The need for comprehensive, standardized phenotyping in the era of genome-wide sequencing

Ada Hamosh, MD, MPH
Scientific Director, OMIM®
Clinical Director, Johns Hopkins University
McKusick-Nathans Institute of Genetic Medicine
Disease Classification

• OMIM: Genetic Traits, Disorders, Diseases
  – Most rare, some not (e.g. cancers)
  – In general, splits disease by molecular basis
  – Fully mapped (and linked) to Orphanet

• Orphanet
  – Rare diseases, not all genetic
  – Organized clinically
  – Fully mapped (and linked) to OMIM
Why Make a Molecular Diagnosis?

• Ends the diagnostic odyssey
• Allows for accurate prognosis and expands reproductive options
• May allow for tailored treatments
• Allows for accurate testing of at risk relatives
  – Eliminates need for screening for those not at risk
Evolution of Molecular Diagnostics:

• Consider a diagnosis (disease), send blood for a single gene mutation (e.g. sickle cell disease) or screen exons, or sequence a single gene (exons and flanking splice sites)

• Then, consider a diagnosis, and send for a panel of genes that are implicated (e.g. Long QT syndrome [n=13+] or hypertrophic cardiomyopathy [n=20+])
OMIM®

Online Mendelian Inheritance in Man®
An Online Catalog of Human Genes and Genetic Disorders
Updated 12 April 2013

Search OMIM

Sample Searches

Advanced Search: OMIM, Clinical Synopses, OMIM Gene Map

McKUSKIE-NATHAN INSTITUTE OF GENETIC MEDICINE

JOHNS HOPKINS UNIVERSITY

National Human Genome Research Institute

NOTE: OMIM is intended for use primarily by physicians and other professionals concerned with genetic disorders, by genetics researchers, and by advanced students in science and medicine. While the OMIM database is open to the public, users seeking information about a personal medical or genetic condition are urged to consult with a qualified physician for diagnosis and for answers to personal questions.

OMIM® and Online Mendelian Inheritance in Man® are registered trademarks of the Johns Hopkins University.

Copyright © 1966-2012 Johns Hopkins University.
About 7000 known disorders cataloged in OMIM

OMIM Morbid Map Scorecard (Updated 12 April 2013):

- Number of phenotypes* for which the molecular basis is known: 4,874
- Number of genes with phenotype-causing mutation: 2,952

* Phenotypes include single-gene mendelian disorders, traits, some susceptibilities to complex disease (e.g., CFH and macular degeneration, 134370.0008), and some somatic cell genetic disease (e.g., FGFR3 and bladder cancer, 134934.0013)
Growth of Gene-Phenotype Relationship
5 March 2013

Source: Online Mendelian Inheritance in Man, Morbid Anatomy of the Human Genome
How Many Unknown (unrecognized) Disorders?

• Currently, about 50% of patients seen in a genetics clinic are not given a specific diagnosis
  – Some with non-syndromic intellectual disability and/or autism
  – Some with clear multi-organ system involvement, but no clear diagnosis
Phenotyping

• Describe the features of a disease
  – To enable a diagnosis (especially of a rare disease)
  – To distinguish between similar disorders
  – To enable genotype-phenotype correlations
Controlled Vocabulary

• Humans don’t need a controlled vocabulary
• Computers do
• The OMIM clinical synopses are (and will remain) written for humans to read
• The Human Phenotype Ontology was developed by Peter Robinson to make an ontology from the OMIM clinical synopses (for computational uses)
Phenotyping Efforts

- Aimed at describing features of a disease
  - London Dysmorphology Database (LDDDB)
  - POSSUM
  - Orphanet terminology
  - HPO
- UMLS (Unified Medical Language System)
- SNOMed-CT
Centers for Mendelian Genomics

• US National Institutes of Health-funded Initiative to identify the causal gene for unsolved Mendelian disorders
  – 3 centers funded:
    • University of Washington (coordinating center)
    • Yale University
    • Baylor-Johns Hopkins

• International collaborative effort
  – Free sequencing for all appropriate
    • Known mendelian disorders
    • Novel disorders with multiplex families
    • Need cases and families
Whole Exome/Genome Sequencing

• Will see variation throughout the genome
• Need to correlate this variation with the phenotype of the person
• Phenotyping cannot be limited to just the disease, but must be of the whole person
• Truly individualized medicine!
PhenoDB (http://phenodb.net)

- A web-based tool for collection, storage, and analysis of standardized phenotypic information
  - Lab tests, images (photos, x-rays, MRIs, videos, digitized pathology slides, etc.)
  - Family history information
  - Consent forms and ELSI discussion
  - Sample tracking (for the project)
  - Analysis and genomic results
Family Submission: CMG1171 - test@test.edu

View / Consent / Updates
Data required before this family can be submitted: disorder type, inheritance, consent, ancestry, patient sex, patient features.

<table>
<thead>
<tr>
<th>Tracking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your local designation for this family (not shared):</td>
</tr>
<tr>
<td>State:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ownership &amp; Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users authorized to access this submission (email addresses):</td>
</tr>
<tr>
<td>If you have a direct collaborator at Baylor or Hopkins, please add their email address.</td>
</tr>
<tr>
<td>Do you have consent to share medical information:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>We organize samples for consideration of sequencing into three categories, pick best fit:</td>
</tr>
<tr>
<td>○ A Mendelian disorder described in OMIM for which the responsible gene has not been identified (example: 223370, Dubowitz syndrome)</td>
</tr>
<tr>
<td>○ A Mendelian disorder with locus heterogeneity (LH) described in OMIM for which the known responsible gene(s) explain only a fraction of the cases and those accounting for more than 25% have been been ruled out in your case, (example: 192600, cardiomyopathy, familial hypertrophic)</td>
</tr>
<tr>
<td>○ An unknown disorder (not described in OMIM) but with segregation in your family consistent with Mendelian Inheritance</td>
</tr>
<tr>
<td>Presumed Inheritance Pattern: Pick best fit...</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
Disorder Types in PhenoDB

• Known Mendelian Disorder, Molecular Basis Unknown

• Known Mendelian Disorder, Locus Heterogeneity (must have ruled out the genes responsible for $>25\%$ of cases)

• Novel (unknown) Disorder
Need Detailed Phenotyping

• To confirm diagnosis, if known
• To sort unknown cases
• To sort which aspect of the individual’s phenotype relates to which variant(s)
• To compare individuals with the same disease
• To compare individuals with similar and/or different phenotypic features
### Lab Tests:

- Was array CGH or other CNV analysis performed on your patient: [ ] Yes  [ ] No  [ ] Unknown

- Were DNA gene tests performed on your patient (e.g. CFTR, BRCA1, BRCA2, etc...): [ ] Yes  [ ] No  [ ] Unknown

- Was whole exome sequencing done before: [ ] Yes  [ ] No  [ ] Unknown

- Were other important tests performed on your patient: [ ] Yes  [ ] No  [ ] Unknown

### Family & Samples:

<table>
<thead>
<tr>
<th>Family Member</th>
<th>Affected</th>
<th>Sample</th>
<th>Sample Type</th>
<th>Phenotypes</th>
<th>Member ID</th>
<th>Sequenced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>M</td>
<td>F</td>
<td>Yes No Unk.</td>
<td>DNA Blood Fibroblasts Lymphoblasts Other</td>
<td>Add features (optional)</td>
<td>CMG1175_1</td>
</tr>
<tr>
<td>Mother</td>
<td></td>
<td></td>
<td></td>
<td>DNA Blood Fibroblasts Lymphoblasts Other</td>
<td>Add features (optional)</td>
<td>CMG1175_2</td>
</tr>
<tr>
<td>Father</td>
<td></td>
<td></td>
<td></td>
<td>DNA Blood Fibroblasts Lymphoblasts Other</td>
<td>Add features (optional)</td>
<td>CMG1175_3</td>
</tr>
<tr>
<td>Sister</td>
<td></td>
<td></td>
<td></td>
<td>DNA Blood Fibroblasts Lymphoblasts Other</td>
<td>Add features (optional)</td>
<td>CMG1175_4</td>
</tr>
</tbody>
</table>

Unk. = Unknown

- Is the family consanguineous: [ ] Yes  [ ] No

- Ancestry: [ ] Pick best fit  [ ] Ancestry Details (optional): 

- Do you have a pedigree: [ ] Yes  [ ] No
Features: CMG1175_1, Patient, Male – test@test.edu – In progress

Family Member:

Birth decade: Unknown
Age at time of evaluation, years: [ ] months: [ ]
Deceased: [ ]

Do you have permission to share photographs: ○ Yes ○ No
Do you have images: ○ Yes ○ No (You may upload x-rays, CT scans, slides, videos, please remove identifying information)
Were other important tests performed on this family member: ○ Yes ○ No ○ Unknown

Features:

Search
You can select features by navigating the hierarchy below, or you can search for them using the search box below. Selecting a feature from the drop-down menu that appears will automatically select it in the hierarchy. Newly selected features will be **highlighted in yellow**.

Search:

<table>
<thead>
<tr>
<th>Features Selected</th>
</tr>
</thead>
</table>

GROWTH & BUILD:
○ Abnormal ○ Normal ○ Unknown
### Features:

You can select features by navigating the hierarchy below, or you can search for them using the search box below. Selecting a feature from the drop-down menu that appears will automatically select it in the hierarchy. Newly selected features will be highlighted in yellow.

<table>
<thead>
<tr>
<th>Search</th>
<th>Features Selected</th>
<th>OMIM Disorders that Match Selected Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MUSCLE, SOFT TISSUE:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NEUROLOGIC:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NEUROPSYCHIATRIC:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SKIN, NAILS, HAIR:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IMMUNOLOGY:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ENDOCRINE FEATURES:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HEMATOLOGY:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>METABOLIC:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NEOPLASIA:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IN UTERO ABNORMALITIES OF THIS PERSON:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KEY LABORATORY ABNORMALITIES:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

Save Features  |  Save Features and Return to Submission

Note that saving the features may take some time if you are uploading a lot of files.
Phenotype - Family ID: 1027, Member ID: 1027_1, Patient - N/A

Birth decade: [Unknown]
Age at time of evaluation, years: [ ] months: [ ]

Do you have permission to share photographs: Yes / No
Do you have images: Yes / No

You can select features by navigating the hierarchy below, or you can search for them using the search box below. Selecting a feature from the popup that appears will automatically select it in the hierarchy. Newly selected features will be highlighted.
Phenotype - Family ID: BH2280, Member ID: BH2280_1, Patient - N/A

Birth decade: Unknown
Age at time of evaluation, years: __________ months: __________

Do you have permission to share photographs: ☐ Yes ☐ No

Do you have images: ☐ Yes ☐ No

You can select features by navigating the hierarchy below, or you can search for them using the search box below. Selecting a feature from the drop-down menu that appears will automatically select it in the hierarchy. Newly selected features will be highlighted in yellow.

Search: [coloboma]

HEAD AND NECK: Eyes > Structure > Iris > Coloboma
HEAD AND NECK: Eyes > Structure > Retina > Coloboma
HEAD AND NECK: Nose > Structure > Alae nasi > Cleft (aka Alae nasi. Notched, Alae nasi coloboma)
HEAD AND NECK: Periorbital region > Eyelids > Eyelid cleft (aka Eyelid coloboma / Eyelid notched)

VOICE: 
☐ Abnormal ☐ Normal ☐ Unknown

CHEST / THORAX:
☐ Abnormal ☐ Normal ☐ Unknown

CARDIOVASCULAR:
☐ Abnormal ☐ Normal ☐ Unknown

RESPIRATORY:
☐ Abnormal ☐ Normal ☐ Unknown

ABDOMEN:
☐ Abnormal ☐ Normal ☐ Unknown

GENITAL SYSTEM:
☐ Abnormal ☐ Normal ☐ Unknown

URINARY SYSTEM:
☐ Abnormal ☐ Normal ☐ Unknown
Pre-populated result for search on “coloboma”, where iris coloboma has been selected.
### Features:

- **Search**
  - You can select features by navigating the hierarchy below, or you can search for them using the search box below. Selecting a feature from the drop-down menu that appears will automatically select it in the hierarchy. Newly selected features will be highlighted in yellow.

- **Inheritance Pattern:** Autosomal dominant

<table>
<thead>
<tr>
<th>Features Selected</th>
<th>OMIM Disorders that Match Selected Inheritance/Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>- GROWTH &amp; BUILD:</td>
<td>- 100070 – AORTIC ANEURYSM, FAMILIAL ABDOMINAL, 1; AAA1</td>
</tr>
<tr>
<td>Current growth</td>
<td>Inheritance, Misc, Vascular,</td>
</tr>
<tr>
<td>and build &gt;</td>
<td>- 154700 – MARFAN SYNDROME; MFS</td>
</tr>
<tr>
<td>Height &gt; Tall</td>
<td>Inheritance, Growth, Head &amp; Neck, Cardiovascular,</td>
</tr>
<tr>
<td>+</td>
<td>Respiratory, Chest, Abdomen, Skeletal, Skin, nails &amp;</td>
</tr>
<tr>
<td></td>
<td>hair, Laboratory abnormalities, Miscellaneous,</td>
</tr>
<tr>
<td></td>
<td>Molecular basis,</td>
</tr>
<tr>
<td></td>
<td>- 132900 – AORTIC ANEURYSM, FAMILIAL THORACIC 4; AAT4</td>
</tr>
<tr>
<td></td>
<td>Cardiac, Eyes, Inheritance, Lab, Vascular,</td>
</tr>
</tbody>
</table>

#### GROWTH & BUILD:

- Abnormal  
- Normal  
- Unknown

**Current growth and build**

- Abnormal  
- Normal  
- Unknown

**Height**

- Short  
- Tall

- Proportionate  
- Disproportionate, long limbs  
- Disproportionate, long trunk

**Weight**

- Abnormal  
- Normal  
- Unknown

**Birth growth parameters**

- Abnormal  
- Normal  
- Unknown

**Other growth characteristics**

- Abnormal  
- Normal  
- Unknown
CMG Early Results

• New Disease Genes
• Phenotypic expansion
  – Involvement of additional organ systems beyond those previously seen in patients with mutations in gene X and disease Y
  – Fewer problems (didn’t meet diagnostic criteria for disease Y)
• Patients with 2 different disorders at the same time
Orphanet Thesaurus Of Clinical Signs

Files available in XML format.

<table>
<thead>
<tr>
<th>Language</th>
<th>Links</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td><a href="http://www.orpha.net/data/xml/en_product5.xml">http://www.orpha.net/data/xml/en_product5.xml</a></td>
<td>0.91 MB</td>
</tr>
<tr>
<td>French</td>
<td><a href="http://www.orpha.net/data/xml/fr_product5.xml">http://www.orpha.net/data/xml/fr_product5.xml</a></td>
<td>0.93 MB</td>
</tr>
<tr>
<td>Spanish</td>
<td><a href="http://www.orpha.net/data/xml/es_product5.xml">http://www.orpha.net/data/xml/es_product5.xml</a></td>
<td>0.00 MB</td>
</tr>
<tr>
<td>Italian</td>
<td><a href="http://www.orpha.net/data/xml/it_product5.xml">http://www.orpha.net/data/xml/it_product5.xml</a></td>
<td>0.92 MB</td>
</tr>
<tr>
<td>German</td>
<td><a href="http://www.orpha.net/data/xml/de_product5.xml">http://www.orpha.net/data/xml/de_product5.xml</a></td>
<td>0.91 MB</td>
</tr>
</tbody>
</table>

User's guide

For other types of products please contact us through the tab "contact"

Orphanet Thesaurus cross-referenced with other terminologies
Orphanet Thesaurus cross-referenced with other terminologies

Files available in OBO format

<table>
<thead>
<tr>
<th>description</th>
<th>links</th>
<th>size</th>
<th>date</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBO file</td>
<td>ORPHA_Mappings.obo</td>
<td>0.83 MB</td>
<td>08/30/12</td>
</tr>
</tbody>
</table>

- The list of Orpha signs and symptoms used to annotate the diseases, cross-referenced with other nomenclatures: HPO, PhenoDB, LDBB and SNOMED-CT.
International Consortium for Phenotype Terminologies

- Assembled by Segolene Ayme
- HPO, DDD, Elements of Morphology, PhenoDB, LDDB, UMLS, SNOMed-CT, others
- Agree on ~2000 high level terms (with definitions and synonyms)
- Make sure that these are used, mapped and have behind them an ontology
- Will have and need more terms for different reasons, but can map narrow to broad
- WHO and SNOMed-CT are committed to adopting these core terms
Maximize Utility of EHR for Individualized Medicine & Research

• Integrate family history information

• Standardized, searchable phenotypic information
  – Comprehensive, head to toe, with positives and negatives

• Genomic information

• To enable Point-of-Care Alerts and decision support

• And public health improvements
URLs

• To participate in the Centers for Mendelian Genomics Project:
  – http://www.mendelian.org or
  – e-mail: gmendel@mendelian.org

• To participate in the Baylor-Hopkins Center for Mendelian Genomics:
  – http://mendeliangenomics.org

• To download PhenoDB for use in your own project:
  – http://phenodbd.net
Acknowledgements

• National Human Genome Research Institute
• Nara Sobreira, Julie Hoover-Fong, Corinne Boehm, Reid Sutton François Schiettecatte, David Valle
• Colleagues at Johns Hopkins, Baylor College of Medicine and around the world
• Patients and their families