

Atlas of Genetics and Cytogenetics in Oncology and Haematology, an Interactive Database

Jean-Loup Huret^{1,*}, Sylvaine Le Minor², Franck Dorkeld², Philippe Dessen² and Alain Bernheim³

¹Genetics, University Hospital, F-86021 Poitiers, France, ²Service de Bioinformatique, UMS 825 CNRS/SC13 INSERM, 7 rue G Môquet, 94801 Villejuif, France and ³Cytogénétique et Génétique Oncologiques, Institut Gustave Roussy PR2, 94805 Villejuif, France

Received October 7, 1999; Accepted October 12, 1999

ABSTRACT

The 'Atlas of Genetics and Cytogenetics in Oncology and Haematology' (<http://www.infobiogen.fr/services/chromcancer>) is a database devoted to chromosome abnormalities in cancer, cancer-prone diseases and genes involved in cancer. Information presented in each page is concise and updated. This database is made for and by: cytogeneticists, molecular biologists, clinicians in oncology and in haematology, and pathologists, who are encouraged to contribute.

INTRODUCTION

Aim

By the end of September 1999, more than 240 sub-types of leukaemias, defined by a given chromosome aberration, often with a clinical and prognostic significance, were recorded by us. It is no longer possible for anyone in the scientific community to know all the chromosome anomalies that may be found in cancer, with their clinical and biological characteristics. It is therefore a necessity for workers to have at their disposal a database with a census of all relevant items, a review and summary on each entity, and updated information.

The 'Atlas of Genetics and Cytogenetics in Oncology and Haematology' (<http://www.infobiogen.fr/services/chromcancer>) is an in-progress database, freely accessible on the Internet, devoted to cytogenetics in cancer. The Atlas includes data on cytogenetics and clinical entities in cancer, on genes, and on cancer-prone diseases. It is made for and by researchers and clinicians in: cytogenetics, molecular biology and genetics in general, oncology, haematology, pathology, cytology and other related specialities. Colleagues are encouraged to contribute. The Atlas is also a useful tool for Teachers and Students.

Organisation

The Atlas is being developed in the frame of INFOBIOGEN by the Bioinformatics unit (<http://www.infobiogen.fr>), the CHU POITIERS (<http://www.chu-poitiers.fr>), under the supervision of the GFCO (French Group for the study of Cytogenetics in

Oncology). Scientific societies in the fields under study are welcome as informal partners.

STRUCTURE OF THE DATABASE

Cards

The purpose is to summarize and review entities in the following fields: (i) genes, (ii) cytogenetic/clinical entities in haematology, (iii) solid tumours: cytogenetic/clinical entities, and (iv) cancer-prone diseases.

Entities are settled relevant according to the following criteria.

- Genes: hybrid gene found in at least one cancer case, or gene amplification or homozygous deletion found in a significant subset of cases in a given cancer-type.
- Chromosome structural aberrations: found in at least two cases in a given cancer, or in only one case if the hybrid gene is also described.
- Cancer-prone disease: significant aggregation of a cancer-type found within families, or a patient with a germline mutation (first allele) and a somatic mutation in the tumour (second allele) of a gene known to be implicated in this tumour type.

Cards are concise and evolve with further improvements and updates from various contributions. They present links to other databases, limited to the most relevant. Lists of entries and indexes facilitate the information search.

Contributors are experts in the fields under study. Guidelines for contributions can be found in the 'how to contribute' section. Contributions are peer reviewed before acceptance. Up to now, about 60 researchers from various countries in Europe and from the US have sent at least one contribution.

Aside

- Deep insights are longer papers focusing on a specific aspect when of particular interest.
- About 80 online journal addresses for content and/or abstract screening can be consulted.
- Introductory items may be used for teaching purposes. We are aware that the Atlas is already used in some University Medical Schools.
- Scientific societies are also encouraged to send us a page explaining what they are and how to contact them. They can advertise freely for their congresses.

*To whom correspondence should be addressed. Tel: +33 5 49 44 45 46; Fax: +33 5 49 44 41 31; Email: jlhuret@infobiogen.fr

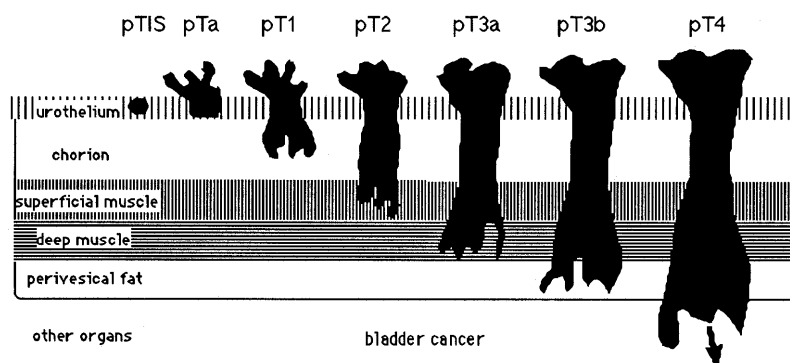


Figure 1. Staging in bladder cancer.

- A large and up-to-date catalogue of database addresses should help workers to save time when accessing relevant information. The Atlas is a link to other databases such as the probe resources and the other various actors of the genome program. Special attention has been given to offering an up-to-date list of routes to mapping resources (often sorted by chromosome). This is of particular importance, as cytogenetic techniques are dramatically improving, making cytogenetics a very powerful tool for research and diagnostic purposes.

- All the information can be accessed by a user-friendly Web interface. The management of all data is performed by Perl scripts, to maintain integrity between all items and especially the construction of indexes and generation of Web pages.

CONTENTS OF THE CARDS

The content of each card has been structured and enriched during the last year.

- Cards on genes (78 entries) include data on: contributors; name(s) of the gene, location; DNA/RNA description; protein description; mutations; where this gene is implicated; diseases, with the prognosis and data on oncogenesis; a selected bibliography with hyperlinks to MEDLINE abstracts (<http://www.ncbi.nlm.nih.gov/pubmed/>); and other external links towards: GeneCards (1), the best cross-roads, GDB (2) for the necessary nomenclature, GenBank (3) where the DNA is described, SWISS-PROT (4), which give extensive data concerning the protein, HGMD (5) for the screening of mutations, OMIM (<http://www.ncbi.nlm.nih.gov/omim/>) for clinical data on inherited diseases.

- Cards on cytogenetic entities and on clinical entities in haematology (92 entries), cards on solid tumours (16 entries) include data on: contributors; name of the disease, with clinics, pathology description, treatment and prognosis; cytogenetics; genes involved: description and protein description; result of the chromosomal anomaly (e.g. translocations): hybrid gene and fusion protein; a selected bibliography with hyperlinks to MEDLINE and another external link towards Resources for Molecular Cytogenetics (<http://bioserver.uniba.it/fish/rocchi/>).

- Cards on cancer-prone diseases (20 entries) include data on: contributors; name(s) of the disease with inheritance mode/epidemiology; clinics: phenotype and clinics, neoplastic risk, prognosis; cytogenetics: inborn condition, cytogenetics of cancers; other findings (ex: cell metabolism); genes and

protein description; mutations: germinal (inborn condition), somatic (tumorigenesis); a selected bibliography with hyperlinks to MEDLINE; other external links above mentioned, and also, whenever possible, external links towards registries and family associations.

IN PROGRESS

- The completion of the Atlas is in progress, but the number of new items to be included is also regularly increasing.

- Iconography (Fig. 1) will be included, in particular drawings of chromosome aberrations, genes and proteins, photographs of G- and R-banded chromosomes, and photographs on morphology and histology of the tumour.

- An extension will concern the link with data in CancerGene (1103 entries), a relational database of genes involved in cancer which has been developed in our Bioinformatics Unit. In this database, cancer genes are defined by their capacity to initiate and/or promote uncontrolled cell growth, using a positive, tumour-inducing, or negative, tumour-suppressing mechanism. The aim of the database is to provide a catalogue of cellular genes involved in different cancers. This database also includes cellular genes associated with the hereditary predisposition to cancer, viral oncogenes and viral cancer-related genes. A part concerns a bibliographic citations database (20 238 references).

- An informal network for clinical and molecular cytogenetics may take place and could include various databases in complementary fields.

COMMENTS

To our knowledge, the Atlas of Genetics and Cytogenetics in Oncology and Haematology is the first and only open database devoted to cytogenetics in cancer, which also include data on the diseases, with particular attention towards clinical aspects and prognosis, an item strikingly discarded from most reports.

This Atlas is intended to give assistance, as none of us can now cover the entire field of knowledge we should cover. It should help cytogeneticists to comment on their findings for clinical diagnosis; help clinicians for treatment decision in cases with a specific prognosis; help students and teachers in oncology, haematology, genetics and pathology; and help authors with confident reviews and valuable external links.

Such will be achieved by the assistance of all high standard specialists in the fields concerned.

REFERENCES

1. Rebhan,M., Chalifa-Caspi,V., Prilusky,J. and Lancet,D. (1998) *Bioinformatics*, **14**, 656–664.
2. Letovsky,S.I., Cottingham,R.W., Porter,C.J. and Li,P.W.D. (1998) *Nucleic Acids Res.*, **26**, 94–99.
3. Benson,D.A., Boguski,M.S., Lipman,D.J., Ostell,J., Ouellette,B.F., Rapp,B.A. and Wheeler,D.L. (1999) *Nucleic Acids Res.*, **27**, 12–17. Updated article in this issue: *Nucleic Acids Res.*, (2000), **28**, 15–18.
4. Bairoch,A. and Apweiler,R. (1999) *Nucleic Acids Res.*, **27**, 49–54. Updated article in this issue: *Nucleic Acids Res.*, (2000), **28**, 45–48.
5. Krawczak,M. and Cooper,D.N. (1997) *Trends Genet.*, **13**, 121–122.