



MIMAS 3.0 is a Multiomics Information Management and Annotation System.

Alexandre Gattiker, Leandro Hermida, Robin Liechti, Ioannis Xenarios, Olivier Collin, Jacques Rougemont, Michael Primig

► To cite this version:

Alexandre Gattiker, Leandro Hermida, Robin Liechti, Ioannis Xenarios, Olivier Collin, et al.. MIMAS 3.0 is a Multiomics Information Management and Annotation System.. BMC Bioinformatics, BioMed Central, 2009, 10, pp.151. <10.1186/1471-2105-10-151>. <inserm-00516451>

HAL Id: inserm-00516451

<http://www.hal.inserm.fr/inserm-00516451>

Submitted on 10 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Database

Open Access

MIMAS 3.0 is a Multiomics Information Management and Annotation System

Alexandre Gattiker^{†1}, Leandro Hermida^{†2}, Robin Liechti³, Ioannis Xenarios³, Olivier Collin⁴, Jacques Rougemont¹ and Michael Primig^{*5}

Address: ¹School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL) and Swiss Institute of Bioinformatics (SIB), CH-1015 Lausanne, Switzerland, ²Friedrich Miescher Institute for Biomedical Research, CH-4058 Basel, Switzerland, ³Swiss Institute of Bioinformatics, Vital-IT, CH-1015 Lausanne, Switzerland, ⁴Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA), Université de Rennes 1, F-35042 Rennes, France and ⁵Inserm, U625, GERHM; IFR-140; Université de Rennes 1, F-35042 Rennes, France

Email: Alexandre Gattiker - alexandre.gattiker@epfl.ch; Leandro Hermida - leandro.hermida@fmi.ch; Robin Liechti - robin.liechti@isb-sib.ch; Ioannis Xenarios - ioannis.xenarios@isb-sib.ch; Olivier Collin - olivier.collin@irisa.fr; Jacques Rougemont - jacques.rougemont@epfl.ch; Michael Primig* - michael.primig@inserm.fr

* Corresponding author †Equal contributors

Published: 18 May 2009

Received: 9 January 2009

BMC Bioinformatics 2009, 10:151 doi:10.1186/1471-2105-10-151

Accepted: 18 May 2009

This article is available from: <http://www.biomedcentral.com/1471-2105/10/151>

© 2009 Gattiker et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: DNA sequence integrity, mRNA concentrations and protein-DNA interactions have been subject to genome-wide analyses based on microarrays with ever increasing efficiency and reliability over the past fifteen years. However, very recently novel technologies for Ultra High-Throughput DNA Sequencing (UHTS) have been harnessed to study these phenomena with unprecedented precision. As a consequence, the extensive bioinformatics environment available for array data management, analysis, interpretation and publication must be extended to include these novel sequencing data types.

Description: MIMAS was originally conceived as a simple, convenient and local Microarray Information Management and Annotation System focused on GeneChips for expression profiling studies. MIMAS 3.0 enables users to manage data from high-density oligonucleotide SNP Chips, expression arrays (both 3'UTR and tiling) and promoter arrays, BeadArrays as well as UHTS data using MIAME-compliant standardized vocabulary. Importantly, researchers can export data in MAGE-TAB format and upload them to the EBI's ArrayExpress certified data repository using a one-step procedure.

Conclusion: We have vastly extended the capability of the system such that it processes the data output of six types of GeneChips (Affymetrix), two different BeadArrays for mRNA and miRNA (Illumina) and the Genome Analyzer (a popular Ultra-High Throughput DNA Sequencer, Illumina), without compromising on its flexibility and user-friendliness. MIMAS, appropriately renamed into Multiomics Information Management and Annotation System, is currently used by scientists working in approximately 50 academic laboratories and genomics platforms in Switzerland and France. MIMAS 3.0 is freely available via <http://multiomics.sourceforge.net/>.

Background

Microarrays are an essential tool for high-throughput analysis of single nucleotide polymorphisms (SNPs), DNA rearrangements, RNA concentrations, exon composition and protein-DNA interactions [1-3]. Microarray technology is based on distinct manufacturing approaches such as robotic application of double stranded DNA fragments onto glass slides (spotted arrays) [4], *in situ* synthesis of high-density oligonucleotide probes (GeneChips) [5] and bead-based systems (BeadArrays) [6]. A comprehensive set of open source and commercial bioinformatics solutions have become available over the last decade that includes certified public array data repositories in Europe, the US and Asia [7-9], a platform for anonymous peer-review of genome biological manuscripts [10] and many web-based [11] or local array data management and analysis solutions [12]. International standards for data acquisition, representation and interchange developed by the Microarray Gene Expression Data Society (MGED, [13]) include the Minimum Information About a Microarray Experiment (MIAME) guidelines [14], the MicroArray and Gene Expression (MAGE) data representation standard [15], the MAGE-TAB interchange format [16] and the MGED Ontology for microarray experiment and biological sample annotation [17]. Annotating microarray data according to MIAME guidelines and depositing them in certified repositories ArrayExpress (EBI), Gene Expression Omnibus (NCBI), and CIBEX is mandatory for publishing in most scientific journals although this policy is regrettably not always rigorously enforced [18].

Most recently, novel ultra-high throughput DNA sequencing (UHTS) technologies have been developed that enable researchers to obtain the complete genomes of model organisms and *H. sapiens* much faster and at a much lower cost than classical methods [19]. Moreover, these technologies appear also to be extremely useful for accurately measuring gene expression (RNA-Seq) [20] and protein-DNA interactions (ChIP-Seq) [21]. However, no widely accepted standard exists as yet within the community to report and archive the output of UHTS experiments. Efforts are under way in collaboration with MGED's Reporting Structure for Biological Investigations (RSBI) work group to develop a single format for annotations across multiple technologies [22] and a standard for UHTS similar to what was developed earlier for microarrays (MINSEQE) [23]. Currently, UHTS experiments are annotated using the MGED Ontology prior to submission to the European Nucleotide Archive [24].

Here we report a substantial extension of our system for consortium-wide microarray data management [25]. The novel web-accessible Multiomics Information Management and Annotation System (MIMAS 3.0) is based upon

an elaborate graphical user interface (GUI) and a scalable relational database. It is designed to store manually annotated expression data from several research facilities that may be organized within a consortium. Version 3.0 has been extended to support data produced by eight different types of microarrays from the two most popular manufacturers as well as annotation and genome location data derived from sequencing experiments. Data representation was standardized according to the MAGE-TAB data exchange format and MGED Ontology. A one-step export feature creating a MAGE-TAB spreadsheet is available facilitating submission to the ArrayExpress repository. MIMAS 3.0 is freely available under the GNU license at <http://multiomics.sourceforge.net/>[26].

Construction and content

The database model

The database model was initially constructed as generically as possible which improves knowledge representation and simplifies maintenance [25]. It was therefore not necessary to modify the model when the system was extended from supporting only GeneChip expression arrays to including seven other different array types (see Additional File 1 in [25]). The software is organized around a database-driven reflective architecture [27]. This is in contrast to a more traditional design where each particular attribute values are stored in a separate column, and specific code is needed to handle the input and storage of each attribute. In an adaptation of the Type Object design pattern [28], attribute definitions are themselves objects in the database. They are defined by their names and value type. Parameters including the default value, whether multiple values are allowed, and the list of possible values if a choice list must be presented. The software reads these definitions to generate fully dynamic web forms, and stores input values in a generic table accommodating all attribute types.

An inconvenience of such generic data storage is the relative inefficiency of queries and data retrieval. However, in our experience, the amount of annotation data is well within the capability of modern relational database management engines, and even complex searches perform almost instantaneously. If search performance were to become an issue, materialized views could easily be introduced to create a de-normalized data mart for efficient searching.

Since the MGED Ontology is limited and does not cover domains such as chemical compounds, MIMAS contains its own controlled vocabulary that greatly overlaps the MGED Ontology. This controlled vocabulary is extended in a controlled manner (via curation) via direct user input. Users can call up a list of all available annotation terms and they can select an option to enter novel terms as they

deem appropriate. These terms are validated by the database curator and included into the controlled vocabulary.

The web application

MIMAS 3.0 GUI is optimized to upload, annotate and manage microarray and UHTS experiments and it includes functionalities to search, download and export data. Its work-flow is designed to allow biological and biomedical researchers with little or no bioinformatics skills to input the data. This facilitates the task of high-throughput data annotation often carried out by research technicians who work at service platforms that are accessible for large numbers of laboratories. It also provides the necessary infrastructure for annotation of biological samples and technical parameters that must follow experimental work to ensure proper data archiving (Figure 1). MIMAS 3.0 does not require any particular operating system, web browser or plug-in software.

New Experiment annotation and management procedure

As described previously [25], MIMAS facilitates organization of samples according to experimental conditions and allows for propagation of annotation information among samples to speed up the process of data description. The intuitive user interface and the design allowing for input from different experts on the same project encourages interaction between technicians and experimental biologists during the annotation process (Figure 1). Another important novel feature is that files and annotation information may be entered into the system in any order, thus encouraging users to process data early on during the experimental work-flow. This is crucial especially for large projects that involve hybridization or sequencing of numerous samples over long periods of time which requires sustained data input to avoid data loss or incorrect annotation.

Several users may be involved at different stages of the annotation process. Since this increases the possibility of errors the data input is validated by a local MIMAS curator prior to uploading into the database. Therefore only curated information is available for data export. Moreover, the annotation data are verified a second time by curators at the EBI's ArrayExpress which further increases the reliability of data archiving. In this context we note that seamless exporting of data to ArrayExpress enables users to exploit its extensive Application Programmer Interface (API) for further programmatic access; we currently favor this generic approach over developing a MIMAS-specific API.

Integration of new high-throughput methods and data formats

MIMAS contains parser modules for the file formats listed in Table 1. Extending our software solution to support

new technologies and data file formats is a straightforward process. The list of supported technologies can be edited via the web interface. Integration of a new data format requires changes to the master parser Perl module that determines a file's format based on its name or its content. It is recommended, but not essential, to write a file-format-specific Perl module for verification of a file's integrity and validity and to extract certain data fields.

Utility

MIMAS has been instrumental for managing and publishing high-throughput microarray data from many laboratories organized within the Swiss Array Consortium, including our own [29-32] and it is now also hosted by the bioinformatics platform [33] of Ouest-Genopole [34]. Its user-friendly interface enables life scientists to locally annotate and upload their raw GeneChip, BeadArray as well as UHTS data to certified repositories. We propose a flexible and scalable solution developed within a network of collaborators that includes bioinformatics experts, research technicians and life scientists. MIMAS is an integral part of our research program and as such it is a sustained software system. This is not to be taken for granted in the fast-paced bioinformatics field of data management solutions that have to deal with rapidly evolving and quickly emerging genome biological technologies.

Discussion

MIMAS 3.0 is a versatile data management solution for the output of microarray experiments based on GeneChips and BeadArrays and data produced with ultra high-throughput DNA sequencing equipment. We do not associate raw data with UHTS experiments because their huge volume (up to 100 gigabytes) and the very large number of text and image files (up to 2000 text files and 14000 image files) produced by them makes storage and access via a web interface impractical. Moreover, no convention exists yet that defines what is considered to be useful raw data such as for example images, fluorescence intensity files, intermediate files or final filtered sequence lists. We therefore decided for the time being to support files in the GFF format which are generated by mapping DNA sequence read output to a reference genome.

The database design facilitates incorporation of novel data formats used in emerging technologies. The system is optimized for a multi-platform and multi-user environment with an up-to-date annotation interface enabling scientists and research technicians to efficiently process large quantities of genome biological data. As such, MIMAS 3.0 is unique among comparable data management software solutions that are often limited to a specific technical platform or that require complicated annotation procedures.

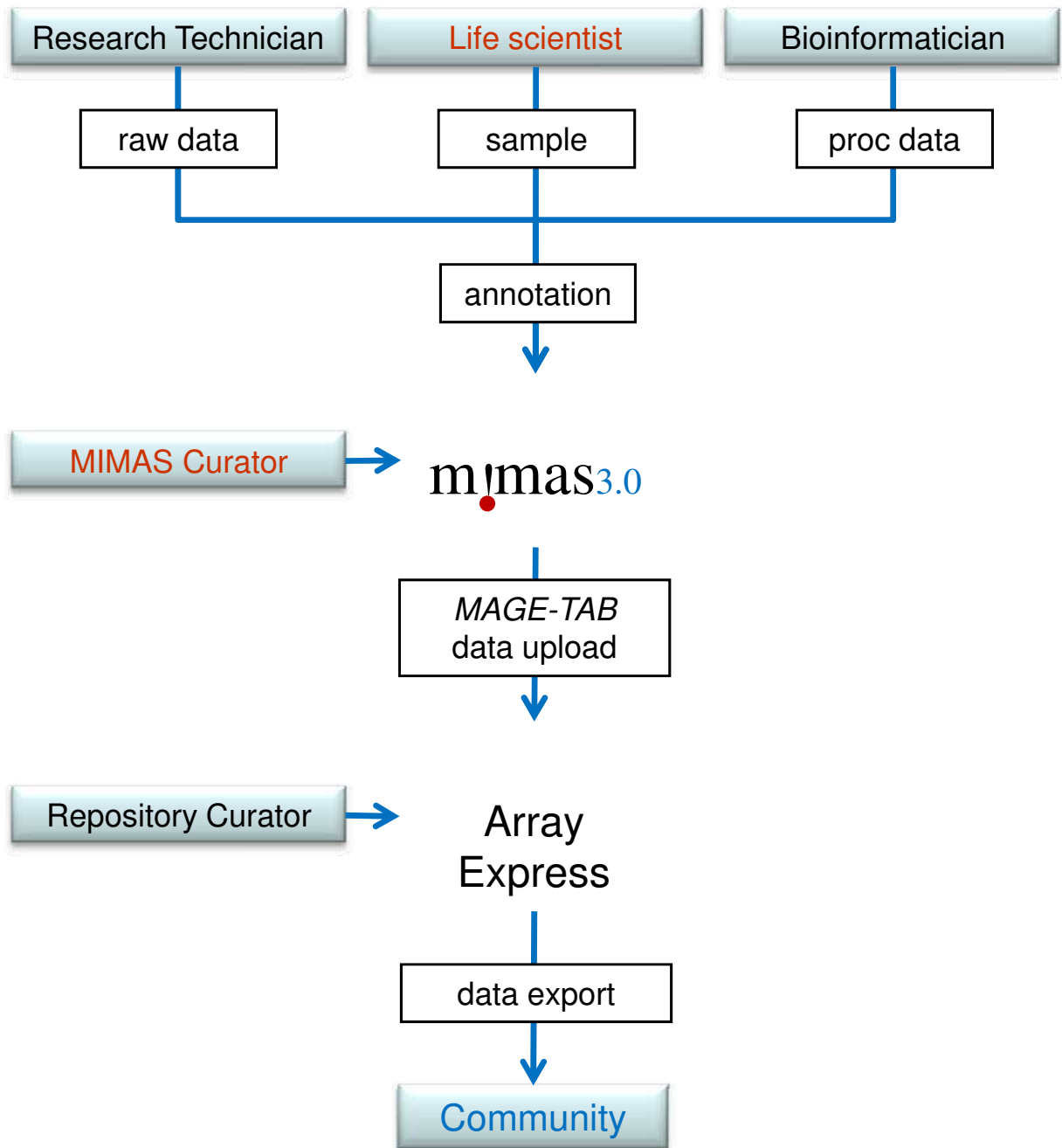


Figure 1
MIMAS 3.0 Data Annotation and Submission Workflow. Tasks performed by different types of users and curators are depicted. A single user can have several roles, depending on his or her technical skills and the organization of the facility or the consortium. The order in which files and annotation data are stored in MIMAS is flexible and organized within user-defined experiments. Raw data are unprocessed CEL (Affymetrix), tabular files and GFF3 files for DNA sequencing (Illumina). Proc data are processed data files. Sample is biomedical input describing the cells and tissues examined.

Table 1: Technologies and data formats supported by MIMAS 3.0.

Technology	Raw data format	Processed data format
Affymetrix GeneChips	Affymetrix CEL file (GCOS versions 3 or 4 or CCG version 1)	Tabular (Affymetrix MAS/GCOS versions 4 or 5, GEO MAS, GEO Series Matrix)
Affymetrix All-Exon Arrays		
Affymetrix Tiling Arrays		
Affymetrix Promoter arrays		
Affymetrix SNP arrays		
Affymetrix Gene Arrays		
Illumina BeadArrays	Tabular (Illumina BeadStudio ArrayExpress-format output plug-in)	
Illumina miRNA arrays		
Illumina Genome Analyzer		GFF 3

Comparison with other solutions

Most of the alternative data management software solutions described below provide data export options in MAGE-ML format. This allows for easy data submission to the ArrayExpress and GEO public repositories. However, the MAGE-ML format is complex and has not evolved since 2002. In contrast, MIMAS supports the MAGE-TAB format which is editable in any spreadsheet software. It is therefore straightforward to review annotation in tabular form prior to submission, and to include additional information beyond the scope of the data management software. This could include for example geographical coordinates of locations where metagenomics samples were collected [35].

BASE (BioArray Software Environment) is a web based system under active development which accommodates data from high-density oligonucleotide microarrays (Affymetrix), BeadArrays (Illumina) and microarrays based on adhesion of DNA fragments onto a glass support from academic or commercial sources (two-color microarrays) [36] (we refer only to the website and not the original publication since version two of the system has been completely rewritten and is as yet unpublished). This software enables users to import and export experimental data in the Tab2MAGE format. However, the current version 2.9 provides only limited support for tiling or SNP arrays and lacks a solution for storing UHTS data. In our opinion this product implements a time consuming, repetitive, and thus error prone annotation procedure (Table 2). MARS (Microarray Analysis and Retrieval System) is a MIAME-compliant software suite for storing,

retrieving, and analyzing multi color microarray but not GeneChip and BeadArray or UHTS data [37]. SBEAMS is a very elaborate system that supports a wide array of functional genomics technologies but does not support exporting in the MAGE-ML or MAGE-TAB formats. Moreover, the complexity of its interface makes it cumbersome to use for experimental biologist, biomedical researchers and clinicians [38]. MiMiR is designed to be used in clinical trials and provides an advanced security infrastructure for that purpose. As such, it provides specialized vocabularies embedded within a complex mapping tool [39]. The maxLoad2 software has not been extended for several years and is a stand-alone solution installed in individual PCs. It does not support a server installation, which prevents collaboration and could lead to problems with data integrity [40]. MIAMExpress is practical especially for researchers who do not have a local data repository and wish to submit their data directly to ArrayExpress. However, it only supports a limited range of microarray platforms and no DNA sequencing platforms [41].

Future work

We intend to maintain and further develop MIMAS 3.0 as a data annotation and archiving solution for technologies that yield information on DNA integrity, gene expression and protein-DNA interactions via microarrays and UHTS methods. Moreover, we wish to include other data types including the output of experiments aiming at RNA and protein expression by *in situ* hybridization and immunohistochemistry [42] and Tissue Microarrays (TMAs) that yield data and images for hundreds of normal and cancerous samples [43]. Finally, we ultimately plan to

Table 2: Comparison of MIMAS 3.0 with other solutions as described in the literature.

	MIMAS 3	BASE 2	MARS	SBEAMS	MiMiR	maxdload2	MIAMExpress
Affymetrix GeneChip	*	*		*	*	*	*
Two-color Microarray		*	*	*	*	*	*
Illumina BeadArray	*	*					
UHTS	*						
MIAME compliant	*	*	*	*	*	*	*
MAGE-TAB export	*	*					*
MAGE-ML export		*	*		*	*	*
MGED Ontology support	*		*		*	*	*
Connectivity to analysis and visualization tools	*	*	*	*	*	*	
User and group management	*	*	*	*	*		
Object sharing and permissions management	*	*	*	*	*		
Web application	*	*	*	*			*
User Notifications/Messages	*	*	*				
Wizard-based annotation and experiment creation	*	*					

incorporate the output of protein expression data measured by mass spectroscopy into MIMAS [44].

Conclusion

MIMAS 3.0 was developed by a network of software developers, computer scientists, life scientists, and research technicians involved in high-throughput data production, analysis and interpretation. In our experience such a constellation spawns solutions that are user-friendly, efficient and durable. A key problem of open source software is often, apart from programming errors, lack of long-term support. Our software has been successfully used for several years by a large number of laboratories in Switzerland and it was recently also set up at the bioinformatics platform of Biogenouest in France. MIMAS 3.0 is an important element of our ongoing genome biological research and will therefore continue to be developed in the foreseeable future. The software is freely available at the Sourceforge repository.

Availability and requirements

- Project home pages: <http://mimas.vital-it.ch/>; <http://mimas.genouest.org/>; download via <http://multiomics.sourceforge.net/>

- Operating system(s): Linux/UNIX, Mac OS X, Windows
- Programming languages: Perl, SQL, JavaScript
- Apache web server with mod_perl extension
- Oracle 9 or later or MySQL 4.1 or later.

Abbreviations

API: (Application Programmer Interface); GCOS: (GeneChip Operating System); GFF3: (Generic Feature Format Version 3); GUI: (Graphical User Interface); MAGE: (MicroArray and Gene Expression); MAS: (Microarray Analysis Suite); MINSEQE: (Minimum Information about a high-throughput Nucleotide Sequencing Experiment); MGED: (Microarray Gene Expression Data); MIAME: (Minimal Information about a Microarray Experiment); UHTS: (Ultra high-throughput DNA Sequencing); TMAs: (Tissue Microarrays).

Authors' contributions

AG and LH designed and developed the database software and AG also contributed to the manuscript. IX and JR contributed to software development and they host the data-

base in Lausanne. RL maintains and curates the database in Lausanne. OC maintains the database in Rennes. MP contributed to the design and wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

We thank F. Chalmel for critical reading of the manuscript. This project was supported by the Swiss Institute of Bioinformatics, the bioinformatics platform of Biogenouest and by Inserm Avenir grant N° R07216NS awarded to M. Primig.

References

- Kapranov P, Sementchenko VI, Gingeras TR: **Beyond expression profiling: next generation uses of high density oligonucleotide arrays.** *Brief Funct Genomic Proteomic* 2003, **2(1)**:47-56.
- Kapranov P, Willingham AT, Gingeras TR: **Genome-wide transcription and the implications for genomic organization.** *Nat Rev Genet* 2007, **8(6)**:413-423.
- Hudson ME, Snyder M: **High-throughput methods of regulatory element discovery.** *Biotechniques* 2006, **41(6)**:673. 675, 677 passim.
- Hager J: **Making and using spotted DNA microarrays in an academic core laboratory.** *Methods Enzymol* 2006, **410**:135-168.
- Dalma-Weiszhausz DD, Warrington J, Tanimoto EY, Miyada CG: **The affymetrix GeneChip platform: an overview.** *Methods Enzymol* 2006, **410**:3-28.
- Fan JB, Gunderson KL, Bibikova M, Yeakley JM, Chen J, Wickham Garcia E, Lebruska LL, Laurent M, Shen R, Barker D: **Illumina universal bead arrays.** *Methods Enzymol* 2006, **410**:57-73.
- Barrett T, Suzek TO, Trup DB, Wilhite SE, Ngau WC, Ledoux P, Rudnev D, Lash AE, Fujibuchi W, Edgar R: **NCBI GEO: mining millions of expression profiles – database and tools.** *Nucleic Acids Res* 2005:D562-566.
- Ikeo K, Ishi-i J, Tamura T, Gojobori T, Tateno Y: **CIBEX: center for information biology gene expression database.** *C R Biol* 2003, **326(10-11)**:1079-1082.
- Parkinson H, Kapushesky M, Shojatalab M, Abeygunawardena N, Coulson R, Farne A, Holloway E, Kolesnykov N, Lilja P, Lukk M, et al.: **ArrayExpress – a public database of microarray experiments and gene expression profiles.** *Nucleic Acids Res* 2007:D747-750.
- Brazma A, Parkinson H: **ArrayExpress service for reviewers/editors of DNA microarray papers.** *Nat Biotechnol* 2006, **24(11)**:1321-1322.
- Nucleic Acids Research Web Server Issue 2008.**
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, et al.: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5(10)**:R80.
- MGED Ontology** [<http://mged.sourceforge.net/ontologies/index.php>]
- Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, et al.: **Minimum information about a microarray experiment (MIAME)-toward standards for microarray data.** *Nat Genet* 2001, **29(4)**:365-371.
- Spellman PT, Miller M, Stewart J, Troup C, Sarkans U, Chervitz S, Bernhart D, Sherlock G, Ball C, Lepage M, et al.: **Design and implementation of microarray gene expression markup language (MAGE-ML).** *Genome Biol* 2002, **3(9)**:RESEARCH0046.
- Rayner TF, Rocca-Serra P, Spellman PT, Causton HC, Farne A, Holloway E, Irizarry RA, Liu J, Maier DS, Miller M, et al.: **A simple spreadsheet-based, MIAME-supportive format for microarray data: MAGE-TAB.** *BMC Bioinformatics* 2006, **7**:489.
- Stoeckert CJ Jr, Causton HC, Ball CA: **Microarray databases: standards and ontologies.** *Nat Genet* 2002, **32(Suppl)**:469-473.
- Ball C, Brazma A, Causton H, Chervitz S, Edgar R, Hingamp P, Matase JC, Icahn C, Parkinson H, Quackenbush J, et al.: **An open letter on microarray data from the MGED Society.** *Microbiology* 2004, **150(Pt 11)**:3522-3524.
- Shendure J, Mitra RD, Varma C, Church GM: **Advanced sequencing technologies: methods and goals.** *Nat Rev Genet* 2004, **5(5)**:335-344.
- Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nat Rev Genet* 2009, **10(1)**:57-63.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al.: **Genome-wide maps of chromatin state in pluripotent and lineage-committed cells.** *Nature* 2007, **448(7153)**:553-560.
- Sansone SA, Rocca-Serra P, Brandizi M, Brazma A, Field D, Fostel J, Garrow AG, Gilbert J, Goodsaid F, Hardy N, et al.: **The first RSBI (ISA-TAB) workshop: "can a simple format work for complex studies?"**. *Omics* 2008, **12(2)**:143-149.
- Minimum Information about a high-throughput Sequencing Experiment** [<http://www.mged.org/minseq/>]
- Cochrane G, Akhtar R, Aldebert P, Althorpe N, Baldwin A, Bates K, Bhattacharyya S, Bonfield J, Bower L, Browne P, et al.: **Priorities for nucleotide trace, sequence and annotation data capture at the Ensembl Trace Archive and the EMBL Nucleotide Sequence Database.** *Nucleic Acids Res* 2008:D5-12.
- Hermida L, Schaad O, Demougin P, Descombes P, Primig M: **MIMAS: an innovative tool for network-based high density oligonucleotide microarray data management and annotation.** *BMC Bioinformatics* 2006, **7**:190.
- Multomics Information Management and Analysis System** [<http://mimas.vital-it.ch/>]. <http://mimas.genouest.org/>
- Buschmann F, Meunier R, Rohnert H, Sommerlad P, Stahl M: **Pattern-Oriented Software Architecture Volume 1: A System of Patterns.** Volume 1. John Wiley & Sons Inc; 1996.
- Johnson R, Woolf B: **"Type object" in "Pattern languages of program design 3"**. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA; 1997.
- Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, Chung SS, Demougin P, Gattiker A, Moore J, Patard JJ, Wolgemuth DJ, Jegou B, et al.: **The conserved transcriptome in human and rodent male gametogenesis.** *Proc Natl Acad Sci USA* 2007, **104(20)**:8346-8351.
- Gonzalez de Aguilar JL, Niederhauser-Wiederkehr C, Halter B, De Tapia M, Di Scala F, Demougin P, Dupuis L, Primig M, Meininger V, Loeffler JP: **Gene profiling of skeletal muscle in an amyotrophic lateral sclerosis mouse model.** *Physiol Genomics* 2008, **32(2)**:207-218.
- Schlecht U, Erb I, Demougin P, Robine N, Borde V, Nimwegen E, Nicolas A, Primig M: **Genome-wide expression profiling, in vivo DNA binding analysis, and probabilistic motif prediction reveal novel Abf1 target genes during fermentation, respiration, and sporulation in yeast.** *Mol Biol Cell* 2008, **19(5)**:2193-2207.
- Spieß AN, Feig C, Schulze W, Chalmel F, Cappallo-Obermann H, Primig M, Kirchhoff C: **Cross-platform gene expression signature of human spermatogenic failure reveals inflammatory-like response.** *Hum Reprod* 2007, **22(11)**:2936-2946.
- Ouest-genopole Bioinformatics Platform** [<http://genouest.org/>]
- Ouest-genopole** [<http://www.ouest-genopole.org/>]
- Morrison N, Wood AJ, Hancock D, Shah S, Hakes L, Gray T, Tiwari B, Kille P, Cossins A, Hegarty M, et al.: **Annotation of environmental OMICS data: application to the transcriptomics domain.** *Omics* 2006, **10(2)**:172-178.
- BioArray Software Environment (BASE)** [<http://base.thep.lu.se>]
- Maurer M, Molidor R, Sturn A, Hartler J, Hackl H, Stocker G, Prokesch A, Scheideler M, Trajanoski Z: **MARS: microarray analysis, retrieval, and storage system.** *BMC Bioinformatics* 2005, **6(1)**:101.
- Marzolf B, Deutsch EW, Moss P, Campbell D, Johnson MH, Galitski T: **SBEAMS-Microarray: database software supporting genomic expression analyses for systems biology.** *BMC Bioinformatics* 2006, **7**:286.
- Tomlinson C, Thimma M, Alexandrakis S, Castillo T, Dennis JL, Brooks A, Bradley T, Turnbull C, Blaveri E, Barton G, et al.: **MiMiR – an integrated platform for microarray data sharing, mining and analysis.** *BMC Bioinformatics* 2008, **9**:379.
- Hancock D, Wilson M, Velarde G, Morrison N, Hayes A, Hulme H, Wood AJ, Nashar K, Kell DB, Brass A: **maxdLoad2 and maxd-Browse: standards-compliant tools for microarray experimental annotation, data management and dissemination.** *BMC Bioinformatics* 2005, **6**:264.

41. Brazma A, Kapushesky M, Parkinson H, Sarkans U, Shojatalab M: **Data storage and analysis in ArrayExpress**. *Methods Enzymol* 2006, **411**:370-386.
42. Deutsch EW, Ball CA, Berman JJ, Bova GS, Brazma A, Bumgarner RE, Campbell D, Causton HC, Christiansen JH, Daian F, et al.: **Minimum information specification for in situ hybridization and immunohistochemistry experiments (MISFISHIE)**. *Nat Biotechnol* 2008, **26(3)**:305-312.
43. Simon R, Mirlacher M, Sauter G: **Tissue microarrays**. *Biotechniques* 2004, **36(1)**:98-105.
44. Taylor CF, Paton NW, Lilley KS, Binz PA, Julian RK Jr, Jones AR, Zhu W, Apweiler R, Aebersold R, Deutsch EW, et al.: **The minimum information about a proteomics experiment (MIAPE)**. *Nat Biotechnol* 2007, **25(8)**:887-893.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

