Impacts of genomics and other 'omics' for the crop, forestry, livestock, fishery and agro-industry sectors in developing countries

1. Introduction

Advances in genomics, the study of all the genetic material (i.e. the genome) of an organism, have been remarkable in recent years. Publication of the first draft of the human genome in 2001 was a milestone, quickly followed by that of the first crop (rice) in 2002 and the first farm animal (chicken) in 2004. Huge technological advancements have meant that sequencing has become dramatically quicker and cheaper over time, so the genomes of many of the important crops, livestock, forest trees, aquatic animals and agricultural pests are now already sequenced or soon will be. The FAO Biotechnology Forum (http://www.fao.org/biotech/biotech-forum/) is hosting this e-mail conference to look at the impacts that genomics, and the other related 'omics', have had so far on food and agriculture in developing countries as well as their potential impacts in the near future.

Before looking at genomics in more detail, a quick overview of some basic genetic concepts can be provided [for more technical details, see FAO (2011a) or the FAO biotechnology glossary (FAO, 2001)]. All living things are made up of cells that contain genetic material called DNA, a molecule made up of a long chain of nitrogen-containing bases (of four kinds: A, C, G and T). DNA is organized as a double helix, where two DNA chains are held together through bonding of the bases, where A bonds with T and C bonds with G. Genomes can be long and 1,000 base pairs (bp) is denoted as 1 kb; 1,000 kb is denoted as 1 Mb; and 1,000 Mb as 1 Gb (i.e. 1,000,000,000 bp). For example, the size of the horse genome is about 2.7 Gb, comparable to that of most mammals, whereas that of the chicken is smaller (about 1 Gb). Compared to animals, genomes sizes in higher plants have a greater range, from about 200 Mb in the horse chestnut and 400 Mb in rice to 2.5 Gb, 12 Gb and 24 Gb in maize, broad bean and slash pine respectively (Murphy, 2011).

Only a small fraction of the genome typically codes for proteins, i.e. where the DNA is first 'transcribed' to a molecule called messenger RNA (mRNA) that is then 'translated' to protein. The remaining and major share of the DNA represents sequences that do not code for proteins and whose role is not yet clearly understood, although this is an area of intense research (The ENCODE Project Consortium, 2012). The genetic material is normally organised into sets of chromosomes (e.g. 5 pairs in *Arabidopsis thaliana*; 30 pairs in cattle), and the entire set is called the genome. In a diploid individual (i.e. where chromosomes are organised in pairs), there are two forms (alleles) of every gene, one from each parent. Note that the genetic material of viruses can be DNA or RNA (so they are called DNA viruses or RNA viruses respectively).

The field of genomics is still relatively young. The first genome sequence of a virus (called phi X 174, containing over 5 kb DNA) was published in 1977 and the first genome of a bacterium (a strain of *Haemophilus influenzae*, almost 2 Mb long) was sequenced nearly 20 years later (Fleischmann et al., 1995). The landmark human genome project began in 1990 and the draft sequence (covering about 90% of the entire 3 Gb genome) was published in 2001 and the full sequence in 2003. This project, in particular, stimulated major technological advancements in DNA sequencing and in bio-informatics (i.e. the use and organization of this kind of biological information, including development of databases, use of computers to analyse the data and integration of information from different sources), so that over time the speed of sequencing has increased rapidly while the corresponding costs have crashed.

Ten years after the draft human genome sequence was published, Lander (2011) noted that sequencing machines could read about 250 billion bases in a week, compared to about 5 million in 2000 and 25,000 in 1990 and that the cost per-base of DNA sequencing had fallen by about 100,000-fold over the decade. Thus, whereas sequencing the first human genome took over a decade and several hundred million US$, it now costs as little as 5,000 to 10,000 US$ (Hayden, 2013) and takes a few weeks. As a result, sequencing the human genome is now becoming much more accessible and the era of 'personal genomics' is beginning, although the costs and time of analyzing the sequence data still remain

substantial (Soon, Hariharan and Snyder, 2013). This (r)evolution is affecting all species, not just humans, and has been made possible by so-called 'next-generation' or 'high-throughput' sequencing technologies, involving the parallel sequencing of large numbers (up to millions) of DNA molecules (Feuillet et al, 2011).

These dramatic improvements have led to a plethora of genome sequencing projects. An overview of these can be seen from the National Center for Biotechnology Information (NCBI), which hosts an important database containing, *inter alia*, information about genomes (http://www.ncbi.nlm.nih.gov/genome/browse/). As at 28 February 2013, the database indicates that the number of species sequenced so far includes 65 animals, 32 plants, 54 fungi and 26 protists as well as 2431 bacteria, 160 archaea and 3329 viruses. All of these completed sequences can be freely downloaded and analysed. This database, as well as others, such as the Genomes OnLine Database (GOLD, http://www.genomesonline.org), also provide information about the several thousand genome sequencing projects currently underway and which are not yet completed.

The huge investments that have been directed towards genomics are driven by the prospects of using the knowledge in several areas, particularly towards human disease diagnosis and prevention and drug development (Soon, Hariharan and Snyder, 2013). In addition, genomics is being used to assist wildlife conservation. For example, the Genome 10K project proposes to sequence the genomes of 10,000 vertebrate species, approximately one for each vertebrate genus, for this purpose (http://genome10k.soe.ucsc.edu/). It is also being used in many different areas of food and agriculture, and this is the focus of the current e-mail conference. A small idea of the enormous amount of ongoing activity in this field can be got by looking at the details of the 159 workshops held during the International Plant & Animal Genome Conference (PAG XXI) which took place in January 2013, describing advances in diverse areas ranging from aquaculture, the buffalo and cacao to rice, sugarcane and wheat rust fungi (http://pag13.mapyourshow.com/5_0/sessions/session_results.cfm?type=SessionType&SessionType=Workshops).

Sequencing the genome of an organism, and the use of bio-informatics to compare it to those of other species, or of individuals from the same species, provides important biological insights, such as the number of genes it contains; organization of its genome; identification of regions that have been strongly affected by genetic selection; and evolution of the genome over time. To find out what each of the genes do, how they are influenced by environmental factors and how they contribute to certain traits (such as growth rate or disease resistance), information on the genome can be supplemented by that from a number of other 'omes', including the transcriptome (i.e. all the mRNA produced by the organism), the proteome (all the protein produced by the organism) and the metabolome (all the small cellular metabolites produced by the organism during cellular metabolic reactions). In this way, information can be gathered on the outputs of the sequential biological processes in the cell whereby DNA is transcribed to mRNA which, in turn, is translated to proteins which are then used, *inter alia*, in cellular metabolic pathways. Study of these 'omes' is called genomics, transcriptomics, proteomics and metabolomics respectively. [Note that in addition to these four main 'omics', several other 'omic' sub-disciplines have also emerged, such as epigenomics which studies genomic-wide epigenetic modifications, like DNA methylation and histone modification, that influence which genes are turned on or off in different cells at different times without changing the DNA sequence].

Note that whereas the genome is quite stable, the levels of mRNA, proteins and metabolites can vary considerably depending on the kind of cells/tissues that are sampled (e.g. the proteins expressed in the udder cells of lactating cows will be different than in their nerve cells), on time (e.g. age of the organism) and on a wide range of environmental factors (e.g. biotic or abiotic stresses).

This Background Document aims to provide information that participants will find useful for the e-mail conference. In Section 2 a brief overview of the current status regarding genomics relevant to food and agriculture is given. Section 3 briefly discusses ways in which the knowledge from genomics can be used. Section 4 presents some specific guidance about the topics that are to be discussed in the

conference. Section 5 provides references of articles mentioned in the document, abbreviations and acknowledgements. Note, for reasons of space, Sections 2 and 3 do not focus on the other 'omics'. These are, however, increasingly being used in each of the sectors to supplement genomics information (see Lidder and Sonnino (2011) for some examples).

2. Status regarding genomics in food and agriculture

2.1 Crops

The first plant genome to be sequenced (in 2000) was *Arabidopsis thaliana*, a small weed of the Brassica family that is much studied as a model plant. In 2002, there was the first crop, when the draft rice genome sequence was published. By the end of 2011, the sequenced genomes included those of the cacao, cassava, cucumber, foxtail millet, grape, maize, peach, pigeonpea, potato, sorghum, soybean as well as relatives of alfalfa and strawberry (Morrell, Buckler and Ross-Ibarra, 2012). Since then, others have been added to the list, including the barley, melon, orange, tomato and watermelon genomes.

To illustrate the typical kinds of strategies used and results obtained from such genomics studies, let us look in more detail at a very recent addition to this list, the chickpea (*Cicer arietinum*), an important legume for food security in many developing countries (Varshney et al, 2013). In their study, they first produced a draft genome sequence of a specific chickpea variety (called CDC Frontier) using the whole genomic shotgun strategy, i.e. where the entire genome is randomly fragmented into small pieces that are sequenced, and the sequences are subsequently assembled using computational methods to produce a consensus sequence (see Feuillet et al (2011) for more details). Their genome 'coverage', indicating how much sequence information they had, was high, at 207x (meaning that, on average, each base was sequenced 207 times). From their sequence, they estimate, *inter alia*, that the chickpea genome is 738 Mb long and contains 28,269 genes (averaging 3,055 bp in length), for whom the function could be tentatively assigned in about 90% of cases. They describe the characteristics of the genome (e.g. nearly half of it consists of 'transposable elements', mobile pieces of DNA that can 'jump' around the genome) and, by comparing it with the genomes from other legume and selected non-legume species, they deduce how the chickpea genome has evolved over the past millions of years.

In the second part of their study, Varshney et al (2013) used whole genome re-sequencing (WGRS) to study the genetic diversity among 29 elite chickpea varieties from different countries. WGRS is an approach where, once a reference genome is available for a given species, genetic variation can be studied by sequencing different individuals of the species and comparing their sequences with the reference genome (Bentley, 2006). They also sequenced 61 additional chickpea accessions using another approach and analysis of the total of 90 genomes made it possible to identify over four million polymorphisms, including single nucleotide polymorphisms (SNPs) and short insertions and deletions (indels), that can be used for future genetic improvement programmes or to identify genes involved in traits of interest. By comparing the genome sequences of landraces and cultivars, they also identified specific regions of the chickpea genome that were potentially under strong genetic selection (i.e. there was a genetic or selective 'sweep') during domestication and breeding.

2.2 Forest trees

Compared to the crop sector, progress in genomics of forest tree species has been much slower. Contributing factors are the very large size of many tree genomes (e.g. ranging from 19 to 24 Gb in spruce and pine species) and limited funding (Neale and Kremer, 2011). The draft sequence of the black cottonwood tree (*Populus trichocarpa)* was the first to be published, in 2006 (see http://www.phytozome.net/poplar.php for more information). Next was the release of the draft sequence of *Eucalyptus grandis* in 2010 (http://www.phytozome.net/eucalyptus.php). There is, however, growing interest in this area (Neale and Kremer, 2011). The ever increasing amount of information from an expanding list of forest tree species contained in specialized tree genomics

databases (such as http://dendrome.ucdavis.edu/treegenes/pubdata/summary_count.php), as well as the papers published in a special issue of the Tree Genetics & Genomics journal in June 2012 (http://link.springer.com/journal/11295/8/3/page/1), all attest to this. As the preface to this special issue notes: "our work on tree genomics is just beginning".

2.3 Livestock

As in the crop sector, the field of genomics is very active in farm animals. During the recent FAO international technical conference on Agricultural Biotechnologies in Developing Countries (FAO, 2011b), which took place in Guadalajara, Mexico in March 2010, the Consultative Group on International Agricultural Research (CGIAR) organized two cross-sectoral sessions on genomics. They indicated that the status of genomic resources in the livestock sector might be slightly more advanced than in plants. The first livestock genome to be sequenced was the chicken (the red jungle fowl, *Gallus gallus*) in 2004 and since then those of the cow, horse, pig, rabbit, sheep and turkey have been released (Fan et al, 2010). The whole genome sequence of the latest addition, the goat, has just been published (Dong et al, 2013).

2.4 Aquatic animals

Following the human genome in 2001, the second vertebrate genome to be released, in 2002, was that of the pufferfish, *Rugu rubripes*, with a genome size of about 365 Mb. Since then, the genomes of other model fish (including the medaka and zebrafish, both model species in developmental biology, and the stickleback, a model for studies of adaptation and speciation), as well as those of important food fish, such as the Atlantic cod and Nile tilapia, have been sequenced in addition to shellfish such as the Pacific oyster. Bernardi et al (2012) report that nearly 60 fish species, including farmed fish and capture fisheries species, are currently being sequenced and an additional 100 species have been identified for sequencing in the near future.

2.5 Micro-organisms

Micro-organisms, including fungi, bacteria and viruses, make up the vast majority of organisms whose genomes have been sequenced, because of their universal importance, particularly in relation to disease, and because their genome sizes are small. Many of those sequenced have direct implications for food and agriculture, including those that cause diseases in plants and animals, are responsible for food spoilage, are used as biocontrol agents and as biofertilizers and in food fermentation processes.

To get a tiny overview of the huge diversity of these microbial genome sequencing projects, we can look at the list of a couple of hundred micro-organisms whose genomes have been sequenced in recent months (Nelson and Garrity, 2012). For bacteria, the list includes strains of *Brucella abortus*, a pathogen that infects livestock and humans; *Cronobacter sakazakii*, a pathogen associated with several outbreaks of food-borne illness; *Dickeya zeae*, the causal agent of rice foot rot in China; *Enterobacter cloacae*, which colonizes rice roots and promotes plant growth by improving plant nutrition; *Geobacillus thermoglucosidans*, a cause of contamination in milk processing plants; *Lactococcus garvieae*, found in a traditional Spanish cheese; *Bifidobacterium bifidum*, a probiotic microorganism that may promote health; *Corynebacterium bovis*, which causes mastitis in dairy cows; and a new bacteria from the *Treponema* genus, found in the cow rumen. The list also includes many viruses, including a number of different strains of the avian influenza virus in China; two types of bluetongue viruses, which infect ruminant animals, in China; two strains of sacbrood viruses, which infect the honeybee, in Korea; an isolate of soybean Putnam virus, which infects soybeans; and an isolate of Peste des petits ruminants virus (PPRV) from wild bharal sheep. The list also includes the filamentous fungus *Aspergillus oryzae* Strain 3.042, which is extensively used for the production of soy sauce and other fermented foods in China.

Metagenomics, the genomics analysis of entire microbial communities, has become increasingly important over the last decade (Relman, 2011). Environmental samples (such as soil; water from lakes,

seas or wastewater treatment plants; or the contents of a rumen) are used to study the genomes of the micro-organisms they contain. An important advantage of the approach is that it is not necessary to culture the micro-organisms involved. Current figures (28 February 2013) indicate that a total of 370 metagenomics projects are ongoing or have been completed, associated with 2737 metagenome samples (http://www.genomesonline.org/cgi-bin/GOLD/metagenomic_classification.cgi).

In this Section 2, the focus has been on the status regarding whole genome sequencing projects. Note that even though a whole sequence might not yet be published for a given species, information on its genome can also be provided through other forms of genomic resources, such as 'libraries' of expressed sequence tags (ESTs) or bacterial artificial chromosomes (BACs) (Quinn et al, 2011). ESTs are short DNA sequences that are complementary to, and derived from, mRNA produced by the organism. BACs are vectors (i.e. DNA molecules which can replicate in a host organism, such as *E. coli*) that can contain large segments (up to 350 kb) of the genome.

3. Applications of genomics in food and agriculture

In this Section, a brief description is provided of some of the main ways in which the knowledge generated through genomics can be used in food and agriculture.

3.1 Genetic improvement of populations

The potential to genetically improve crop, livestock, forest tree, aquatic animal or microbial populations for specific purposes is probably the major driving force behind the use of genomics and the other 'omics' in food and agriculture. They allow researchers and breeders to gain direct access to knowledge about the make-up and functioning of the genetic blueprint of the population or species of interest and to use the knowledge for their genetic improvement.

There are several ways in which this can be done. For example, as illustrated by the example of Varshney et al (2013) in Section 2.1, genome sequencing leads to the identification of unprecedented numbers of molecular markers (such as SNPs) spread throughout the entire genome. Their association with genes of interest can be harnessed for genetic improvement, an approach called marker-assisted selection (see FAO (2007) for an extensive overview in the different sectors). Commercially available "SNP chips" have been developed that allow individuals to be genotyped for tens of thousands of SNP markers distributed across the genome. A number of strategies are available to use the large number of markers for genetic improvement. One strategy is the use of association mapping or genome-wide association studies (GWAS), where genome-wide marker alleles associated with the trait of interest (e.g. yield) are first identified and then, in a second step, markers with significant associations with the trait are used to predict breeding values (Hayes and Goddard, 2010). Another strategy is genomic selection where, unlike GWAS, all of the genome-wide markers are used to predict breeding values. This second strategy requires a reference (training) population of individuals that have been recorded for the trait and genotyped for the markers. Both strategies have been used in animal and plant breeding (Hayes and Goddard, 2010; Varshney et al, 2012). As an alternative to using molecular markers to predict breeding values, it has been proposed that the full genome sequence of individuals be used in the future (Hayes, Lewin and Goddard, 2013).

Genomics and the other 'omics' make it possible to identify individual genes affecting important traits and to understand how they function. This knowledge can be used for genetic improvement within the population, e.g. using marker-assisted backcrossing to transfer an important identified gene from a donor to an elite variety (Varshney et al, 2012), or to transfer the gene to another species for development of a genetically modified organism (GMO).

3.2 Characterization and management of genetic resources for food and agriculture

Domestication of plants and animals began about 10,000 years ago and studies of the timing, location and selection pressures behind the different domestication events provide valuable information about

the current genetic resources for food and agriculture. Genomics has provided a whole range of new tools to explore such issues. For example, Groenen et al (2012) compared the genomes of wild and domestic pigs from Europe and Asia and their analyses indicated, firstly, that the Asian and European wild boars separated from each other roughly one million years ago and, secondly, that there was a clear distinction between European and Asian breeds, thus confirming the hypothesis that beginning about 10,000 years ago, pigs were independently domesticated in western Eurasia and East Asia.

Genomics and genomics-derived molecular markers are also playing an important role in the characterization, study and preservation of wild populations, such as capture fish and forest tree populations. For example, they have been used to estimate the effective population size (a measure of the rate of loss of genetic diversity and rate of increase in inbreeding in a population or species) of the critically endangered North Sea houting (*Coregonus oxyrhynchus*) and to study genetic interactions between stocked hatchery strain brown trout (*Salmo trutta*) and wild brown trout populations (FAO, 2008).

Metagenomics is now being used to characterize and study the diversity of different complex microbial ecosystems that are relevant to food and agriculture. For example, McSweeney and Mackie (2012) describe recent initiatives in this area to study the microbial community in the rumen.

3.3 Food and agricultural product authentication

Genomics and molecular markers derived from genome sequencing can be used to confirm the authenticity of commercially available food and agricultural products (such as timber, fish, seafood, meat, milk, fibre, vegetables or processed foods). This can involve identification at the species level or even at the within species level. For example, Wilkinson et al (2012) describe the use of a high density SNP genotyping assay (developed from SNPs identified and characterized by next generation sequencing) to authenticate pork products derived from specific British pig breeds. The use of genetic markers to identify timber species in order to curb illegal and unsustainable logging (http://www.globaltimbertrackingnetwork.org) and to recently identify horsemeat falsely sold as beef in Europe are typical applications in this area.

3.4 Pathogen detection

Knowledge of the genome sequence of specific strains of micro-organisms (see Section 2.5) makes it possible to accurately identify the agents causing food contamination as well as plant and animal diseases. The specificity involved also makes it possible to trace the source of the pathogen and to monitor its geographical and temporal spread, such as that of the highly virulent race of wheat stem rust Ug99 (http://www.fao.org/agriculture/crops/rust/stem/rust-report/stem-ug99racettksk/en/).

3.5 Vaccine development

Genomics is also used to develop vaccines to manage diseases in livestock and fish. One of the key steps in vaccine development is the identification of potential antigen candidates that may be effective in vaccines (an antigen is a molecule, usually a protein foreign to the animal, which elicits an immune response on first exposure to the immune system by stimulating the production of antibodies specific to its various antigenic determinants. During subsequent exposures, the antigen is bound and inactivated by these antibodies). As noted by OIE (2010), “the field of genomics and related areas has revolutionised the manner in which microbial antigens are identified”. Study of the host genome can also assist in vaccine development. For example, publication of the Atlantic cod genome gave new information about the genetic control of its immune system, so the authors note “our novel findings regarding the immune system will allow for more targeted vaccine development, aiding disease management and the process of domestication of Atlantic cod” (Star et al, 2011).

4. Topics to be discussed in this e-mail conference

This is the 19th e-mail conference to be hosted by the FAO Biotechnology Forum (http://www.fao.org/biotech/biotech-forum/en/) since it was launched in the year 2000. As with each conference hosted by the Forum, the focus is on applications in developing countries.

As seen in Section 2, the first sequenced genomes of domesticated animals and plants were released about 10 years ago and an ever increasing number of important species have been sequenced each year since then. Also, the genomes of several thousand micro-organisms have now been sequenced. As seen in Section 3, the knowledge generated from genomics can be applied in several different ways. In this context, the first main question to be addressed in the conference is:

4.1 What have been the impacts (positive and/or negative) so far of genomics and the other 'omics' for the crop, forestry, livestock, fishery and agro-industry sectors in developing countries?

In addressing this question, the specific kinds of issues that participants might wish to discuss are:

- How exactly was the knowledge derived from genomics and the other 'omics' used?
- What kind of impacts did they have?
- Were the impacts the same in different sectors? If not, why not?
- Were the impacts the same in different developing world regions? If not, why not?
- What products, if any, were derived from the knowledge? If products were developed, how were intellectual property rights issues dealt with (an issue discussed in the second genomics-related session at ABDC-10 (FAO, 2011b))?
- Which specific issues enabled genomics and the other 'omics' to have positive impacts (e.g. government policies, international collaboration, public-private partnerships, complementary infrastructure, germplasm distribution networks)?
- Which specific issues prevented them from having positive impacts (e.g. costs, intellectual property rights, the species sequenced)?
- Were the impacts influenced by the sequencing strategy (e.g. generating genome sequence data in-house in developing countries versus outsourcing this work)?
- The relative importance of genomics versus transcriptomics, proteomics, metabolomics or other 'omics'

While the first main question looks at the past and the present, the second main question looks to the near future:

4.2 What are the impacts (positive and/or negative) of genomics and other 'omics' likely to be in the near future (e.g. the next five years) for the crop, forestry, livestock, fishery and agro-industry sectors in developing countries?

In addressing this question, the kind of issues that participants might wish to discuss might be:

- Whether the magnitude of their impacts is likely to be big or small, and why?
- In which food and agricultural sectors are the impacts likely to be largest?
- In which developing world regions are the impacts likely to be largest?
- What can be changed so that genomics and the other 'omics' can have positive impacts on food security and sustainable development in developing countries in the near future?
- The relative importance of genomics versus transcriptomics, proteomics, metabolomics or other 'omics' in the near future

4.3 Instructions for sending a message

Before submitting a message to the e-mail conference, participants are requested to:

a) Ensure that it addresses the topics mentioned in Section 4 above. (A specific comment: This Background Document briefly mentioned GMOs, and this may also be the case during the conference. If so, discussions in this conference should not consider the issues of whether GMOs should, or should not, be used per se; their regulation; or the attributes, positive or negative, of GMOs themselves).
b) Limit its length to a maximum of 600 words.
c) Follow the 'Guidelines for Sending Messages' contained at the end of the Welcome Text that participants receive when they subscribe to the conference. Among other things, the Guidelines note that participants: are assumed to be speaking on their own behalf and not on behalf of their employers (unless they indicate otherwise); should introduce themselves briefly in their first posting to the conference, providing also their full work address at the end of the message; and may not post libellous, insulting or defamatory messages or materials, or links to such materials and should exercise tolerance and respect toward other participants whose views may differ from your own.

5. References, abbreviations and acknowledgements

Bentley, D.R. 2006. Whole-genome re-sequencing. Current Opinion in Genetics & Development 16: 545–552.

Bernardi, G., Wiley, E.O., Mansour, H., Miller, M.R., Orti, G., Haussler, D., O'Brien, S.J., Ryder, O.A. and B. Venkatesh. 2012. The fishes of Genome 10K. Marine Genomics 7: 3–6.

Dong, Y., Xie, M., Jiang, Y. et al. 2013. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus). Nature Biotechnology 31: 135-141. http://www.nature.com/nbt/journal/v31/n2/full/nbt.2478.html

Fan, B., Du, Z-H., Gorbach, D.M. and Rothschild, M.F. 2010. Development and application of high-density SNP arrays in genomic studies of domestic animals. Asian-Australasian Journal of Animal Science 23: 833-847. http://www.ajas.info/Editor/manuscript/upload/23-109.pdf (330 KB).

FAO, 2001. Glossary of biotechnology for food and agriculture. http://www.fao.org/biotech/biotech-glossary/en/ (available in English, French, Spanish, Chinese, Arabic, Russian, Polish, Serbian, Vietnamese and Kazakh).

FAO, 2007. Marker-assisted selection: Current status, and future perspectives in crops, livestock, forestry and fish. Edited by E.P. Guimaraes, J. Ruane, B.D. Scherf, A. Sonnino, and J.D. Dargie. http://www.fao.org/docrep/010/a1120e/a1120e00.htm

FAO, 2008. Biotechnology as a toolbox to study and monitor agricultural genetic resources. Side event to the 13th meeting of the Subsidiary Body on Scientific, Technical and Technological Advice (SBSSTA) of the Convention on Biological Diversity, 22 February 2008, FAO Headquarters, Rome. http://www.fao.org/biotech/docs/sbstta.htm

FAO, 2011a. Introduction to molecular biology and genetic engineering. Module A, by O. Brandenberg, Z. Dhlamini, A. Sensi, K. Ghosh and A. Sonnino, in the 5-module 'Biosafety Resource Book'. http://www.fao.org/docrep/014/i1905e/i1905e00.pdf (1.2 MB)

FAO, 2011b. Biotechnologies for agricultural development: Proceedings of the FAO international technical conference on "Agricultural biotechnologies in developing countries: Options and opportunities in crops, forestry, livestock, fisheries and agro-industry to face the challenges of food insecurity and climate change" (ABDC-10) http://www.fao.org/docrep/014/i2300e/i2300e00.htm

Feuillet, C., Leach, J.E., Rogers, J., Schnable, P.S. and K. Eversole. 2011. Crop genome sequencing: Lessons and rationales. Trends in Plant Science 16: 77-88.

Fleischmann R.D., M.D. Adams, O. White et al. 1995. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 269: 496-512.

Groenen, M.A.M., Archibald, A.L., Uenishi, H. et al. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491: 393–398. http://www.nature.com/nature/journal/v491/n7424/full/nature11622.html

Hayden, E.C. 2013. Gene sequencing leaves the laboratory: Maturing technology speeds medical diagnoses. Nature 494: 290-291.

Hayes, B. and M. Goddard. 2010. Genome-wide association and genomic selection in animal breeding. Genome 53: 876–883.

Hayes, B., Lewin, H.A. and M. Goddard. 2013. The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. Trends in Genetics. In Press.

Lander, E.S. 2011. Initial impact of the sequencing of the human genome. Nature 470: 187-197.

Lidder, P. and A. Sonnino, 2011. Biotechnologies for the management of genetic resources for food and agriculture. Background Study Paper No. 53. FAO Commission on Genetic Resources for Food and Agriculture. http://www.fao.org/docrep/meeting/022/mb387e.pdf (1.2 MB)

McSweeney, C. and R. Mackie. 2012. Micro-organisms and ruminant digestion: State of knowledge, trends and future prospects. Background Study Paper No. 61. FAO Commission on Genetic Resources for Food and Agriculture. http://www.fao.org/docrep/016/me992e/me992e.pdf (0.9 MB)

Morrell, P.L., Buckler, E.S. and J. Ross-Ibarra. 2012. Crop genomics: advances and applications. Nature Reviews Genetics 13: 85-96.

Murphy, D.J. 2011. Plants, biotechnology, and agriculture. CABI Press, UK.

Neale, D.B. and A. Kremer. 2011. Forest tree genomics: Growing resources and applications. Nature Review Genetics 12: 111-122.

Nelson, O.W. and G.M. Garrity. 2012. Genome sequences published outside of Standards in Genomic Sciences, July-October 2012. Standards in Genomic Sciences 7: 131-149. http://standardsingenomics.org/index.php/sigen/article/view/sigs.3416907/814

OIE, 2012. The application of biotechnology to the development of veterinary vaccines. Guideline 3.3 (adopted by OIE World Assembly in May 2010) of the Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. World Organisation for Animal Health (OIE). http://www.oie.int/international-standard-setting/terrestrial-manual/access-online/

Quinn, N.L., Gutierrez, A.P., Koop, B.F. and W.S. Davidson. 2012. Genomics and genome sequencing: Benefits for finfish aquaculture. In 'Aquaculture', edited by Z. Muchlisin. http://www.intechopen.com/books/aquaculture/genomics-and-genome-sequencing-benefits-for-finfish-aquaculture

Relman, D.A. 2011. Microbial genomics and infectious diseases. New England Journal of Medicine 365: 347-357. http://www.nejm.org/doi/pdf/10.1056/NEJMra1003071

Soon, W.W., Hariharan, M. and M.P. Snyder. 2013. High-throughput sequencing for biology and medicine. Molecular Systems Biology 9: 640. 14 pp. http://www.nature.com/msb/journal/v9/n1/full/msb201261.html

Star, B., Nederbragt, A.J., Jentoft, S. et al. 2011. The genome sequence of Atlantic cod reveals a unique immune system. Nature 477: 207–210. http://www.nature.com/nature/journal/v477/n7363/full/nature10342.html

The ENCODE Project Consortium, 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57–74.

Varshney, R.K., Ribaut, J-M., Buckler, E.S., Tuberosa, R., Rafalski, J.A. and P. Langridge. 2012. Can genomics boost productivity of orphan crops? Nature Biotechnology, 12, 1172-1176. http://www.nature.com/nbt/journal/v30/n12/full/nbt.2440.html

Varshney, R.K., Song, C., Saxena, R.K. et al. 2013. Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nature Biotechnology. In Press. http://www.nature.com/nbt/journal/vaop/ncurrent/full/nbt.2491.html

Wilkinson, S., Archibald, A.L., Haley, C.S., Megens, H-J., Crooijmans, R.P.M.A., Groenen, M.A.M., Wiener, P. and R. Ogden. 2012. Development of a genetic tool for product regulation in the diverse British pig breed market. BMC Genomics 2012, 13: 580. http://www.biomedcentral.com/1471-2164/13/580

****

ABBREVIATIONS: ABDC-10 = FAO conference on 'Agricultural Biotechnologies in Developing Countries'; BAC = Bacterial artificial chromosome; Bp = Base pairs; DNA = Deoxyribonucleic acid; EST = Expressed sequence tag; FAO = UN Food and Agriculture Organization; Gb = Gigabase pairs (1,000 Mb); GWAS = Genome-wide association studies; Indels = Insertion and deletion polymorphisms; Kb = Kilobase pairs (1,000 bp); Mb = Megabase pairs (1,000,000 bp); mRNA = Messenger RNA; RNA = Ribonucleic acid; SNPs = Single nucleotide polymorphisms; WGRS = Whole genome re-sequencing

ACKNOWLEDGEMENTS: This document was prepared by John Ruane, from the FAO Research and Extension Branch. Grateful appreciation is expressed to the following people for their comments on the document: To the external referees: Milton Kanashiro (Brazilian Agricultural Research Corporation (Embrapa), Brazil); Denis J. Murphy (University of Glamorgan, United Kingdom, http://staff.glam.ac.uk/users/184); Miguel Perez-Enciso (Centre for Research in Agricultural Genomics (CRAG), Spain, http://www.icrea.cat/Web/ScientificStaff/Miguel-Perez-Enciso-255); Raul W. Ponzoni (WorldFish, Malaysia, http://cvlist.cslive.org/EmpDetails.aspx?empNum=000557); and Rajeev Varshney (International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India http://www.icrisat.org/CEG/index.htm); as well as to Paul Boettcher (FAO Animal Genetics Resources Branch) and Andrea Sonnino (FAO Research and Extension Branch).

FAO, 28 February 2013.

Recommended reference for this publication:
FAO, 2013. Impacts of genomics and other 'omics' for the crop, forestry, livestock, fishery and agro-industry sectors in developing countries. Background Document to Conference 19 of the FAO Biotechnology Forum (4-24 March 2013). http://www.fao.org/biotech/biotech-forum/