10.1039/C6MB00220J

Breukelen, S. Mohammed, E. Cuppen, A. J. R. Heck and V. Guryev, *Cell Reports*, 2013, **5**, 1469-1478.
41. S. Pineda, P. Gomez-Rubio, A. Picornell, K. Bessonov, M. Marquez, M. Kogevinas, F. X. Real, K. Van Steen and N. Malats, *Human Heredity*, 2015, **79**, 124-136.
42. N. Guillemin, M. Bonnet, C. Jurie and B. Picard, *Journal of Proteomics*, 2011, **75**, 352-365.
43. N. Guillemin, I. Dufort and M.-A. Sirard, *Gene Expression to Genetical Genomics*, 2015, **8**, 9-18.
44. S. Van Borm, S. Belák, G. Freimanis, A. Fusaro, F. Granberg, D. Höper, D. King, I. Monne, R. Orton and T. Rosseel, in *Veterinary Infection Biology: Molecular Diagnostics and High-Throughput Strategies*, eds. M. V. Cunha and J. Inácio, Springer New York, 2015, vol. 1247, ch. 30, pp. 415-436.
45. L. Hood, J. R. Heath, M. E. Phelps and B. Y. Lin, *Science*, 2004, **306**, 640-643.
46. S. S. Nielsen and N. Toft, *Journal of Dairy Science*, 2011, **94**, 1849-1857.
47. A. Movahedi and D. J. Hampson, *Veterinary Microbiology*, 2008, **131**, 1-13.
48. K. Rue-Albrecht, D. A. Magee, K. E. Killick, N. C. Nalpas, S. V. Gordon and D. E. MacHugh, *Frontiers in Immunology*, 2014, **5**.
49. A. K. Githeko, S. W. Lindsay, U. E. Confalonieri and J. A. Patz, *Bulletin of the World Health Organization*, 2000, **78**, 1136-1147.
50. B. Hoffmann, M. Scheuch, D. Hoper, R. Jungblut, M. Holsteg, H. Schirrmeier, M. Eschbaumer, K. V. Goller, K. Wernike, M. Fischer, A. Breithaupt, T. C. Mettenleiter and M. Beer, *Emerging Infectious Diseases*, 2012, **18**, 469-472.