The MetaFMF website: a high quality tool for meta-analysis of FMF

Denis Pugnère*, Manuel Ruiz1, Cyril Sarrauste de Menthière, Benjamin Masdoua, Jacques Demaille and Isabelle Touitou2

Institute of Human Genetics—CNRS UP R1142, 141, rue de la Cardonille, 34396 Montpellier Cedex 5, France, 1CIRAD AMIS, Biotrop, UMR 1096—TA40/03, Avenue Agropolis, 34398 Montpellier Cedex 5, France and 2Laboratory of Genetics, A de Villeneuve Hospital, 34295 Montpellier Cedex 5, France

Received August 13, 2002; Accepted October 2, 2002

ABSTRACT

We present here the MetaFMF database (freely accessible at http://fmf.igh.cnrs.fr/metaFMF/index_us.html) that attempts to gather and unify, in a common resource, data on phenotype–genotype correlation in familial Mediterranean fever (FMF). A single accession form, including a large number of quality controls, has been implemented such that data, collected worldwide, are included in an homogeneous manner. The inclusion criterion has the objective to avoid interpretational bias: patients will be included only if they bear at least two mutations. The clinical form has been set up by an International editorial board (12 FMF expert centres), which guarantees the validity of the data. Data are anonymous and submitted by a secure interface, in which the researcher is logged in with a specific ID and password. A pilot study on 211 patients has shown the feasibility and relevance of this project. We anticipate that the use of MetaFMF will enable reliable assessment of phenotype–genotype correlations in FMF, and define a set of severe versus mild mutations/genotypes. It should also highlight reasons for previous inconsistencies in such correlations.

INTRODUCTION

Familial Mediterranean fever (FMF) is the prototype of Hereditary Inflammatory disorders. It is transmitted as an autosomal recessive trait. The gene was identified in 1997 (1,2), and, since then, a set of phenotype–genotype correlation features has been reported by different groups in the literature [reviewed in (3)], and during the two first International conferences on FMF (Jerusalem, Israel, September 1997, Antalya, Turkey, May 2000). We decided to create MetaFMF, a website devoted to the study of phenotype–genotype correlation in FMF, to overcome the problems of some previous inconsistent/unexplained results. Among the main questions we treated are: (1) is the M694V homozygous genotype associated with renal amyloidosis? and (2) is E148Q a benign polymorphism or a pathogenic mutation?

Thus this site is quite different from other disease-specific websites, like Cystic Fibrosis (http://www.genet.sickkids.on.ca/cftr/). The metaFMF site is not a mutation database, nor a patient registry or a list of symptoms. The main purpose of the site is to conjointly collect detailed mutational and clinical description of FMF patients. The inclusion data form has been primarily designed to optimize the forthcoming statistical handling of the included data. The first step was the creation of an International Editorial board, which met for the first time on November 3, 2000, in the National Institute of Health (Bethesda, USA). The inclusion criterion, and the demographical/clinical items to be included were defined during this meeting. The inclusion criterion was chosen to avoid interpretational bias: patients will be included only if they bear two mutated MEFV alleles. It was decided that each centre would provide 10–20 patient data to a test metaFMF site, so as to perform a pilot-study, prior to the launching of the definitive project. The pilot study was conducted between June 1st and September 15, 2001, and included 211 patients from 11 of the 12 aforementioned centres. This pilot study allowed improvement of the form and definitively proved the reliability and the straightforwardness of the site.

SITE PRESENTATION

The FMF website is divided into 3 main parts: MetaFMF database, the InFEVER database and the FMF2002 (23–27 September 2002, La Grande Motte, France) congress site. A large part of the MetaFMF site is devoted to allow contributors to enter data in a safe and easy to understand way. It is also possible to search the database through an intuitive interface. The MetaFMF site hosts several automatics and human checks.

*To whom correspondence should be addressed. Tel: +33 499619909; Fax: +33 499619901; Email: denis.pugner@igh.cnrs.fr

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

#2003 Oxford University Press

DOI: 10.1093/nar/gkg097
CONTRIBUTOR SUBSCRIPTION

To access the site for submission, each contributor needs to be registered in the database. The contributor has for this purpose a special form (Web page) which allows him or her to be registered. Upon confirmation by the administrator(s), an identification code will be sent to the contributor, who can then type this identification code and password, to be authorized to submit data into the MetaFMF database.

DATA SUBMISSION PROCESS

The single accession form contains 6 Steps corresponding to different data types for each patient (the main items are presented in Table 1):

Step 1. demographic and familial data (corresponding to 13 fields in the database).

Step 2. genetic data (26 fields), some data are linked with ‘Internet periodic FEVERS’ (Infevers) database (see accompanying paper).

Steps 3–5. clinical data (147 fields).

Step 6. biological data (20 fields).

Step 7 (final), automatically generated data summary.

Most items of the forms work on a Yes/No/Unknown basis entry, to limit lacking data. At Step 7 of the form, researchers and clinicians have the possibility to re-access previous steps to modify or complete submitted data.

DATA QUALITY CONTROL SYSTEMS

Many different control systems have been implemented to ensure metaFMF data coherence. The WWW data submission form includes a large number of instantaneous quality controls developed in ECMAScript (4) language in order to detect mistakes and incoherent choices.

A percentage of completion is automatically calculated and displayed at Step 7 of the submission form. It is calculated as follows: percentage of the 12 main items (Fever, Abdominal signs, Articular signs, Muscular signs, Thoracic signs, Spleenomegalic, Cutaneous signs, Renal complications, Age at onset, Frequency of attacks, Duration of attacks, Colchicine treatment) with answers other than ‘Unknown’. If the percentage of completion is less than 75%, a message is automatically displayed at the top of the data summary page, asking users to complete data. Entries with a percent of completion less than 75% are not validated.

All the submitted entries are checked by the metaFMF site administrator before definitive validation. Different status of data validation have been created: status ‘W’ (Waiting) corresponding to new data waiting for metaFMF administrator validation, ‘M’ (Modified) for updated data having to be validated, ‘I’ (Incomplete) for data with percentage of completion less than 75%, and status ‘A’ (Accepted) for definitively validated data (Fig. 1). Only data in status ‘A’ can be retrieved with the Search module.

The MetaFMF administrator automatically receives an Email after each new submission. An intranet data management site permits administrator to check new submitted entries, modified entries, data status and allows to validate or not the data. This system prevents data corruption within the database and unauthorized data modification. A list of all previous modifications is available allowing to keep track of data evolution.

After control of each patient by the metaFMF site administrator, contributors receive an automatically generated Email with confirmation of data validation or an indication by the metaFMF administrator of detected incoherence in the submitted data.

VIEW AND MODIFICATION MODULES

MetaFMF site contributors can use the View and Modify modules available online, to respectively view and update their personal submitted patient data. With the View module, researchers or clinicians can visualize the list of their submitted patients with a data summary for each patient. With the Modify module, contributors can easily access the different steps of the data submission form to update any previously submitted data. The corresponding changed information is displayed in red in the page allowing easy checking of the modifications.

DATA SEARCHING AND RETRIEVAL

In order to allow contributors and clinicians to compare symptoms from their patients with patients submitted from others contributors, we designed a database search module which initially shows only a subset of fields to be selected. This search form allows them to firstly choose some major symptoms. The search module then displays patients’ data in a tabular format. In this array, the user can then select graphically other symptoms, and each sub-selection extends the search with the new parameter.

DESIGN AND IMPLEMENTATION

The data is managed using the open source MySQL (http://www.mysql.com) relational database management system. An ODBC link with Microsoft Access has been setup to help the validator to track inconsistencies in the database and statistician to analyze the data. The full Web site is driven with Perl/CGI scripts. Due to the large set of data related to each patient and to ease data submission process, we used ECMAScript (4) controls within HTML dynamic pages. In order to be sure that every contributor could submit his patients, nearly all currently available browsers from various types of computers (PC, Mac, Unix) were tested in the submission process: Netscape, Internet Explorer, Mozilla… both old and current versions.

When the human interface was built, one of our specifications was to offer the most comprehensive and intuitive interface possible. The pilot study has proven to be a valuable study to discover what parts of the Web site might be misunderstood.

To enhance both the database and the web site, the developers use an alternate (development) site independent from the public one. With this organization, all enhancements to the site are activated with minimum downtime.
<table>
<thead>
<tr>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
<th>Step 5</th>
<th>Step 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of birth&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Mutation detected on Allele 1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Age at onset</td>
<td>Skin: Attack duration, attack frequency, fever? (Erysipelas like erythema)</td>
<td>Association with other diseases: Behcet, Inflammatory bowel disease, PAN, HSP, Unclassified vasculitis, Spondylarthritis, Rheumatoid arthritis, JLA</td>
<td>Erythrocyte sedimentation rate (ESR): During acute attack During attack free period Association with attacks unknown</td>
</tr>
<tr>
<td>Sex</td>
<td>Attack frequency</td>
<td>Thoracic signs: Attack duration, attack frequency, fever? types (Unilateral pain, Pleuritis, Pericarditis)</td>
<td>Treatment: Colchicine (dose, age at beginning, efficiency, compliancy), other</td>
<td>C-reactive protein (CRP): During acute attack During attack free period Association with attacks unknown</td>
<td></td>
</tr>
<tr>
<td>Country of Residency</td>
<td>Mutation detected on Allele 2&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Attack duration</td>
<td>Scrotal involvement&lt;sup&gt;e&lt;/sup&gt;: Attack duration, attack frequency, fever? (Unilateral pain)</td>
<td>Leucocyte count (WBC): During acute attack During attack free period Association with attacks unknown</td>
<td></td>
</tr>
<tr>
<td>Ancestry</td>
<td>Mutations tested</td>
<td>Relation with menstruations, pregnancy, lactation&lt;sup&gt;d&lt;/sup&gt;: Menstruation and attacks frequently coincide, attacks changed during pregnancy, Attack changed during lactation</td>
<td>Renal involvement: Hematuria, Proteinuria, Leucocyturia, Urine cast, Non amyloid glomerular disease, Nephrotic syndrome, Renal failure, Renal biopsy</td>
<td>Fibrinogene: During acute attack During attack free period Association with attacks unknown</td>
<td></td>
</tr>
<tr>
<td>Consanguinity: degree between parents?</td>
<td>Technique used: Mutation spécifique Screening technique</td>
<td>Location/type of attack: Isolated fever</td>
<td>Amyloidosis, location (Kidney, adrenal, GI, heart, Thyroid . . .)</td>
<td>Splenomegaly</td>
<td></td>
</tr>
<tr>
<td>Familial Case: Is proband? which degree?</td>
<td></td>
<td>Abdominal signs: Attack duration, attack frequency, what sign? (Pain, peritonitis, diarrhea, constipation . . .)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Articular signs: Attack duration, attack frequency, Location? (ankle, elbow, hip, knee, neck, SI, shoulders, wrist)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Muscular signs: Attack duration, attack frequency, type? (Exertional pain, short attacks of myalgia, Fibromyalgia, Protracted myalgia)</td>
<td>Miscellaneous: Aseptic meningitis LP based, Hepatomegaly, Recurrent hyperbilirunemia not Gilbert disease, Gallstone</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Items in bold: more precise data is required behind this item. <sup>a</sup>Automatically transformed into current age before being posted on the search module. <sup>b</sup>If ‘no’ selected, contributors are forwarded directly from Step 2 to Step 7 (summary). <sup>c</sup>Scrolling list derived from the FMF entry of the INFEVERS database. <sup>d</sup>Item only available for females. <sup>e</sup>Item only available for males.
MetaFMF data quality control systems

![Diagram of MetaFMF data quality control systems]

Figure 1. MetaFMF data quality control systems.
DISCUSSION AND FURTHER DEVELOPMENTS

Our objectives for MetaFMF are to provide an easy protocol for data submission, together with a systematic control of data quality. The WWW interfaces linked to a database can be used to gather genetic and clinical data worldwide. Moreover, MetaFMF is a unique tool to create a critical mass of data on rare mutations. This information is highly valued for medical and clinical research. The final functional version of the site was launched in February 2002. Data from more than 2100 patients have already been submitted to date (August 2002). We anticipate that the use of MetaFMF will provide reliable knowledge about phenotype–genotype correlations in FMF, and define a set of severe versus mild mutations/genotypes. This should help to discriminate between true mutations and polymorphisms, an example being the controversial debate over E148Q. It should also highlight reasons for the previous inconsistencies (for example, differences in the mode of recruitment of patients and in statistical methodology used, variable clinical criteria, late treatment, existence of modifiers and/or environmental factors). Finally, this project should provide a useful picture of the differences in genotype distribution between countries, including those not classically affected by FMF.

ACKNOWLEDGEMENTS

We want to thank the International Editorial board for its inestimable help in the elaboration of the data entry form: Dr Eldad Ben-Chétrit (Israel), Dr David Booth (Great Britain), Dr Daniel Cattan (France), Dr Daniel Kastner (USA), Dr Avi Livneh (Israel), Dr Hassan Majeed (Jordan), Dr Huri Ozdogan (Turkey), Dr Seza Ozen (Turkey), Dr Tamara Sarkisian (Armenia), Dr Isabelle Touitou (France), Dr Mehmet Tunca (Turkey), Dr Fatos Yalcinkaya (Turkey). We thank Kabir Kbiri for his participation in the creation of the MetaFMF website. This project is supported by the European Community.

CITING MetaFMF

If you use the MetaFMF database as a tool for your published research, we ask that this paper be cited.

REFERENCES