

Landscape of Genomics Technology Platforms in Switzerland

Patrick Descombes^{a*} and Edward J. Oakeley^b

Abstract: This review describes the innovative technological developments that have brought genomics from its beginnings in the late 1980s to the present day and then discusses the ways in which genomics platforms are deployed across Switzerland to play a key role in supporting basic and applied research in both academia and industry.

Keywords: Bioinformatics · Core facilities · Genomics · Microarrays · Sequencing

Introduction

Since the late 1980s, the genomics landscape has evolved at a tremendously fast pace. It began with labor-intensive ‘home-brew’ methods to identify, sequence and analyze single genes at a time and over the following decades transformed into high-throughput platforms and analytical solutions that have opened the door to the comprehensive integration of multiple -omics data types and spawned the analysis of numerous clinical, nutritional and environmental variables. This review will first summarize these innovative and exciting technological developments and then shift to discuss the ways in which genomics platforms have been deployed across Switzerland to play a key role in supporting basic and applied research in both academia and industry. This has made a significant contribution to the shape of research in the country.

Speed-of-light Evolution of Genomics Technologies from the 1980s

In the mid-1970s, F. Sanger invented a method for DNA sequencing using a DNA polymerase to incorporate dideoxynucleotides into radio-labeled DNA strands.^[1] This breakthrough, for which F. Sanger received the Nobel Prize for Chemistry in 1980, paved the way for the modern genomics era with the development of various types of automated sequencers, which would commonly be referred later as ‘first-generation’ sequencing. At the same time, the nascent field of molecular biology sprung to life with the discovery of: 1) restriction enzymes to cut DNA at specific sequences for which W. Arber received the Nobel Prize in Physiology or Medicine 1978; 2) DNA cloning vectors such as plasmids, bacterial artificial chromosomes (BACs), and yeast artificial chromosomes (YACs) to propagate isolated DNA fragments of various length; and 3) the Polymerase Chain Reaction (PCR) which enables the specific amplification of focused DNA sequences, and earned its inventor K. Mullis the Nobel Prize for Chemistry in 1993.^[2] This culminated, at the turn of the century, in the very ambitious Human Genome project. From its beginning, the Human Genome sequencing project faced numerous technical, financial, analytical and ethical challenges.^[3] The first major challenge was that the value of the genome was closely tied to our ability to measure the activities of the various genes that it encoded and the second was that ‘first-generation’ sequencing was just too slow and expensive for this kind of work. This culminated in the development of two major technologies: Microarrays, which opened the door to affordable gene activity and genotypic measurements; and high-throughput DNA sequencing, better known as ‘Next Generation’ Sequencing (NGS) or ‘second-generation’ sequencing,

which made affordable genome sequencing a reality.

Microarrays

A quantum leap in the field was reached with the invention of microarrays in the late 1990s, which provided for the first time a method by which data could be collected in a truly comprehensive way. The first generation of arrays delivered transcriptome analysis,^[4] and were rapidly followed by genotyping,^[5] and chromatin IP analysis (ChIP on chip).^[6]

The principle of microarrays lies in exploiting the complementarity between the two strands of DNA, pioneered by the work of E. Southern.^[7] DNA probes specific for genomic DNA or mRNA targets are immobilized at known locations on a glass surface. The probes were initially either produced from synthetic oligonucleotides or PCR products amplified from DNA or cDNA (DNA copy of RNA) and printed using custom spotting robots.^[8] The ‘home-brew’ printers used in the early days were quickly replaced by commercial solutions, mainly proposing oligonucleotides arrays produced by either *in situ* synthesis by photolithography,^[8] ink-jet synthesis^[9] or coupling of oligonucleotides to beads,^[10] as the need to control batch variability pushed this technology into the mainstream. Fluorescently labeled copies of the sequence of interest such as cDNA or cRNA for gene expression, or DNA for genotyping or ChIP on chip, can then be hybridized to the arrays and photographed by fluorescence microscopy. The signal intensity specifically associated with each spot can be estimated and translated into a relative expression value (for gene expression purposes, see ref. [4] for review), a genotype (presence of a specific polymorphism, for genotyping studies, see ref. [5] for review), or a measure of the relative

*Correspondence: Dr. P. Descombes^a

^aNestlé Institute of Health Sciences SA

EPFL Innovation Park, Bâtiment G and EPFL, SV
CH-1015 Lausanne

E-mail: patrick.descombes@rd.nestle.com

^bNovartis Institutes for Biomedical Research
Novartis Campus, CH-4058 Basel

amount of a protein of interest recruited at a particular genomic location for ChIP on chip.

The gene expression arrays were initially developed based on data coming from massive Expressed Sequenced Tags (ESTs) projects by Sanger sequencing of cDNA libraries originating from various tissues of key model organisms, combined with the development of transcriptome profiling methods such as differential displays^[11] and serial analysis of gene expression (SAGE),^[12] all deposited in public databases (<http://www.ncbi.nlm.nih.gov/nucest>). This was further augmented in later generations using whole genome sequencing data as a reference for the designs. The speed of this process is impressive and reflects the demand for data. The first complete bacterial genome (*Haemophilus influenza*) was published in 1995,^[13] followed one year later by the first complete eukaryotic genome (yeast *Saccharomyces cerevisiae*^[14]). These initial milestones were quickly followed by the first insect (*Drosophila melanogaster*^[15]) and plant genomes (*Arabidopsis thaliana*^[16]) in 2000, and the first draft of the Human Genome in 2001.^[17] Today more than 3,000 eukaryotic genomes, 68,000 prokaryotic genomes and 5,000 viral genomes have either been completed or are in draft versions in the NCBI database (<http://www.ncbi.nlm.nih.gov/genome/browse/>).

Thanks to the wealth of information collected and publicly available, commercial suppliers rapidly provided a portfolio of solutions to interrogate the transcriptomes of numerous species.

Likewise, progress in our understanding of the spread of human genetic variation, thanks to initiatives such as the Haplotype Map project (<http://hapmap.ncbi.nlm.nih.gov/>), coupled with the first medium scale population sequencing projects, lead to the development of high-density genotyping microarrays that enable to interrogate hundreds of thousands to millions of polymorphic regions of the human genome and infer copy number variation.^[5]

Finally with the completion of genome sequencing projects, tiling arrays covering the complete genomes (except repetitive regions) of major model organisms (such as human, mouse, *Drosophila melanogaster*, *Arabidopsis thaliana*, and *Caenorhabditis elegans*) became available, opening the door to genome-wide studies of the transcriptome or the complete landscape of binding sites of transcription factors.^[6,18]

Concomitantly, the field of bioinformatics and associated developments in statistical methods emerged in order to address the significant problem of high volume data storage, analysis and interpretation that IT departments were not equipped for.^[19]

In summary, the major breakthrough brought by microarrays created, for the first time in the history of biology, the ability to globally interrogate the transcriptome and the genome in highly parallel manner with an unprecedented fully comprehensive approach, combined with (relatively) low costs and short time. As we will discuss below for Switzerland, this time also coincides with the decision of many Universities and companies to coordinate and group the implementation and access to these technologies into shared core facilities.

Next-generation Sequencing and Beyond

The heavy cost of the Human Genome project showed the limits of traditional Sanger sequencing, resulting in a desperate need for more appropriate methods to sequence large numbers of genomes at a significantly higher pace and lower cost. In 1990, two scientists sat in a bar in Cambridge (UK) and designed a method for highly parallelized sequencing that led to the formation of the company Solexa. By the mid-2000s this project (now the market leader) and others had matured the current NGS technologies, which now permit industrial scale sequencing. It is worth mentioning here the key local contribution of L. Farinelli, E. Kawashima and P. Mayer (Glaxo-Weclome's Geneva Biomedical Research Institute) with the invention of a DNA clustering method patented in 1997. This clustering method, combined with the Solexa sequencing method forms the basis of the systems currently marketed by the NGS leader Illumina. While far cheaper than the Human Genome project, all these NGS instruments still required substantial investment and technical expertise to run efficiently and again the genomics core facilities played an instrumental role in bringing this advance to the wider scientific community.

It is out of the scope of this review to detail the various sequencing technologies developed, which have been described in depth recently in excellent reviews.^[20] In summary, the NGS technologies have in common the following steps: the unknown DNA (or cDNA derived from RNA) is processed through various molecular biology steps, which fragment it into a desired target size; and then attach flanking adapters of known sequence; this may or may not be further amplified by PCR to give a so-called sequencing library. The library is immobilized onto a solid substrate (a glass plate or a bead) and clonally amplified to generate sufficient DNA molecules for appropriate detection of signal-over-noise during the sequencing

process. Depending on the technology, sequencing is performed either by synthesis or ligation, and base addition is detected by recording either emitted fluorescence following incorporation of labeled bases, or pH changes due to the release of hydrogen ions during base incorporation. The various current suppliers propose instruments with distinct specificities such as read length between 50 to 300 bases, and different output from a few megabases (Mb) to terabases (Tb) per run. For example, the highest throughput systems (HiSeq X-Ten) enables a single laboratory to sequence up to 1.8 Tb per day, corresponding to 18,000 human genomes at 30X coverage per year.^[20b,c] Consequently, the bottleneck has clearly shifted from data production to data storage, analysis and interpretation. The steps started with the microarrays to develop the fields of computing, bioinformatics and bio statistics have evolved to face these new challenges.^[21] The field is still under strong pressure to continue to develop appropriate solutions especially with more data now being processed in the Cloud with solutions for storage and suites of optimized analysis and interpretation pipelines rather than on private networks,^[22] with the ultimate goal to convert 'big data' into 'smart and exploitable data'. One interesting corollary is the shift of contribution of the various factors to global costs of sequencing projects, from being mostly devoted to sequencing with 'first-generation' sequencing to storage, computing and analysis nowadays.^[22b] The revolution brought by NGS has opened new avenues to unprecedented biomedical research such as clinical diagnostics and personalized medicine,^[23] confirmed by the first authorization of NGS for diagnostic purpose by the FDA,^[24] as well numerous examples in the areas of basic research such as for example studies of genomes, transcriptomes, DNA-protein interactions, and metagenomics.^[25]

One of the main limitations of NGS technologies is the relatively short read length, which precludes resolution of the numerous repetitive elements present in genomes. The only 'third-generation' sequencing technology commercially available today for longer reads is the single-molecule real-time (SMRT) sequencing from Pacific Biosciences. In SMRT sequencing, single DNA molecules are individually sequenced in real time by DNA polymerases immobilized at the bottom of zeptoliter-sized chambers. Owing to the library preparation and sequencing process, which do not require any amplification, the technology generates more uniform genome coverage. The latest version of the chemistry produces on average a read length above 14 kilobases (kb), with the longest reads exceeding 60 kb, and about 1 gigab-

ase (Gb) of data per SMRT cell.^[20b,c] These specificities make SMRT sequencing ideally tailored to address the challenge of *de novo* genome assembly, and long reads are providing an additional level of unprecedented structural information for both the genome^[26] and the transcriptome.^[27] However, the current throughput and costs limit the deployment of SMRT sequencing for large genome-wide studies.

Alternative ‘third-generation’ sequencing platforms such as the emerging nanopore-based technologies might provide yet another avenue to further increase read length. Nanopore sequencing is performed by translating peaks of electrical current into a base sequence as long DNA stretches pass through a small channel such as a protein pore or a hole in a graphene layer. However, the current high error rate and low throughput of nanopore sequencing needs to be overcome before it can become competitive against the established second- and third-generation technologies.^[20b,c]

Genomics Technology Platforms in Switzerland

In the late 1990s various initiatives from different Swiss academic actors were launched, with the aim of creating centralized genomics core facilities across the country. The primary motivations of these initiatives were the following: centralize into laboratories accessible to the research community the expensive and complex technologies as well as the knowledge and expertise; ensure an efficient and continuous connection and collaboration with the key technology suppliers to set out consistent pricing and access to the latest state-of-the-art technologies and applications; guarantee a high-level maintenance and availability of these delicate instruments to ensure uninterrupted delivery of high quality data; stimulate the interactions and exchanges between the experts working in the different technology platforms in Switzerland, as well as within the scientific community of users to match the supply to the demand and foster creative brainstorming.

The unifying theme in those early days was the concerted need to bring the new microarray technology and associated analysis to eager scientists keen to leverage this new power for gene expression and variant analysis.

In parallel, the field of bioinformatics was in its infancy but recognized as a much-needed discipline to bring biology from a qualitative, descriptive science to a quantitative computational discipline that could support, provide expertise, and advise researchers on matters such as: data storage; analysis; interpretation;

and dissemination. The Swiss Institute of Bioinformatics (SIB) was founded in 1998 to address these needs. The SIB has since grown significantly and is now an important partner for the researchers and biotechnological platforms of academic institutions across Switzerland. The Vital-IT high-performance computing infrastructure developed by the SIB and its associated bioinformatics core facilities have greatly facilitated the analysis of the genomic data deluge.

In the early 2000s, various complementary strategies were approved across the country with the common, but independent, goal of establishing genomics core facilities to serve local needs. These included:

- The Friedrich Miescher Institute (part of the Novartis Research Foundation) established the Functional Genomics core facility in the Rosental Campus of Basel in 2000. The FMI hosted experts in genomics, proteomics, imaging and bioinformatics to provide a system biology focus to their research. Ties between the Francophone and Alemannic cantons were strengthened when, in 2004, a new director for the FMI in Basel was appointed from the University of Geneva. At this time, the FMI was accepted as one of the first Swiss-German institutes to host bioinformaticians who were also members of the SIB.
- The Genomic Research Laboratory was created in 2000 within the division of infectious diseases of Geneva University Hospitals. This structure – supported by the SNF and several grants – was developed to enable basic research into the pathogenesis of *Staphylococcus aureus*. Research approaches used bacterial genetics, home-brew high-density microarrays and, more recently, NGS. This microbial focus led to the creation of a dedicated bioinformatics group, which subsequently expanded its focus to include medically relevant questions around clinical metagenomics in collaboration with local and international consortia.
- The Life Sciences Training Facility (LSTF) was established at the University of Basel in 2001 with joint support and funding from the Biozentrum and the Department of Biomedicine (DBM). The LSTF’s purpose was to provide researchers with training and access to equipment and software solutions for GeneChip-based profiling experiments. The facility’s success is based on three elements: (i) a strong focus on analyzing RNA with a proven technology (high-density oligonucleotide microarrays), (ii) the possibility for users to acquire the skills for genome-wide

studies rather than outsourcing the work, and (iii) the environment for validating, optimizing and implementing new procedures adapted to the particular needs of a given research project. In 2014 the LSTF entered the Next-generation sequencing era by teaming up with the Genomics Facility Basel (see below). In summary, the LSTF is clearly an outstanding example of a long-term collaboration between the Universities of Basel and Zurich in Switzerland.

- The Genomics Platform of the National Center for Competence in Research (NCCR) ‘Frontiers in Genetics’ was established in 2001 to support the research programs of the local scientific community including the members of the NCCR network, the University of Geneva, and the wider academic and private research communities around the Lake of Geneva and further. Following the initial investment by the NCCR ‘Frontiers in Genetics’, the Genomics Platform expanded in the frame of the ‘Science – Vie – Société’ program in 2003, concomitantly with the implementation of the DNA Array Facility of the University of Lausanne (see below). In 2013, the NCCR ‘Frontiers in Genetics’ program came to an end and the governance of the Genomics Platform was transferred to the newly established Institute of Genetics and Genomics of Geneva (IGE3).
- The DNA Array Facility (DAF) of the NCCR ‘Molecular Oncology’ at the University of Lausanne/ISREC was established in 2001 to provide support to the members of the ‘Molecular Oncology’ network and by extension to the local scientific community. It would merge with the Lausanne DNA Array facility in 2003 (see below)
- The Functional Genomics Center (FGCZ) of the University of Zurich and the Eidgenössische Technische Hochschule (ETH) Zurich opened in 2002. The FGCZ is a research and training facility established to provide an inter-institutional technology platform to support gene expression (with microarray technologies) and proteomics research (by mass spectrometry).
- The DNA Array Facility (DAF) of the University of Lausanne was established in 2003 as one of the core facilities of the Center for Integrative Genomics to provide data generation and data analysis services for the UNIL, CHUV, ISREC and EPFL as part of the ‘Science - Vie - Société’ program. As already mentioned, the University of Lausanne DAF and the ISREC/NCCR ‘Molecular Oncology’ DAF merged in 2003. In 2008 the facility was renamed

‘Genomic Technologies Facility’ to represent more faithfully the support in additional genomics technologies beyond microarrays.

- The Genomic Core Lab of the Institute of Oncology Research was established at the Oncology Institute of Southern Switzerland in Bellinzona in 2003. The main interest is in the application of genome-wide technologies for characterization of lymphoid tumors on the evaluation of anti-lymphoma effect of new compounds. The Genomic Core lab is dedicated to local users – Institute of Oncology Research (IOR) and Institute for Research in Biomedicine (IRB) labs – but is also open to external researchers.
- In order to take advantage of emerging Next Generation Sequencing technologies scientists from three institutions in Basel established an NGS unit at the Department of Biosystems Science and Engineering (D-BSSE of the ETH Zurich in Basel) in 2008. The placement at the D-BSSE ensured high-quality bioinformatics and IT support of the project, by the highly interdisciplinary expertise established at the department, including many groups working in computational biology and the group of software engineers working at the SystemsX.ch unit CISD (Center for Information Sciences and Databases; now part of the Scientific IT Services of ETH Zurich). This concerted initiative culminated in the establishment of the Genomics Facility Basel of the University of Basel and ETH Zurich in 2014 to allow researchers of the life science community in Basel and at ETH Zurich direct access to state-of-the-art genomics technologies.

The microarray technology can be considered as a harmonizing factor, adopted and supported at the same time by a group of investigators across Switzerland. In 2001, M. Primig (at the time at the Biozentrum, Basel) called a meeting of the leaders of these newly established core platforms who agreed to meet regularly and share their respective expertise, successes and issues. This initial meeting resulted in the creation of the Swiss Array Consortium (SAC). The SAC turned out to be a very efficient forum for pre-competitive exchange. It would promote and accelerate post-genomic biological research by working together to evaluate technologies and, through shared experiences and needs, develop informatics tools to enhance research in Switzerland. This initiative has fostered strong connections within the Swiss genomics community, which, despite the waning interest in microarrays, remains a force in our Nation’s scientific endeavors.

The technology platforms of the SAC members have followed the evolution of genomics technologies and have embraced NGS, ‘third-generation’ sequencing, as well as more targeted approaches such as real-time PCR, digital PCR, single-cell analysis, or Nanostring nCounter. Some have expanded their roles to include complementary profiling activities such as proteomic, metabolomic and imaging to address the fast evolving needs of their users. They have also deployed some level of data management and analysis, in order to complement the support provided by initiatives such as the Vital IT platform and to meet the raising expectations of users with differing levels of expertise and available resources in bioinformatics and statistics.

In parallel to these academic initiatives, industry followed a similar path, as exemplified by the following examples. This list is far from exhaustive and is meant to highlight a couple of companies, which have been heavily involved in genomics for many years and have also embraced the concept of genomics technology platforms.

- Novartis was an early industrial adopter of genomics. It recognized this need within both its academic (FMI) and corporate research arms. In 2000, ‘The Genome Factory’ (Novartis AG) was founded to serve the microarray needs of corporate research. The creation of the Novartis Institutes for Biomedical Research (NIBR) in 2004 resulted in the transfer of ‘The Genome Factory’ from Novartis to NIBR and its rebranding as ‘Genomic and Genetic Applications’ (GGA) within the Biomarker Development (BMD) group. In 2009, GGA acquired its first NGS systems. At the end of 2015, a multidepartmental center of excellence for genomics research and collaboration was established on the Novartis Campus in Basel combining clinical, pre-clinical and exploratory research in one location.
- Roche sponsored the Basel Institute for Immunology (BII) between 1971 and 2000 to exploit and develop the new technologies of molecular biology and genetics. In 2000, the BII was closed and replaced by the Roche Center for Medical Genomics, Ltd which focused on molecular pathology and diagnostics. Roche has also been a key player in the long read sequencing field with the marketing of the 454 system (acquired from 454 Life Sciences in 2007) and more recently the co-development of the Sequel system, together with Pacific Biosciences, for use in clinical diagnostics.
- In 2008, Philip Morris International’s hired a senior member of SIB and to-

gether they established the Biological Systems Research department. As a part of it, the Research Technologies group was set up to support the systems toxicology research activities with access to the latest ‘omics technologies and a high performance-computing environment. In 2011, the array-based transcriptomics and the proteomics activities were centralized in Neuchâtel. To complement them, a genomics platform was established and started providing short read sequencing in 2012. It was expanded at the end of 2015 to also provide long reads. In the same timeframe, targeted and non-targeted analysis capabilities were implemented by the proteomics platform, and the transcriptomics platform scaled up to address the increased demand.

- In 2010, Nestlé established the Nestlé Institute of Health Sciences (NIHS) on the EPFL campus in Lausanne. NIHS is a basic research institute deploying complementary ‘omics technologies (genomics, proteomics, lipidomics, metabolomics, and nutrient profiling), and bioinformatics to support pioneering fundamental research programs aimed at understanding health and disease, and developing science-based targeted nutritional solutions for the maintenance of health. The institute is focusing on three main areas: metabolic health (obesity and diabetes), brain health and ageing (cognitive decline and muscle ageing) and gastrointestinal health (microbiome). In this context, the Functional Genomics group has developed its expertise in various state-of-the-art genomics technologies including short- and long-read sequencing, microarrays and targeted technologies, and applications such as gene expression, genotyping, microbiome characterization, *de novo* bacterial and plant genome assembly, as well as bioinformatics and analytical competencies to support the projects of the institute, as well as the global Nestlé R&D.

It is finally worth mentioning that a number of high profile service providers have been established and actively participate in shaping the Swiss genomics landscape. These laboratories have implemented various genomics technologies and applications, from nucleic acid extraction to sequencing and data analysis.

Closing Remarks

The current genomics situation in Switzerland is the emulation of the revolution in the field of genomics, combined with the work of motivated scientists who

engaged to adopt in the early days these fast evolving genomics technologies, were and are still federated into an informal and friendly network which provides a highly collaborative forum for exchange and brainstorming. The evolution of the careers of some of us from academia to industry and reciprocally has further contributed to the efficient networking and strengthening connections between academia and industry.

Genomics is not an exception and core technology platforms have also been deployed in other biology fields in order to address the needs of the community in these various complementary technologies, as presented elsewhere in this issue: chemical biology screening by D. Moreau and J. Gruenberg;^[28] mass spectrometry (proteins) by O. Rinner;^[29] bioimaging by C. R. Bauer.^[30]

Genomics technologies are in constant evolution; new instruments to sequence genomes at even higher pace and lower costs are expected to become available in the very near future. NGS is now fully integrated in biological research. Precision medicine initiatives aim at bringing this revolution to the beds of patients in order to improve the treatment and prevention of their diseases. The goal is to assess likelihood of outcome given prior expectations coming from: genetics; age; the environment; and lifestyle. These advances promises to improve human health but raise numerous questions about data privacy and ethics as access to these data will require gathering across previously distinct domains of information. Refer to the review of D. Trono^[31] in this issue for a presentation of the precision medicine initiative in Switzerland.

One thing is clear, as the years pass, the genomics revolution has a bigger and bigger impact on ever wider fields of research outside of the initial biomedical and basic research domains: agriculture, with assisted breeding; food authenticity; forensic pathology; food safety with food-borne pathogen tracking; paternity testing; and long-term data storage. Predicting the future is always a foolhardy endeavor, given the difficulty we seem to have with keeping track of the present, however it is clear that biology and information technology have become firm friends. It is impressive to witness that the expansion of genomic data is growing so much faster than the famous Moore's law of the last century that a new paradigm of computing is required. As Switzerland excels in the fields of biotechnology, nanotechnology, physics and engineering it will be interesting to watch how the close links that have been forged between these formally distinct disciplines evolve into new ways to explore the world. Interesting and exciting times indeed.

Acknowledgments

We thank the following colleagues and friends for their contributions with key information about their respective laboratories, listed here by alphabetical order: C. Beisel and P. Demougin (Biozentrum, Basel); M. Docquier (Genomics Platform IGe3, University of Geneva); K. Harshman (Genomic Technologies Facility, University of Lausanne); M. Primig (University of Rennes, France); A. Rinaldi and F. Bertoni (Institute of Southern Switzerland Bellinzona); R. Schlapbach (Functional Genomics Center Zurich, ETH Zurich); J. Schrenzel (Genomics Laboratory, Geneva University Hospitals); N. Sierrro, N. Ivanov and J. Thomas (Philip Morris International, Switzerland).

Websites of the various Genomics Technology and informatics Platforms in Switzerland cited in this review

- Functional Genomics Facility, Friedrich Miescher Institute Basel <http://www.fmi.ch/research/platforms/platform.html?plt=104>
- Genomic Platform, IGe3 University of Geneva: <http://www.ige3.unige.ch/genomics-platform.php>
- Functional Genomics Center ETH & University of Zurich: <http://www.fgcz.ch/>
- Genomic Technologies Facility University of Lausanne: <https://www.unil.ch/gtf/home.html>
- Genomics Facility, Institute of Southern Switzerland Bellinzona: http://ior.iosi.ch/site/?page_id=124
- Quantitative Genomics Facility, Biozentrum Basel: <http://www.biozentrum.unibas.ch/research/groups-platforms/overview/umit/qgf/>
- Genomic laboratory, Geneva University Hospital: <http://www.genomic.ch/index.php>
- Vital-IT, SIB: <http://www.vital-it.ch/>
- Novartis Institute for Biomedical Research, Basel: <https://www.nibr.com/stories/medicines/rediscovering-power-natures-exquisite-complexity>
- Functional Genomics, Nestlé Institute of Health Sciences, Lausanne: <https://www.nestleinstitutehealthsciences.com/science-and-technology/omics#topic01>
- Genetics and Genomics, Roche Basel: http://www.roche.com/research_and_development/what_we_are_working_on/research_technologies/informatics-based_technologies/genetics_and_genomics.htm

Received: August 22, 2016

- [1] a) F. Sanger, A. R. Coulson, *J. Mol. Biol.* **1975**, *94*, 441; b) F. Sanger, J. E. Donelson, A. R. Coulson, H. Kossel, D. Fischer, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 1209.
- [2] K. Mullis, F. Faloon, S. Scharf, R. Saiki, G. Horn, H. Erlich, 'Cold Spring Harbor symposia on quantitative biology', **1986**, 51 Pt 1, 263.
- [3] C. DeLisi, *Nature* **2008**, *455*, 876.
- [4] A. Schulze, J. Downward, *Nat. Cell Biol.* **2001**, *3*, E190.
- [5] A. C. Syvanen, *Nature Genet.* **2005**, *37* Suppl, S5.
- [6] T. C. Mockler, S. Chan, A. Sundaresan, H. Chen, S. E. Jacobsen, J. R. Ecker, *Genomics* **2005**, *85*, 1.
- [7] E. M. Southern, U. Maskos, J. K. Elder, *Genomics* **1992**, *13*, 1008.

- [8] D. Gershon, *Nature* **2002**, *416*, 885.
- [9] T. R. Hughes, M. Mao, A. R. Jones, J. Burchard, M. J. Marton, K. W. Shannon, S. M. Lefkowitz, M. Ziman, J. M. Schelter, M. R. Meyer, S. Kobayashi, C. Davis, H. Dai, Y. D. He, S. B. Stephanians, G. Cavet, W. L. Walker, A. West, E. Coffey, D. D. Shoemaker, R. Stoughton, A. P. Blanchard, S. H. Friend, P. S. Linsley, *Nat. Biotechnol.* **2001**, *19*, 342.
- [10] A. Oliphant, D. L. Barker, J. R. Stuelpnagel, M. S. Chee, *BioTechniques* **2002**, Suppl, 56.
- [11] a) B. Kornmann, N. Preitner, D. Rifat, F. Fleury-Olela, U. Schibler, *Nucleic Acids Res.* **2001**, *29*, E51; b) P. Liang, A. B. Pardee, *Science* **1992**, *257*, 967.
- [12] V. E. Velculescu, L. Zhang, B. Vogelstein, K. W. Kinzler, *Science* **1995**, *270*, 484.
- [13] R. D. Fleischmann, M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, *et al.*, *Science* **1995**, *269*, 496.
- [14] A. Goffeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S. G. Oliver, *Science* **1996**, *274*, 546.
- [15] M. D. Adams, S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, R. A. George, S. E. Lewis, S. Richards, M. Ashburner, S. N. Henderson, G. G. Sutton, J. R. Wortman, M. D. Yandell, Q. Zhang, L. X. Chen, R. C. Brandon, Y. H. Rogers, R. G. Blazej, M. Champe, B. D. Pfeiffer, K. H. Wan, C. Doyle, E. G. Baxter, G. Helt, C. R. Nelson, G. L. Gabor, J. F. Abril, A. Agbayani, H. J. An, C. Andrews-Pfannkoch, D. Baldwin, R. M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E. M. Beasley, K. Y. Beeson, P. V. Benos, B. P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M. R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K. C. Burtis, D. A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J. M. Cherry, S. Cawley, C. Dahlke, L. B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A. D. Mays, I. Dew, S. M. Dietz, K. Dodson, L. E. Doup, M. Downes, S. Dugan-Rocha, B. C. Dunkov, P. Dunn, K. J. Durbin, C. C. Evangelista, C. Ferraz, S. Ferreira, W. Fleischmann, C. Fosler, A. E. Gabriellian, N. S. Garg, W. M. Gelbart, K. Glasser, A. Glodek, F. Gong, J. H. Gorrell, Z. Gu, P. Guan, M. Harris, N. L. Harris, D. Harvey, T. J. Heiman, J. R. Hernandez, J. Houck, D. Hostin, K. A. Houston, T. J. Howland, M. H. Wei, C. Ibegwam, M. Jalali, F. Kalush, G. H. Karpen, Z. Ke, J. A. Kennison, K. A. Ketchum, B. E. Kimmel, C. D. Kodira, C. Kraft, S. Kravitz, D. Kulp, Z. Lai, P. Lasko, Y. Lei, A. A. Levitsky, J. Li, Z. Li, Y. Liang, X. Lin, X. Liu, B. Mattei, T. C. McIntosh, M. P. McLeod, D. McPherson, G. Merkulov, N. V. Milshina, C. Mobarry, J. Morris, A. Moshrefi, S. M. Mout, M. Moy, B. Murphy, L. Murphy, D. M. Muzny, D. L. Nelson, D. R. Nelson, K. A. Nelson, K. Nixon, D. R. Nusskern, J. M. Pacleb, M. Palazzolo, G. S. Pittman, S. Pan, J. Pollard, V. Puri, M. G. Reese, K. Reinert, K. Remington, R. D. Saunders, F. Scheeler, H. Shen, B. C. Shue, I. Siden-Kiamos, M. Simpson, M. P. Skupski, T. Smith, E. Spier, A. C. Spradling, M. Stapleton, R. Strong, E. Sun, R. Svirskas, C. Tector, R. Turner, E. Venter, A. H. Wang, X. Wang, Z. Y. Wang, D. A. Wassarman, G. M. Weinstock, J. Weissenbach, S. M. Williams, WoodageT, K. C. Worley, D. Wu, S. Yang, Q. A. Yao, J. Ye, R. F. Yeh, J. S. Zaveri, M. Zhan, G. Zhang, Q. Zhao, L. Zheng, X. H. Zheng, F. N. Zhong, W. Zhong, X. Zhou, S. Zhu, X. Zhu, H. O. Smith, R. A. Gibbs, E. W. Myers, G. M. Rubin, J. C. Venter, *Science* **2000**, *287*, 2185.

- [16] The Arabidopsis Initiative, *Nature* **2000**, *408*, 796.
- [17] a) E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczy, R. LeVine, P. McEwan, K. McKernan, J. Meldrum, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, Y. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, F. Uberbacher, M. Brazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kasprzyk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, J. Szustakowki, International Human Genome Sequencing, *Nature* **2001**, *409*, 860; b) J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M. L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y. H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N. N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J. F. Abril, R. Guigo, M. J. Campbell, B. V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y. H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh, X. Zhu, *Science* **2001**, *291*, 1304.
- [18] a) P. Bertone, V. Trifonov, J. S. Rozowsky, F. Schubert, O. Emanuelsson, J. Karro, M. Y. Kao, M. Snyder, M. Gerstein, *Genome Res.* **2006**, *16*, 271; b) P. Kapranov, A. T. Willingham, T. R. Gingeras, *Nat. Rev. Genet.* **2007**, *8*, 413.
- [19] D. B. Allison, X. Cui, G. P. Page, M. Sabripour, *Nat. Rev. Genet.* **2006**, *7*, 55.
- [20] a) M. L. Metzker, *Nat. Rev. Genet.* **2010**, *11*, 31; b) J. A. Reuter, D. V. Spacek, M. P. Snyder, *Mol. Cell.* **2015**, *58*, 586; c) S. Goodwin, J. D. McPherson, W. R. McCombie, *Nat. Rev. Genet.* **2016**, *17*, 333.
- [21] E. E. Schadt, M. D. Linderman, J. Sorenson, L. Lee, G. P. Nolan, *Nat. Rev. Genet.* **2010**, *11*, 647.
- [22] a) P. Muir, S. Li, S. Lou, D. Wang, D. J. Spakowicz, L. Salichos, J. Zhang, G. M. Weinstock, F. Isaacs, J. Rozowsky, M. Gerstein, *Genome Biol.* **2016**, *17*, 53; b) A. Sboner, X. J. Mu, D. Greenbaum, R. K. Auerbach, M. B. Gerstein, *Genome Biol.* **2011**, *12*, 125.
- [23] a) D. C. Koboldt, K. M. Steinberg, D. E. Larson, R. K. Wilson, E. R. Mardis, *Cell* **2013**, *155*, 27; b) E. R. Mardis, *Nature* **2011**, *470*, 198; c) S. J. Aronson, H. L. Rehm, *Nature* **2015**, *526*, 336.
- [24] F. S. Collins, M. A. Hamburg, *N. Engl. J. Med.* **2013**, *369*, 2369.
- [25] H. P. Buermans, J. T. den Dunnen, *Biochim. Biophys. Acta* **2014**, *1842*, 1932.
- [26] M. J. Chaisson, J. Huddleston, M. Y. Dennis, P. H. Sudmant, M. Malig, F. Hormozdiari, F. Antonacci, U. Surti, R. Sandstrom, M. Boitano, J. M. Landolin, J. A. Stamatoyannopoulos, M. W. Hunkapiller, J. Korlach, E. E. Eichler, *Nature* **2015**, *517*, 608.
- [27] D. Sharon, H. Tilgner, F. Grubert, M. Snyder, *Nat. Biotechnol.* **2013**, *31*, 1009.
- [28] D. Moreau, J. Gruenberg, *Chimia* **2016**, *70*, 878.
- [28] O. Rinner, *Chimia* **2016**, *70*, 860.
- [29] C. R. Bauer, *Chimia* **2016**, *70*, 889.
- [30] D. Trono, *Chimia* **2016**, *70*, 851.