European Cytogeneticists Association
Register of Unbalanced Chromosome Aberrations (ECARUCA); an online database for rare chromosome abnormalities


Department of Human Genetics, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands
School of Computing Science, Middlesex University, London, UK
Institute of Child Health, University College London, London, UK
Institute of Medical Genetics, University of Zurich, Zurich, Switzerland

Available online 07 November 2005

Abstract

During recent years a considerable improvement in diagnostic techniques has enabled cytogeneticists to find more and smaller chromosomal aberrations. However, accurate clinical knowledge about rare chromosome disorders is frequently lacking, mostly due to a significant decline in publishable cases. On the other hand, there is an increasing demand from parents and physicians for reliable information. In order to improve the quality and the quantity of data available, we designed a new database named the European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA) at http://www.ecaruca.net. This Internet-database contains cytogenetic and clinical data of patients with rare chromosome abnormalities, including microscopically visible aberrations, as well as microdeletions and -duplications. Cases with certain breakpoints collected in the Zurich Cytogenetic Database were transferred to ECARUCA. The advantages of ECARUCA compared to existing sources are that ECARUCA is interactive, dynamic and has long-term possibilities to store cytogenetic, molecular and clinical data. Professionals can login to submit new cases and perform searches in the database through the Inter-
Currently the database contains 1500 unique chromosomal aberrations from almost 4000 patients. A frequent submission of new data ensures the up-to-date quality of the collection. Individual parent accounts allow parents to inform the ECARUCA team about the follow-up of their child. The ECARUCA database provides health care workers with accurate information on clinical aspects of rare chromosome disorders. Additionally, detailed correlations between chromosome aberrations and their phenotypes are of invaluable help in localising genes for mental retardation and congenital anomalies.

Keywords: Rare chromosome aberrations; Online database; Genotype–phenotype correlation

1. Introduction

Up to 0.6% of the general population have an unbalanced chromosome aberration [1]. Depending on the diagnostic methods used and the population studied, microscopically visible chromosomal aberrations occur in 4–28% of mentally retarded patients, making chromosomal abnormalities a major cause of mental retardation [2,3]. Relatively common conditions like Turner syndrome and Down syndrome are clinically well known and recognisable, in contrast to the very limited knowledge on many rare chromosome abnormalities. These rare chromosome disorders have a total incidence of at least 0.07% [4] and with an annual birth rate of slightly more than 7 million in Europe [5], it can be estimated that currently 3000–5000 children with a rare chromosome aberration are born each year on this continent only.

Moreover, the number of more or less unique chromosomal aberrations that are identified is rapidly increasing due to the development of new molecular and cytogenetic techniques such as multiplex ligation-dependent probe amplification (MPLA), multiplex amplifiable probe hybridisation (MAPH) and array-based comparative genomic hybridisation (array CGH) [6–9]. Subtelomere screening by fluorescence in situ hybridisation (FISH) and genetic markers resulted in the identification of submicroscopic subtelomeric rearrangements in approximately 5% of mentally retarded patients [10,11]. More advanced techniques such as genome wide high resolution array CGH, will allow the detection of submicroscopic interstitial deletions and duplications, which consequently will considerably increase the number of detectable chromosomal aberrations. However, while the number of chromosomally defined syndromes is rising, the clinical knowledge about individual syndromes remains limited due to the low number of patients that will actually be published in detail.

The source for clinical information concerning specific chromosomal disorders that is available for clinicians, counsellors, researchers and parents is the (inter)national medical literature. The main sources of information in scientific journals are case reports and occasionally a review discussing a chromosomal syndrome like Wolf–Hirschhorn or Cri du Chat [12,13]. Many reports have been collected in the “Catalogue of Unbalanced Chromosome Aberrations in Man” written by Schinzel [4]. This standard work is based on the literature of the last 30 years and contains around 2000 descriptions of patients with a rare chromosome aberration. In addition, the commercially available Zurich Cytogenetic Database which contains over 7200 cases can be used as a digital resource for information [14].

Although all references mentioned are extremely valuable, the number of reports per individual chromosome aberration remains rather limited. This is mostly due to the declining number of published or publishable papers describing the clinical features of a single patient with a
rare chromosome abnormality. Furthermore, articles illustrating the diagnostic improvements of new cytogenetic techniques usually provide only a limited clinical description of a large number of patients. Moreover, patients with a rare chromosome aberration are mostly diagnosed and described in scientific journals at a (very) young age. Follow-up information in these cases is hard to obtain, resulting in a lack of valuable information.

In conclusion, there is a need for an accessible dynamic database to fulfil the demand for information from clinicians, scientists and parents about rare (sub) microscopic chromosome aberrations. The Internet has been chosen as the medium for establishing the interactive online European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA).

In this article the development of the ECARUCA project and the structure of the database are described.

2. Materials and methods

2.1. Objectives

The objectives of ECARUCA are to improve two supplementary areas: the medical and scientific working fields of rare chromosome aberrations (see Fig. 1).

The ECARUCA project aims to increase knowledge on clinical features in patients with rare chromosome aberrations. Within the medical field, physicians, genetic counsellors and other health care providers are explicitly calling for an improvement of availability of medical and psychosocial information. Due to the absence of detailed knowledge on the clinical features and follow-up in rare chromosome aberrations, parents may currently not receive the optimal clinical information related to the chromosomal aberration of their child, especially concerning complications that occur later in life and information about the achievement of developmental milestones.

Furthermore, contact between parents of children with a similar chromosome aberration has shown to be very valuable and should be facilitated.

For the reliability of the database the accuracy of the cytogenetic studies performed in each case is very important. Therefore ECARUCA has set up a program for less facilitated centres to get free support in further cytogenetic characterisation of patient samples. Through this program, the number of cases in the database will increase. Moreover, exact determination of the breakpoints enables the medical professional to search for clinical information described in similar patients, thereby informing parents more accurately.

<table>
<thead>
<tr>
<th>Medical issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Increase of information on clinical features and complications in patients with (sub)-microscopic chromosome aberrations</td>
</tr>
<tr>
<td>- Encourage and facilitate contact between patients &amp; parents</td>
</tr>
<tr>
<td>- Support less facilitated centres in specifying chromosome aberrations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scientific issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Compile a resource of biological material useful for mapping candidate genes for specific clinical abnormalities and identification of new syndromes</td>
</tr>
<tr>
<td>- Enhance co-operation between centres</td>
</tr>
</tbody>
</table>

Fig. 1. The main objectives of the ECARUCA project.
For scientists, ECARUCA aims to be a resource of information that is useful in localising candidate genes as well as the identification of new syndromes. As (cytogenetic) techniques improve and smaller, unique genomic aberrations are detected, the collective registration of the clinical features occurring in these patients is of utmost importance. Therefore, cooperation between different centres and the exchange of knowledge will be initiated by ECARUCA.

Integrating the medical and scientific objectives should eventually lead to the improvement of quantity and quality of knowledge on rare chromosome aberrations available to both professionals and patients.

2.2. European network

In order to facilitate communication and cooperation between countries, physicians and scientists, a European network has been set up (Fig. 2).

First, a National Coordinator has been assigned in each participating country. The National Coordinator acts as an intermediary between the professionals of his or her country and ECARUCA. The National Coordinator represents ECARUCA during national meetings and informs colleagues on new developments within the project. Furthermore, the National Coordinator can assist and stimulate colleagues to submit new cases to the ECARUCA database. Sometimes, language difficulties can discourage a person to take part in an international project. To lower this threshold, the National Coordinator acts as the first contact person to whom people can turn to for information. Conversely, the Coordinator can inform ECARUCA on important developments on clinical genetics within his or her country. An up-to-date list of

![Fig. 2. Overview of the European Network of the ECARUCA project.](image-url)
all National Coordinators including contact information is available on the ECARUCA web-
site.

Another important role in the European network is fulfilled by the patient organisations. Several support groups in different European countries support the ECARUCA project. Patient organisations play an important role in supplying information to family members of patients with a rare chromosomal aberration. They also offer a contact service between patients and they can provide patient data.

Furthermore, a Microdeletion Research Network composed of several research centres around Europe has been set up. Now and in the future, collaboration between centres working on rare chromosome aberrations will be of great importance. A better understanding of the underlying genetic mechanisms and clinical outcome can only be reached through collective, standardised registration of these unique patients in one database, accessible to all participants.

2.3. Database model and confidentiality

The cytogenetic, molecular and clinical data collected by ECARUCA is stored in a relational database management system. All cases receive a unique case ID number and all data like cytogenetic, clinical and pedigree information are linked to that ID number. The database can be queried about aberrations of chromosome regions according to the ISCN 1995 nomenclature. The design of the ECARUCA data model also allows storage of aberrations at base pair level in order to collect molecular data of patients with a submicroscopic aberration.

The database is situated on a secure server at the Department of Human Genetics in Nijmegen. A web interface, using state of the art Java technology, enables users to view and submit data to ECARUCA safely and promptly on the website.

The security and integrity of the data is not only ensured by the correct implementation of hardware and software solutions, but also at the level of the User, by only allowing account holders to have access to the data. Accounts are granted exclusively to professionals in the field of human genetics and health care, in order to ensure confidentiality and correct interpretation of (clinical) data.

Before submission of data of a new patient, parents or the legal guardian need to sign an informed consent form. This form is stored in the patient file of the submitting centre.

3. Results

3.1. Database contents

Currently, the database contains around 4000 cases with almost 1500 unique aberrations. A large number of cases have been derived from the Zurich Cytogenetic Database, established by A. Schinzel and collaborators. This database contains cytogenetic and clinical data of published cases. Excluded are the three most frequent autosomal trisomies (for chromosomes 13, 18 and 21), and X and Y chromosome aberrations unless combined with an autosomal aberration. Before transferring cases to the ECARUCA database, all cases with uncertain breakpoints were excluded. Breakpoints were considered to be uncertain if the original publication did not provide a complete confirmation of breakpoints in the patient described. Moreover, because of the different data models of the Zurich and ECARUCA databases, all individual aberrations needed to be redefined. In the original database they are stored as text fields, while in ECAR-
UCA they are stored as alphanumerical values, allowing conversion to base pair level. An aberration is defined by the start and end position in base pair running from pter to qter. In this way all molecular and cytogenetic data can be optimally integrated. Currently we are working on the implementation of query possibilities and display of data on the base pair level.

Fig. 3 shows an overview of the number of deletions and duplications per chromosome. The aberrations are distributed over the entire genome and include the following types: deletions, duplications, rings, uniparental disomies, trisomies, triploidies and tetraploidies. The majority are deletions ($N = 2296$), followed by duplications ($N = 1773$). As expected, some chromosomes contain more abnormalities (e.g. 4, 11, 13, 15, 18, 22) than others (e.g. 12, 14, 16, 19), partly due to differences in size and gene density. Moreover, a number of chromosomes are involved in more common aberrations such as the 138 cases of DiGeorge syndrome, with a deletion of 22q11.2.

3.2. Website

The collection and distribution of information on rare chromosome aberrations takes place through the ECARUCA website (www.ecaruca.net). The website consists of two sections: public pages that are freely available to each Internet user and restricted pages containing patient information that are only accessible to account holders.

Via the homepage (Fig. 4), an Internet user can access four main areas: Submit Cases, Query Database, Cytogenetic Verification and Frequently Asked Questions. The menu on the left side of the homepage gives entry to remaining topics such as an overview of the data and a frequently changing introduction of a European patient organisation.

On a monthly basis, an interesting patient that has been submitted to the database will be published on the website. A general description of this patient is given at the public pages, whereas the restricted pages contain detailed cytogenetic and clinical information and illustrative clinical pictures. Cases considered for this topic are selected by the Clinical Database Managers (I.F and D.A.K.) in accordance with the Project Management Board, consisting of
the delegates from the centres in London, Zurich and Nijmegen. Furthermore, professionals are encouraged to bring interesting patients to the attention of ECARUCA.

In addition to the public pages, the following restricted options on the ECARUCA website are available to professional account holders:

- Submit cytogenetic and molecular information.
- Submit clinical information.
- Search by chromosomal aberration.
- Search by clinical feature(s).
- List of all cases submitted by the centre of the account holder.
- List of all participating centres.
- Detailed Case of the Month information.

Furthermore, account holders receive an overview by email of all cases submitted to the ECARUCA database in the past month, thereby providing up-to-date clinical information and the possibility for publications together with colleagues.

A unique feature of the website is the restricted page for parents of children whose data has been submitted to the ECARUCA database. A case-specific parent account is created for all newly submitted cases and sent to the referring clinician. With this account parents have access to the information of their own child in the database and can directly send follow-up information to the Clinical Database Managers.
3.3. Submission of cases

Submission of new cases takes place through a clear online process. This process is composed of the submission of cytogenetic and/or molecular data and, subsequently, the submission of accompanying clinical data.

All new cases are presented to ECARUCA by entering the required cytogenetic and/or molecular information (Fig. 5a). Data that are required include the karyotype, the exact breakpoints, ISCN quality and, if performed, the outcome of molecular techniques used. Furthermore, the name and email address of the clinician involved should be reported in case the cytogenetic submitter and the clinician are not the same person. The data submitted are checked by the Clinical Database Manager and, upon approval, the clinical data can be submitted by means of a separate data registration interface (Fig. 5b). After entering all the data in the interface the User submits the information to the database with a single mouse click. Any additional cytogenetic, molecular data and clinical follow-up information is submitted by email. The Clinical Database Manager carefully reviews all incoming data. Only when no further essential information is expected for a particular case and all data are confirmed to be certain, the case is made available for viewing by other account holders.

3.4. Search functionalities

Account holders have access to query the database either by chromosome aberration or by clinical features according to a select list derived from the Winter–Baraitser London Dysmorphology Database [15].

The option “Search by chromosomal aberration” allows the user to retrieve information about cases that concern a particular region on a chromosome of choice (Fig. 6a). In the example, the search is specified to include all deletions in the region between the bands q21.3 and qter of chromosome 18. The search result provides not only cases for which the aberration is located within the boundaries of the chosen chromosome bands, but also those, which overlap the specified region. The User will see a list of all aberrations present in the database that are present within the region of interest (Fig. 6b). By clicking on the aberration of interest, all features that accompany that aberration are listed. The final result (Fig. 6c) is an overview of major and other clinical features of the total number of cases of that specific aberration. Also, information on average age at last examination and IQ are given together with a list of relevant publications.

The option “Search by clinical feature(s)” allows the user to find all chromosomal aberrations in the database that are associated with specified features. An example for the features platelet abnormalities, ptosis and ventricular septal defect is shown in Fig. 7. The clinical features can be selected from a user-friendly expandable selection “tree”, which is based on the feature list present in the widely known Winter–Baraitser London Dysmorphology Database. The User needs to indicate how many of the chosen features are minimally required to be associated with an aberration for it to be included in the search results. As a result, all types of aberrations for which patients display these clinical features are shown.

This search strategy can provide a possible diagnosis and can in addition help to identify chromosomal regions and candidate genes for specific feature(s).
In case a professional has specific interest in a particular aberration, he or she can contact the Clinical Database Manager in order to receive more detailed information on an individual case.

3.5. Cytogenetic verification

Another important component of the ECARUCA project is a collaboration between centres in Europe and the Mediterranean that provide cytogenetic, molecular and, if needed, clinical diagnostics on cases with a rare chromosomal aberration or cases with incomplete investigations.

In a number of genetic centres, especially in Eastern Europe, the financial or technical possibilities to perform high quality cytogenetic analyses are limited. As a consequence, breakpoints cannot be determined accurately. Because ECARUCA aims to be a reliable, high-stan-
dard database, these cases cannot be entered into the database without additional studies, resulting in a loss of a potentially high volume of interesting data. This may subsequently contribute to the growing gap between laboratories in Western Europe and those in the less developed countries.

In order to prevent the issue outlined above, less facilitated centres can receive assistance with the verification of cytogenetic data. Material can be sent to the Institute of Medical Genetics in Zurich, where the exact breakpoints will be determined. Subsequently, the case will be entered in the ECARUCA database.

In the year 2004, the samples of 57 patients originating from 10 different countries have been investigated. Techniques used include routine karyotyping, reverse painting, micro satellite analysis, chromosome micro dissection, FISH and array CGH. For most of these cases, the advanced cytogenetic analysis with high-resolution techniques resulted in a more precise delineation of the breakpoints. Subsequently, in a number of cases the initial designation of the
breakpoints had to be revised, thereby having implications for karyotype–phenotype correlation [16,17].

The correct determination of the chromosome aberration is not only important for the reliability of the ECARUCA database, but also has implications for clinical, counselling and scientific research activities.

4. Discussion

4.1. Future aims

The main feature of the ECARUCA database is its interactivity: users can query the database, and at the same time they can submit cases in order to increase the number of cases in the database, making it more powerful and informative. Submitting an unpublished case to the database will create a platform for such rare cases to be viewed by colleagues who might themselves have observed a similar case, leading to joint publications. At present, almost 4000 patients are collected in the database. In Europe alone, an annual number of 3000–5000 children carrying a rare chromosome aberration are born. Therefore it is expected that a considerable number of new cases coming from all over the world will be submitted to the database each year. Furthermore, by reviewing the medical literature on case reports and other publications, a supplemental number can be included every year.
The functionality of the ECARUCA database and the website are constantly being improved to meet the needs of its users. We try to make the submission procedure as complete as possible and simultaneously less time consuming. Furthermore, the adaptation of the database regarding the implementation of a search and entering facility at base pair level will ensure that the ECARUCA database can accommodate future techniques that have much higher resolution than that of the routine cytogenetic techniques presently used. New high-tech methods like array CGH are currently used in a limited number of laboratories, but the general expectation is that this technique will become a standard diagnostic method in the near future. The possibility of collecting aberrations at the level of base pairs will have positive implications for the scientific use of the ECARUCA database. First of all, it will become easier for non-clinical professionals to interpret and use the available data. Secondly, scientists, notably in the field of molecular genetics, require a standardised way of storing data at the submicroscopic level, which will be provided by ECARUCA. Storing of molecular data generated by high quality experiments in a standardised format greatly enhances the international search for new genes.

The easy and free access to the database for doctors and scientists is a good basis for collaboration in genotype–phenotype studies. The fast progress in the development of new techniques provides an increasing knowledge on genetic abnormalities. However, only by understanding the clinical consequences of these genetic changes can progress be made towards gene identification and ultimately perhaps diagnostic interventions.

A limited other Internet sources are available to look for information on rare chromosome aberrations, e.g. Decipher and the Mendelian Cytogenetics Network. However, the first database only collects submicroscopic aberrations while the second one collects and provides information on balanced chromosomal rearrangements. Therefore, the ECARUCA database complies with the need to fill this gap.

In the future, ECARUCA will also focus on the follow-up of patients present in the database. This will give physicians more insight into how their patient will most likely develop and whether there are any specific new motor skills or physical abnormalities that need special attention.

In addition to the expected increase in knowledge among professionals, we wish to develop correct, understandable and clear information for parents and other interested parties. This should comprise a textual summary of the main clinical and developmental aspects of the aberrations included in the database. Since it would be impossible to include a description of all aberrations, we will restrict the information to a certain extension of chromosome regions or to more frequently arising small aberrations. The physician in attendance can retrieve more specific information that matches the exact deleted or duplicated region in the patient from the database.

4.2. Concluding remarks

The Internet era has given rise to all kinds of new possibilities in information processing. The Internet is especially helpful in finding information about unusual matters like rare chromosome aberrations. Until the existence of ECARUCA, however, an Internet search on a particular rare chromosome disorder usually yielded only a couple of websites with a limited amount of information, sometimes out-of-date and never including follow-up data of patients. Although professionals have the opportunity to look up information in medical journals, books
and other sources, these have the limitation, except for review articles, that only one or two patients are described independently and not the frequency of clinical features occurring in a group of patients with the same aberration. Furthermore, follow-up data of published patients is hard to retrieve.

Altogether, ECARUCA can be considered to be a new system that meets the needs of scientists, physicians and patients and their family members involved with rare chromosomal aberrations, in an interactive way.

Acknowledgements

Funding was obtained from the Fifth Framework Program of the European Union entitled “Quality of Life and Management of Living Resources” (project number QLRI-CT-2002-02746) and the Netherlands Organisation for Health Research and Development (BBAAdV, ZonMW 907-00-058; DAK, ZonMW 920-03-338).

References