PhenoDB, Variant Analysis Tool and GeneMatcher for Whole Exome/Genome Sequencing Analysis

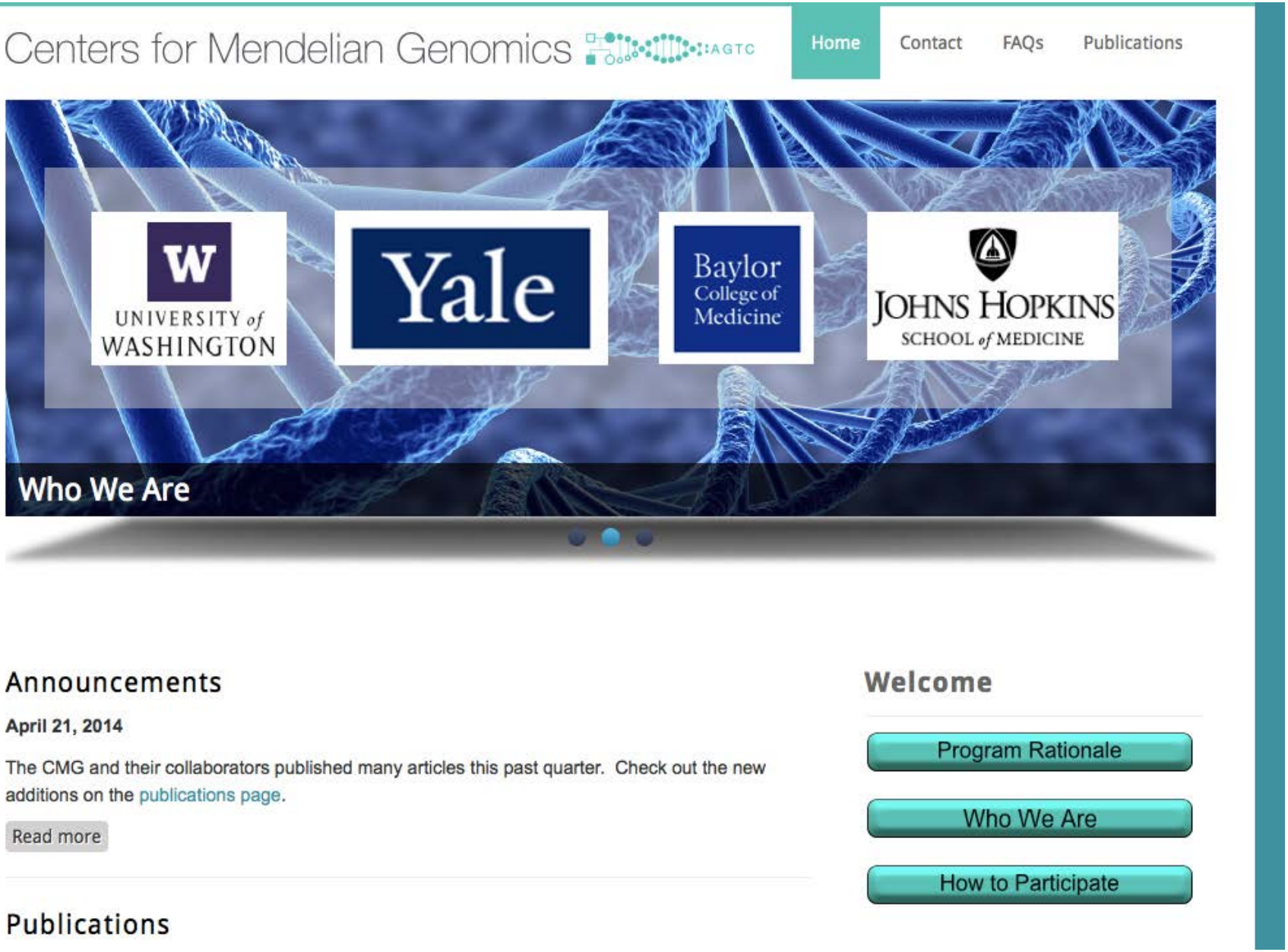
Ada Hamosh, MD & Nara Sobreira,MD, PhD

Johns Hopkins University

McKusick-Nathans Institute of Genetic Medicine

Centers for Mendelian Genomics

- US National Institutes of Health-funded initiative to identify the causal gene for unsolved Mendelian disorders
- International collaborative effort
  - Free sequencing for all appropriate
    - Known mendelian disorders
    - Novel disorders with multiplex families
    - Need cases and families
- The data belong to the submitters
  - Help with analysis, if desired.



BHCMG - http://mendeliangenomics.org

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

PhenoDB: A New Web-Based Tool for the Collection, Storage, and Analysis of Phenotypic Features

Ada Hamosh,[1]* Nara Sobreira,[1] Julie Hoover-Fong,[1] V. Reid Sutton,[2] Corinne Boehm,[1] François Schiettecatte,[3] and David Valle[1]

[1]McKusick-Nathans Institute of Genetic Medicine Johns Hopkins University, Baltimore, Maryland; [2]Department of Molecular & Human Genetics Baylor College of Medicine, Houston, Texas; [3]FS Consulting, Salem, Massachusetts

Hum Mut 34:561, 2013

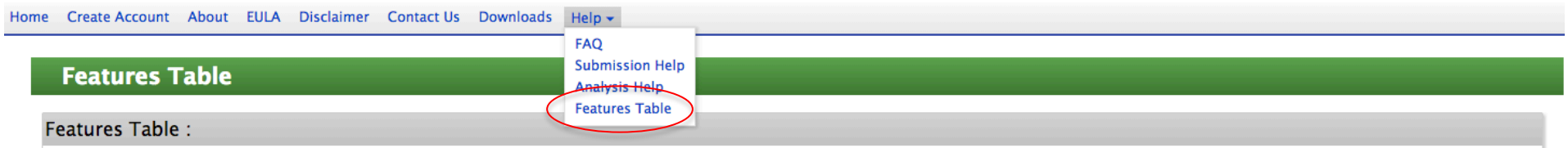
- Rapid and efficient entry of families or cohorts
- Provides unique identifier for each family and subject
- Clinical features based on OMIM Clinical Synopses
- Accepts image data
- Searchable
- Organizes phenotypic features in standard format for easy review
- Variant analysis tool

http://phenodbresearch.net OR http://phenodb.org

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

PhenoDB Terms



Features Table

Features Table :

PhenoDB ID	PhenoDB Feature Path	HPO ID	ICHPT ID Elements of Morphology ID
phenodb:0001	GROWTH & BUILD:		
phenodb:0002	GROWTH & BUILD: Current growth and build		
phenodb:0003	GROWTH & BUILD: Current growth and build > Height	HP:0000002	
phenodb:0004	GROWTH & BUILD: Current growth and build > Height > Short	HP:0004322	T2839
phenodb:0005	GROWTH & BUILD: Current growth and build > Height > Short > Proportionate	HP:0003508	T2840
phenodb:0006	GROWTH & BUILD: Current growth and build > Height > Short > Disproportionate, short limbs	HP:0008873	T2841
phenodb:0007	GROWTH & BUILD: Current growth and build > Height > Short > Disproportionate, short trunk	HP:0008922	
phenodb:0008	GROWTH & BUILD: Current growth and build > Height > Short > Other feature		
phenodb:0009	GROWTH & BUILD: Current growth and build > Height > Tall	HP:0000098	
phenodb:0010	GROWTH & BUILD: Current growth and build > Height > Tall > Proportionate	HP:0011407	T2835
phenodb:2925	GROWTH & BUILD: Current growth and build > Height > Tall > Disproportionate	HP:0001519	T2931
phenodb:0011	GROWTH & BUILD: Current growth and build > Height > Tall > Disproportionate, long limbs		
phenodb:0012	GROWTH & BUILD: Current growth and build > Height > Tall > Disproportionate, long trunk		
phenodb:0013	GROWTH & BUILD: Current growth and build > Height > Tall > Other feature		
phenodb:0014	GROWTH & BUILD: Current growth and build > Weight	HP:0004323	
phenodb:0015	GROWTH & BUILD: Current growth and build > Weight > Underweight	HP:0004325	
phenodb:0016	GROWTH & BUILD: Current growth and build > Weight > Underweight > Isolated (height average or tall)		
phenodb:0017	GROWTH & BUILD: Current growth and build > Weight > Underweight > With short stature		
phenodb:0018	GROWTH & BUILD: Current growth and build > Weight > Underweight > Other feature		
phenodb:0019	GROWTH & BUILD: Current growth and build > Weight > Overweight	HP:0004324	T2825
phenodb:0020	GROWTH & BUILD: Current growth and build > Weight > Overweight > Isolated (height average or short)		
phenodb:0021	GROWTH & BUILD: Current growth and build > Weight > Overweight > With tall stature		
phenodb:0022	GROWTH & BUILD: Current growth and build > Weight > Overweight > Other feature		
phenodb:0023	GROWTH & BUILD: Birth growth parameters		
phenodb:0024	GROWTH & BUILD: Birth growth parameters > Length		
phenodb:0025	GROWTH & BUILD: Birth growth parameters > Length > Small for gestational age (<10%)	HP:0001518	T2843
phenodb:0026	GROWTH & BUILD: Birth growth parameters > Length > Large for gestational age (>90%)	HP:0001520	
phenodb:0027	GROWTH & BUILD: Birth growth parameters > Length > Other feature		
phenodb:0028	GROWTH & BUILD: Birth growth parameters > Weight		
phenodb:0029	GROWTH & BUILD: Birth growth parameters > Weight > Small for gestational age (<10%)	HP:0001518	
phenodb:0030	GROWTH & BUILD: Birth growth parameters > Weight > Large for gestational age (>90%)	HP:0001520	
phenodb:0031	GROWTH & BUILD: Birth growth parameters > Weight > Other feature		
phenodb:0032	GROWTH & BUILD: Birth growth parameters > Head circumference		
phenodb:0033	GROWTH & BUILD: Birth growth parameters > Head circumference > Small for gestational age (<10%)	HP:0011451	
phenodb:0034	GROWTH & BUILD: Birth growth parameters > Head circumference > Large for gestational age (>90%)	HP:0004488	
phenodb:0035	GROWTH & BUILD: Birth growth parameters > Head circumference > Other feature		
phenodb:0036	GROWTH & BUILD: Other growth characteristics		

Submission Page

View / Consent / Updates

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

Tracking :

Your local designation for this family (not shared) :

State : -

Ownership & Access :

Users authorized to access this submission (email addresses) : ◯ View ◯ Edit

If you have a direct collaborator at Baylor or Hopkins,
please add their email address.

Do you have consent to share medical information : ◯ Yes ◯ No

Disorder :

We organize samples for consideration of sequencing into three categories, pick best fit:

- ◯ A Mendelian disorder described in OMIM for which the responsible gene has not been identified (example: 223370, Dubowitz syndrome)
- ◯ 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)
- ◯ An unknown disorder (not described in OMIM) but with segregation in your family consistent with Mendelian Inheritance

Presumed Inheritance Pattern : Pick best fit...

Laboratory Tests Previously Performed

Lab Tests :

Was array CGH or other CNV analysis performed on your patient : (•) Yes ( ) No ( ) Unknown

- Oligo array : ( ) Yes (•) No
- SNP array : (•) Yes ( ) No
- Other array : ( ) Yes (•) No
- Unknown platform : [ ]
- Submit array results report : ( ) Upload ( ) Fax ( ) Mail

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

- Genes screened with negative results : *ACVR1* | *PTPN11* |
- Genes screened with inconclusive results :
- Submit all DNA results reports : ( ) Upload ( ) Fax ( ) Mail

---

Was whole exome sequencing done before : ( ) Yes ( ) No (•) Unknown

---

Were other important tests performed on your patient : ( ) Yes ( ) No (•) Unknown

Family History and Pedigree

Family & Samples :

Family Member	Affected	Sample	Sample Type	Phenotypes	Member ID	Sequenced
Patient ○ M ◉ F	◉ Yes ○ No ○ Unk.	◉ Yes ○ No	☑ DNA ☐ Blood ☐ Fibroblasts ☐ Lymphoblasts ☐ Other DNA sample type : ☑ Blood ☐ Saliva ☐ Fibroblasts ☐ Lymphoblasts ☐ Other	Edit features Diagnosis search	BH2028_1	-
Mother	◉ Yes ○ No ○ Unk.	◉ Yes ○ No	☑ DNA ☐ Blood ☐ Fibroblasts ☐ Lymphoblasts ☐ Other DNA sample type : ☑ Blood ☐ Saliva ☐ Fibroblasts ☐ Lymphoblasts ☐ Other	Edit features Diagnosis search	BH2028_2	Yes
Father	○ Yes ◉ No ○ Unk.	◉ Yes ○ No	☐ DNA ☐ Blood ☐ Fibroblasts ☐ Lymphoblasts ☑ Other saliva	Add features (optional)	BH2028_3	-
Sister	◉ Yes ○ No ○ Unk.	◉ Yes ○ No	☑ DNA ☐ Blood ☐ Fibroblasts ☐ Lymphoblasts ☐ Other DNA sample type : ☑ Blood ☐ Saliva ☐ Fibroblasts ☐ Lymphoblasts ☐ Other	Edit features Diagnosis search	BH2028_4	-
Sister	○ Yes ◉ No ○ Unk.	◉ Yes ○ No	☐ DNA ☐ Blood ☐ Fibroblasts ☐ Lymphoblasts ☑ Other saliva	Add features (optional)	BH2028_5	-
Maternal aunt	◉ Yes ○ No ○ Unk.	○ Yes ◉ No	Sample obtainable: ○ Yes ◉ No ○ Unk.	Edit features Diagnosis search	BH2028_6	-
Maternal male cousin	◉ Yes ○ No ○ Unk.	○ Yes ◉ No	Sample obtainable: ○ Yes ◉ No ○ Unk.	Edit features Diagnosis search	BH2028_7	-
Maternal female cousin	◉ Yes ○ No ○ Unk.	○ Yes ◉ No	Sample obtainable: ○ Yes ◉ No ○ Unk.	Edit features Diagnosis search	BH2028_8	-
Other (fill-in) 2028_6's husband ◉ M ○ F	○ Yes ○ No ◉ Unk.	○ Yes ◉ No	Sample obtainable: ○ Yes ◉ No ○ Unk.	Add features (optional)	BH2028_9	-
Select relation to patient						
	Unk. = Unknown					

Is the family consanguineous : ◉ Yes ○ No

Please show the relationship in the pedigree and describe :

Consanguinity tested : ○ Yes ◉ No

Ancestry : Other Ancestry details (optional) :

Ancestry tested : ◉ Yes ○ No

Confirmed ancestry : See PCA plot, mother in-between CEU and YRI clusters

Do you have a pedigree : ◉ Yes ○ No

Please remove all identifiers from the pedigree & label it with the Member IDs from the table above.

Submit pedigree (please indicate your patient with an arrow) : ◉ Upload ○ Fax ○ Mail

Upload new pedigree file : Choose File no file selected

Pedigree file

Clinical Features Information

View / Edit / Consent / Updates

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	Features Selected
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 :	

GROWTH & BUILD:

Abnormal Normal Unknown

Features :

Search	Features Selected	OMIM Disorders that Match Selected 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.

Search :

GROWTH & BUILD:
Abnormal Normal Unknown

HEAD AND NECK:
Abnormal Normal Unknown

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

SKELETAL:
Abnormal Normal Unknown

MUSCLE, SOFT TISSUE:
Abnormal Normal Unknown

NEUROLOGIC:
Abnormal Normal Unknown

NEUROPSYCHIATRIC:
Abnormal Normal Unknown

SKIN, NAILS, HAIR:
Abnormal Normal Unknown

IMMUNOLOGY:
Abnormal Normal Unknown

ENDOCRINE FEATURES:
Abnormal Normal Unknown

HEMATOLOGY:
Abnormal Normal Unknown

METABOLIC:
Abnormal Normal Unknown

NEOPLASIA:
Yes No Unknown

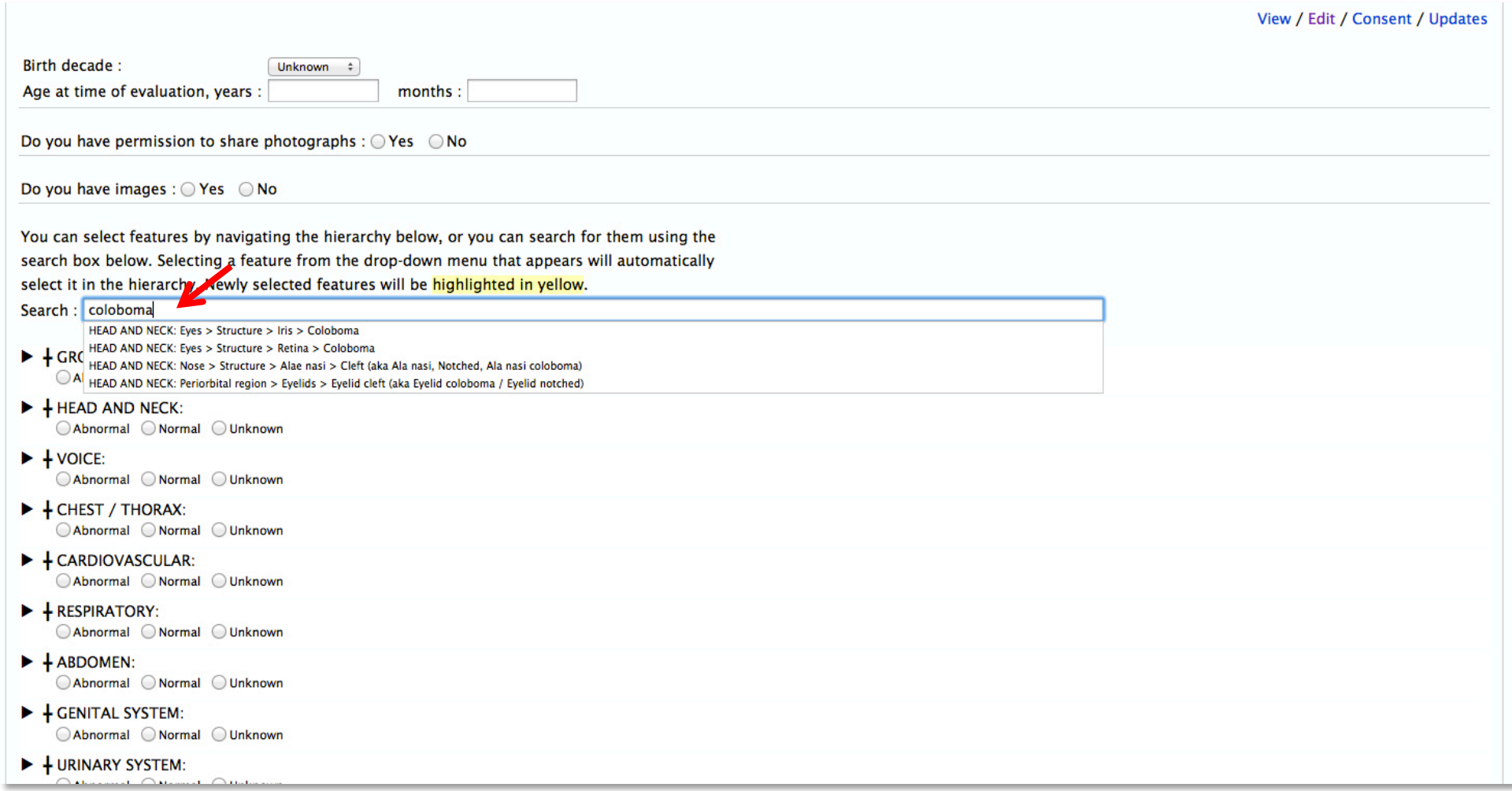
IN UTERO ABNORMALITIES OF THIS PERSON:
Yes No Unknown

KEY LABORATORY ABNORMALITIES:
Yes No Unknown

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.

Entering Features



Features Selected

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.

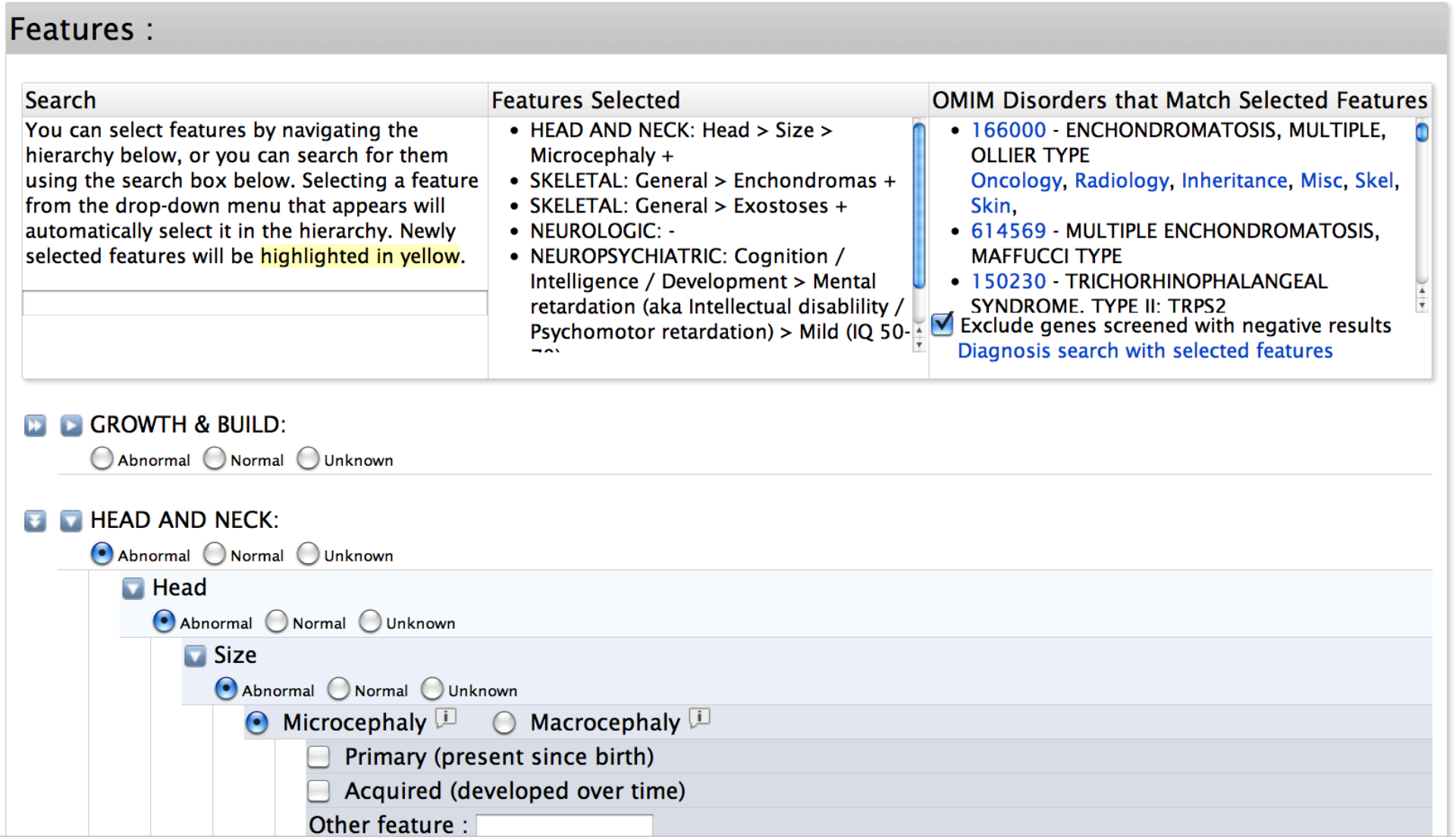
Search :

The feature: "HEAD AND NECK: Eyes > Structure > Iris > Coloboma" was selected in the hierarchy below.

- ▶ ┼ GROWTH & BUILD:
  - ○ Abnormal ○ Normal ○ Unknown
- ▶ — HEAD AND NECK:
  - ◉ Abnormal ○ Normal ○ Unknown
    - ┼ Head
      - ○ Abnormal ○ Normal ○ Unknown
    - ┼ Face
      - ○ Abnormal ○ Normal ○ Unknown
    - — Eyes
      - ◉ Abnormal ○ Normal ○ Unknown
        - ┼ Function
          - ○ Abnormal ○ Normal ○ Unknown
        - — Structure
          - ◉ Abnormal ○ Normal ○ Unknown
            - ┼ General (globe)
              - ○ Abnormal ○ Normal ○ Unknown
            - ┼ Anterior chamber
              - ○ Abnormal ○ Normal ○ Unknown
            - ┼ Cornea
              - ○ Abnormal ○ Normal ○ Unknown
            - ┼ Lens
              - ○ Abnormal ○ Normal ○ Unknown
            - ☐ Epibulbar dermoids
            - ┼ Glaucoma
              - ○ Abnormal ○ Normal ○ Unknown
            - — Iris
              - ◉ Abnormal ○ Normal ○ Unknown
                - ☐ Absent / Aniridia
                - ☐ Bright blue
                - ☑ Coloboma
                - ☐ Hypoplasia

Pre-populated result for search on “coloboma”, where iris coloboma has been selected.

Clinical Features and Possible Diagnosis



Clinical Features

Features :

☑ Show unknown and unaffected members ☑ Show normal and unknown features

Family Members/Features	Patient - Female BH2028_1	Mother BH2028_2	Sister BH2028_4	Maternal aunt BH2028_6	Maternal male cousin BH2028_7	Maternal female cousin BH2028_8	2028_6's husband - Male BH2028_9	Father BH2028_3	Sister BH2028_5
Affected :	Yes	Yes	Yes	Yes	Yes	Yes	Unknown	No	No
HEAD AND NECK: Head > Size > Microcephaly	+		+						
SKELETAL: General > Enchondromas	+	+	+	+	+	+			
SKELETAL: General > Exostoses	+	+	+	+	+	+			
NEUROLOGIC:	-								
NEUROPSYCHIATRIC: Cognition / Intelligence / Development > Mental retardation (aka Intellectual disablility / Psychomotor retardation) > Mild (IQ 50-70)	+		+						

+ = abnormal/yes, - = normal/no, ? = unknown

Download features table as tab-delimited text, Microsoft Excel file.

Diagnosis Suggested

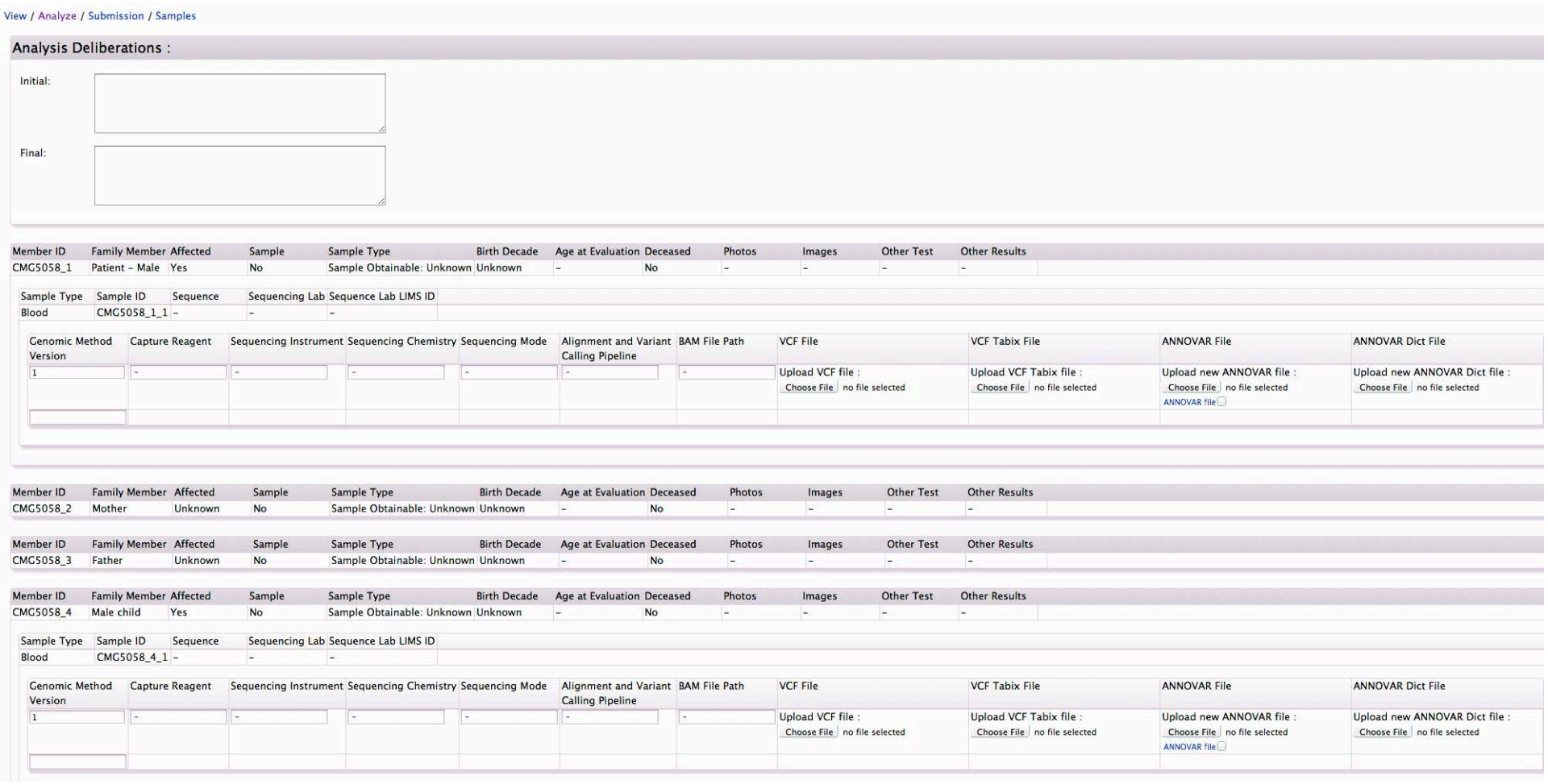
Patient Features MIM Search Result:

166000 - ENCHONDROMATOSIS, MULTIPLE, OLLIER TYPE
614569 - MULTIPLE ENCHONDROMATOSIS, MAFFUCCI TYPE
133700 - EXOSTOSES, MULTIPLE, TYPE I
133701 - EXOSTOSES, MULTIPLE, TYPE II
128300 - EAR EXOSTOSES
150230 - TRICHORHINOPHALANGEAL SYNDROME, TYPE II; TRPS2
133690 - EXOSTOSES WITH ANETODERMIA AND BRACHYDACTYLY, TYPE E
133600 - EXOSTOSES OF HEEL
600209 - EXOSTOSES, MULTIPLE, TYPE III; EXT3
175450 - POLYPOSIS, INTESTINAL, WITH MULTIPLE EXOSTOSES
158345 - MULTIPLE EXOSTOSES WITH SPASTIC TETRAPARESIS
190350 - TRICHORHINOPHALANGEAL SYNDROME, TYPE I; TRPS1
614875 - METAPHYSEAL ENCHONDROMATOSIS WITH D-2-HYDROXYGLUTARIC ACIDURIA
614701 - CORNELIA DE LANGE SYNDROME 4; CDLS4
605946 - METAPHYSEAL DYSPLASIA, BRAUN-TINSCHERT TYPE
613330 - SPONDYLO-MEGAEPIPHYSEAL-METAPHYSEAL DYSPLASIA; SMMD
190351 - TRICHORHINOPHALANGEAL SYNDROME, TYPE III; TRPS3
137800 - GLIOMA SUSCEPTIBILITY 1; GLM1
164210 - HEMIFACIAL MICROSOMIA; HFM
164680 - ONYCHOGRYPOSIS, PEDAL, WITH KERATOSIS PLANTARIS AND COARSE HAIR

Variant Analysis Tool

Nara Sobreira, MD, PhD

PhenoDB analysis tool: Sequence data for analysis



View / Analyze / Submission / Samples

Analysis Deliberations :

Initial:

Final:

Member ID	Family Member	Affected	Sample	Sample Type	Birth Decade	Age at Evaluation	Deceased	Photos	Images	Other Test	Other Results
CMG5058_1	Patient – Male	Yes	No	Sample Obtainable: Unknown	Unknown	–	No	–	–	–	–

Sample Type	Sample ID	Sequence	Sequencing Lab	Sequence Lab LIMS ID
Blood	CMG5058_1_1	–	–	–

Genomic Method Version	Capture Reagent	Sequencing Instrument	Sequencing Chemistry	Sequencing Mode	Alignment and Variant Calling Pipeline	BAM File Path	VCF File	VCF Tabix File	ANNOVAR File	ANNOVAR Dict File
1	-	-	-	-	-	-	Upload VCF file : Choose File no file selected	Upload VCF Tabix file : Choose File no file selected	Upload new ANNOVAR file : Choose File no file selected ANNOVAR file	Upload new ANNOVAR Dict file : Choose File no file selected

Member ID	Family Member	Affected	Sample	Sample Type	Birth Decade	Age at Evaluation	Deceased	Photos	Images	Other Test	Other Results
CMG5058_2	Mother	Unknown	No	Sample Obtainable: Unknown	Unknown	–	No	–	–	–	–

Member ID	Family Member	Affected	Sample	Sample Type	Birth Decade	Age at Evaluation	Deceased	Photos	Images	Other Test	Other Results
CMG5058_3	Father	Unknown	No	Sample Obtainable: Unknown	Unknown	–	No	–	–	–	–

Member ID	Family Member	Affected	Sample	Sample Type	Birth Decade	Age at Evaluation	Deceased	Photos	Images	Other Test	Other Results
CMG5058_4	Male child	Yes	No	Sample Obtainable: Unknown	Unknown	–	No	–	–	–	–

Sample Type	Sample ID	Sequence	Sequencing Lab	Sequence Lab LIMS ID
Blood	CMG5058_4_1	–	–	–

Genomic Method Version	Capture Reagent	Sequencing Instrument	Sequencing Chemistry	Sequencing Mode	Alignment and Variant Calling Pipeline	BAM File Path	VCF File	VCF Tabix File	ANNOVAR File	ANNOVAR Dict File
1	-	-	-	-	-	-	Upload VCF file : Choose File no file selected	Upload VCF Tabix file : Choose File no file selected	Upload new ANNOVAR file : Choose File no file selected ANNOVAR file	Upload new ANNOVAR Dict file : Choose File no file selected

PhenoDB analysis tool: Sequence data for analysis

Final Results :

GWAS array type :

BAM file sent :

Upload genotyping array data (PLINK) file : Choose File no file selected

Upload SNP array report files :

Upload CNVs report file : Choose File no file selected

Upload LOH report file : Choose File no file selected

Upload B_Allele_Freq and LogRratio chromosome plot file : Choose File no file selected

Upload PCA plot file : Choose File no file selected

Upload relatedness check file : Choose File no file selected

Upload .ped File : Choose File no file selected

Upload QC report file : Choose File no file selected

Upload final results file : Choose File no file selected

Save Analysis

Variant Analysis Design

New Analysis :

Member ID	Family Member	Affected	Sample ID - Genomic Version - Lab LIMS ID - ANNOVAR File Name	Include in Analysis
BH2028_1	Patient - Female	Yes		
BH2028_2	Mother	Yes	JH2028_2 - 1 - 200494807 - 200494807@1072257546_annovar_report.txt	☑
BH2028_3	Father	No		
BH2028_4	Sister	Yes		
BH2028_5	Sister	No		
BH2028_6	Maternal aunt	Yes	-	
BH2028_7	Maternal male cousin	Yes	-	
BH2028_8	Maternal female cousin	Yes	-	
BH2028_9	2028_6's husband - Male	Unknown	-	

Run name :

Analysis type

- ✓ Select analysis type...
- Autosomal recessive – Compound heterozygous
- Autosomal recessive – Homozygous
- X-Linked recessive
- Autosomal dominant – New mutation
- Autosomal dominant – Inherited mutation
- Autosomal dominant – Variants

Run Analysis

Refgene gene location : 'exonic', 'exonic splicing', 'splicing'

Exclude if SNP present : dbSNP126 : ☑ dbSNP129 : ☑ dbSNP131 : ☑

Percentage cutoff : 0.01

(1,000 genome and Exome variant server)

Exclude chromosome X : ☐

Indel span : +/- 50 base-pairs

Total depth cutoff :

Dropped variant row numbers :

(Logs the stage when the specified rows are dropped in the analysis process)

Analysis Results Log:

Run name: Supplemental Figure 1
Analysis type: Autosomal recessive - Compound heterozygous
Override, excluding if SNP present in: 'snp126', 'snp131', 'snp129'
Override, precentage cutoff: 0.01
Override, indel exclusion list span: +/- 50 base pairs
Parsing ANNOVAR file for BH2041_1 - Patient - Female - Affected: Yes
Parsed 88,533 variants from the ANNOVAR file '200494814@1072257478.OnBait.recode_ANNOVAR_REPORT.txt'
Parsing ANNOVAR file for BH2041_2 - Mother - Female - Affected: No
Parsed 86,349 variants from the ANNOVAR file '200494815@1072257497.OnBait.recode_ANNOVAR_REPORT.txt'
Parsing ANNOVAR file for BH2041_3 - Father - Male - Affected: No
Parsed 80,263 variants from the ANNOVAR file '200494816@1072257502.OnBait.recode_ANNOVAR_REPORT.txt'
Patient ANNOVAR, including 'het' in 'Genotype', 88,532 variants -> 57,387 variants
Patient ANNOVAR, including 'exonic;splicing', 'exonic', 'splicing' in 'RefgeneGeneLocation', 57,387 variants -> 14,736 variants
Patient ANNOVAR, excluding 'synonymous SNV' in 'RefgeneExonFunction', 14,736 variants -> 6,913 variants
Patient ANNOVAR, excluding if SNP present in 'snp126', 'snp131', 'snp129', 6,913 variants -> 953 variants
Patient ANNOVAR, missing column 'hg19_esp6500_all'
Patient ANNOVAR, missing column 'hg19_esp6500_ea'
Patient ANNOVAR, missing column 'hg19_esp6500_aa'
Patient ANNOVAR, excluding if value is greater than 0.01 in 'hg19_esp6500si_aa', 'hg19_esp6500_all', 'hg19_esp6500_ea', 'hg19_esp6500si_all', 'hg19_esp6500si_ea', 'hg19_esp6500_aa', 953 variants -> 594 variants
Patient ANNOVAR, excluding if value is greater than 0.01 in 'Afalt_1000g2012apr_asn', 'Afalt_1000g2012Apr_eur', 'Afalt_1000g2012Apr_afr', 'Afalt_1000g2012Apr_amr', 'Afalt_1000g2012Apr_all', 594 variants -> 523 variants
Read 429,902 entries from 'snv' exclusion file
Patient ANNOVAR, excluding the 'snv' VariantType on the snv exclusion list, 523 variants -> 458 variants
Read 4,786,664 entries from 'indel' exclusion file
Patient ANNOVAR, excluding the 'indel' VariantType on the indel exclusion list, +/- 50 base pairs, 458 variants -> 365 variants
Patient ANNOVAR, excluding unique value in ' RefgeneGeneName', 365 variants -> 29 variants
Patient ANNOVAR, labeling parents, 28 variants labeled, mother only: 8 variants, father only: 17 variants, mother & father: 3 variants (29 variants total)
Patient ANNOVAR, excluding same parents origin in 'RefgeneGeneName', 29 variants -> 12 variants
Patient ANNOVAR, adding approved symbol data, 12 variants updated (12 variants total)
Patient ANNOVAR, adding OMIM gene data, 2 variants updated (12 variants total)
Patient ANNOVAR, adding mouse model data (MGI), 4 variants updated (12 variants total)
Patient ANNOVAR, adding mammalian phenotype data (MGI), 12 variants updated (12 variants total)
Patient ANNOVAR, adding expression data (Gepsis Tissue), 12 variants updated (12 variants total)
Patient ANNOVAR, adding interaction data (GeneCards), 10 variants updated (12 variants total)
Patient ANNOVAR, adding Gene Ontology links, 12 variants updated (12 variants total)
Patient ANNOVAR, adding Mendel Samples data, 12 variants updated (12 variants total)
Patient ANNOVAR, adding PubMed links, 12 variants updated (12 variants total)
Patient ANNOVAR, adding Emory variant classification, 0 variants updated (12 variants total)
Patient ANNOVAR, adding ClinVar variant classification, 1 variants updated (12 variants total)
Patient ANNOVAR, adding UniProt links, 12 variants updated (12 variants total)
Patient ANNOVAR, adding CCDS data, 12 variants updated (12 variants total)
Patient ANNOVAR, adding CADD data, 12 variants updated (12 variants total)
Final count: 12

Patient Features MIM Search Result:
164280 - FEINGOLD SYNDROME 1; FGLDS1
614756 - CEREBELLAR ATAXIA, NONPROGRESSIVE, WITH MENTAL RETARDATION; CANPMR
214800 - CHARGE SYNDROME
612863 - CHROMOSOME 6q24-q25 DELETION SYNDROME
123560 - CRYPTOMICROTIA-BRACHYDACTYLY SYNDROME
112430 - LONG-THUMB BRACHYDACTYLY SYNDROME
186500 - MULTIPLE SYNOSTOSES SYNDROME 1; SYNS1
609307 - SPINOCEREBELLAR ATAXIA 27; SCA27
613792 - CHROMOSOME 3pter-p25 DELETION SYNDROME
108800 - ATRIAL SEPTAL DEFECT 1; ASD1
108650 - SPASTIC ATAXIA 7, AUTOSOMAL DOMINANT; SPAX7
139210 - MYHRE SYNDROME; MYHRS
605259 - SPINOCEREBELLAR ATAXIA 13; SCA13
113301 - BRACHYDACTYLY, TYPE E, WITH ATRIAL SEPTAL DEFECT, TYPE II
164210 - HEMIFACIAL MICROSOMIA; HFM
607136 - SPINOCEREBELLAR ATAXIA 17; SCA17
607346 - SPINOCEREBELLAR ATAXIA 19; SCA19
188400 - DIGEORGE SYNDROME; DGS
206900 - MICROPHTHALMIA, SYNDROMIC 3; MCOPS3
156200 - MENTAL RETARDATION, AUTOSOMAL DOMINANT 1; MRD1

Analysis design and samples included

Variants to include

Comparing to other family members

Adding biological information

OMIM search result

Analysis Result File

Prioritization of the variant in the gene known to cause one of the diagnosis suggested

	BR	BS	BT	BU	BV	BW	BX	BY
1	Approved Symbol	OMIM	OMIM Matching Phenotypes	OMIM Phenotypes	Mouse Model	Mammalian Phenotype	Expression	Interaction
2	EXT2	http://omim.or	Exostoses, multiple, type 2	Exostoses, multiple, type 2	Homozygous nu	http://www.informatics.ja	Adipose tissue	AGRN, ANXA
3	OR5J2					http://www.informatics.jax.org/batch/su		ADRBK2, ARF
4	OR4D9					http://www.informatics.jax.org/batch/su		ADRBK2, ARF
5	LRP5	http://omim.org/entry/603506		Exudative vitreoretinopathy 4	Homozygous m	http://www.informatics.ja	Brain / Colon /	APCDD1, AXI
6	MRPL21					http://www.informatics.ja	Adrenal / Bone	AARS2, AASS
7	CTTN				Mice homozygo	http://www.informatics.ja	Adrenal / Bone	ABCF1, ABCF
8	ARHGEF17					http://www.informatics.ja	Adipose tissue	A2M, ARAP2,
9	XRRA1					http://www.informatics.ja	Brain / Breast / Kidney / Nerv	
10	RSF1					http://www.informatics.jax.org/batch/su		CASC5, CENF

PhenoDB analysis tool: Analyses saved

Analysis Results :

Run Name	Analysis Type	Refgene Gene Location	Exclude SNP	Percentage Cutoff	Exclude Chrosomome X	Indel Span	Initial Count	Final Count	Analysis Results	Date Created	Log
AR-HOMO	Autosomal recessive - Homozygous	-	dbSNP131, dbSNP129, dbSNP126	0.01	-	50	65,485	22	tab, xls	Jul 07, 2013 07:10:13	Show
AR-CH	Autosomal recessive - Compound heterozygous	-	dbSNP131, dbSNP129, dbSNP126	0.01	-	50	65,485	10	tab, xls	Jul 07, 2013 07:12:16	Show
AD	Autosomal dominant - Variants	-	dbSNP131, dbSNP129, dbSNP126	0.01	-	50	65,485	254	tab, xls	Jul 07, 2013 07:13:05	Show

Run name :

Overlap : Select optional overlap...

Merge Selected Analysis Results | Create Final Results From Selected Analysis Results | Delete Selected Analysis Results

Download analysis results table as tab-delimited text, Microsoft Excel file.

Analysis design page – Filter function

Filter :

Member ID	Family Member	Affected	Sample ID - Genomic Version - Lab LIMS ID - ANNOVAR File Name
BH3200_1	Patient - Female	Yes	○ BH3200_1_1 - 1 - 200855166 - 200855166@1097030495_MS_OnBait_ANNOVAR_REPORT.txt
BH3200_2	Mother	No	-
BH3200_3	Father	No	-
BH3200_4	Brother	No	-
BH3200_5	maternal male cousin - Male	Yes	○ BH3200_5_1 - 1 - 200855167 - 200855167@1097030496_MS_OnBait_ANNOVAR_REPORT.txt
BH3200_6	Maternal aunt	No	-
BH3200_7	Father of _5 - Male	No	-
BH3200_8	Maternal grand-mother	No	-
BH3200_9	Maternal grand-father	No	-

	Run Name	Analysis Type	Final Count	Analysis Results	Log
○	AR	Autosomal recessive - Homozygous	2	tab, xls	Show
○	XL	X-Linked recessive	-	tab, xls	Show
○	CH	Autosomal recessive - Compound heterozygous	2	tab, xls	Show
○	AD	Autosomal dominant - Inherited mutation	43	tab, xls	Show

Run name :

Filter :

- ✓ Select filter...
- Gene names
- ACMG incidental findings
- OMIM phenotypic series
- Interactions

PhenoDB analysis tool: Searching all complete analyses

Submissions by State | Inbox/Tagged | Submissions by Group | Search | Analyses | Analyses Search

- Submission ID : [ ] Edit View Analyze Filter
- Create new analysis sandbox
- List all analysis sandboxes
- List analyses :
  - List all analyses
  - List 'Autosomal recessive – Compound heterozygous' analyses
  - List 'Autosomal recessive – Homozygous' analyses
  - List 'X-Linked recessive' analyses
  - List 'Autosomal dominant – New mutation' analyses
  - List 'Autosomal dominant – Inherited mutation' analyses
  - List 'Autosomal dominant – Variants analyses'
  - List 'Merge' analyses
- Upload genomic methods spreadsheet: Choose File no file selected Submit
  Empty cells in the spreadsheet will clear corresponding fields in the database : [ ]
  Download sample genomic methods spreadsheet for uploading.

PhenoDB analysis tool: Cohort analysis tool

☐	CMG2265 View / Edit / Submission / Samples	bone_cone_2	Autosomal Recessive – Compound Heterozygous	–	–	0.01	–	50	65,485	14	tab, xls	Dec 20, 2012 09:36:24	Show
☐	CMG2265 View / Edit / Submission / Samples	bone_cone_2	Autosomal Recessive – Homozygous	–	–	0.01	–	50	65,485	32	tab, xls	Dec 20, 2012 09:37:21	Show
☐	CMG2265 View / Edit / Submission / Samples	cone_bone_2	Autosomal Dominant – Variants	–	–	0.01	–	50	65,485	312	tab, xls	Dec 20, 2012 09:37:59	Show
☐	CMG2283 View / Edit / Submission / Samples	bone_cone_3	Autosomal Recessive – Compound Heterozygous	–	–	0.01	–	50	73,809	21	tab, xls	Dec 20, 2012 09:38:50	Show
☐	CMG2283 View / Edit / Submission / Samples	bone_cone_3	Autosomal Recessive – Homozygous	–	–	0.01	–	50	73,809	6	tab, xls	Dec 20, 2012 09:39:24	Show
☐	CMG2283 View / Edit / Submission / Samples	bone_cone_3	Autosomal Dominant – Variants	–	–	0.01	–	50	73,809	307	tab, xls	Dec 20, 2012 09:39:57	Show
☐	CMG2318 View / Edit / Submission / Samples	proband_parents_brother	Autosomal Dominant – New Mutation	–	–	0.01	–	50	76,778	14	tab, xls	Dec 20, 2012 10:06:15	Show
☐	CMG2318 View / Edit / Submission / Samples	–	Autosomal Recessive – Homozygous	–	–	0.01	–	50	76,778	16	tab, xls	Dec 20, 2012 10:17:06	Show
☐	CMG2318 View / Edit / Submission / Samples	1013	Autosomal Recessive – Homozygous	–	–	0.01	–	50	63,486	29	tab, xls	Dec 21, 2012 12:12:44	Show
☐	CMG2318 View / Edit / Submission / Samples	test	Autosomal Recessive – Compound Heterozygous	–	–	–	Yes	–	76,778	11	tab, xls	Jan 02, 2013 14:10:42	Show
☐	CMG2342 View / Edit / Submission / Samples	3_brothers	Autosomal Recessive – Compound Heterozygous	–	–	0.01	–	50	76,327	4	tab, xls	Dec 20, 2012 10:41:11	Show
☐	CMG2342 View / Edit / Submission / Samples	3_brothers	Autosomal Recessive – Homozygous	–	–	0.01	–	50	76,327	6	tab, xls	Dec 20, 2012 10:45:29	Show
☐	CMG2342 View / Edit / Submission / Samples	3_brothers	Autosomal Dominant – Inherited Mutation	–	–	0.01	–	50	76,327	67	tab, xls	Dec 20, 2012 10:48:25	Show
☐	CMG2342 View / Edit / Submission / Samples	test	Autosomal Recessive – Compound Heterozygous	–	–	–	Yes	–	76,327	25	tab, xls	Dec 29, 2012 09:35:31	Show
☐	CMG2342 View / Edit / Submission / Samples	1001	Autosomal Dominant – Inherited Mutation	–	–	0.01	Yes	50	63,000	43	tab, xls	Jan 15, 2013 16:58:20	Show
☐	CMG2363 View / Edit / Submission / Samples	2_SISTERS	Autosomal Recessive – Homozygous	–	–	0.01	Yes	50	75,667	11	tab, xls	Dec 20, 2012 10:52:14	Show
☐	CMG5058 View / Edit / Analyze / Submission / Samples	5individuals	Autosomal Dominant – Inherited Mutation	–	dbSNP131, dbSNP129, dbSNP126	0.01	–	50	77,561	9	tab, xls	Mar 12, 2013 13:39:33	Show
☐	CMG5058 View / Edit / Analyze / Submission / Samples	5'UTR	Autosomal Dominant – Inherited Mutation	'UTR3;UTR5', 'UTR5', 'UTR5;UTR3'	dbSNP131, dbSNP129, dbSNP126	0.01	–	50	77,561	–	tab, xls	Mar 12, 2013 15:54:22	Show
☐	CMG5058 View / Edit / Analyze / Submission / Samples	3'UTR	Autosomal Dominant – Inherited Mutation	'UTR3', 'UTR3;UTR5', 'UTR5;UTR3'	dbSNP131, dbSNP129, dbSNP126	0.01	–	50	77,561	5	tab, xls	Mar 12, 2013 15:57:43	Show
☐	CMG5058 View / Edit / Analyze / Submission / Samples	SYNONYMOUS	Autosomal Dominant – Inherited Mutation	'synonymous'	dbSNP131, dbSNP129, dbSNP126	0.01	–	50	77,561	–	tab, xls	Mar 12, 2013 16:06:03	Show
☐	CMG5058 View / Edit / Analyze / Submission / Samples	5individuals_10bp	Autosomal Dominant – Inherited Mutation	–	dbSNP131, dbSNP129, dbSNP126	0.01	–	10	77,561	10	tab, xls	Mar 14, 2013 06:52:25	Show

Run Name :

Overlap : Select optional overlap...

Merge Selected Analysis Results

Delete Selected Analysis Results

Download analysis results table as tab-delimited text, Microsoft Excel file.

Research PhenoDB

- Designed for BHCMG, but freely available for download and modification
- Includes ELSI module and Samples Module
- Collects images and reports
- Space for documenting considerations, analyses, conclusions

BHCMG PhenoDB

- Holds data on 3280 submissions
  - ✓ Including cohorts ranging from 5-300
- Comprising over 5,000 individuals
- Size of dataset allows robust comparisons

http://phenodb.org

Welcome to myPhenoDB

(PDB)

Please Sign In

Current User :

Email :

Password :

Login

Forgotten your password?

New to the database?

- Create an account...

McKUSICK-NATHANS Institute of Genetic Medicine

PhenoDB.org

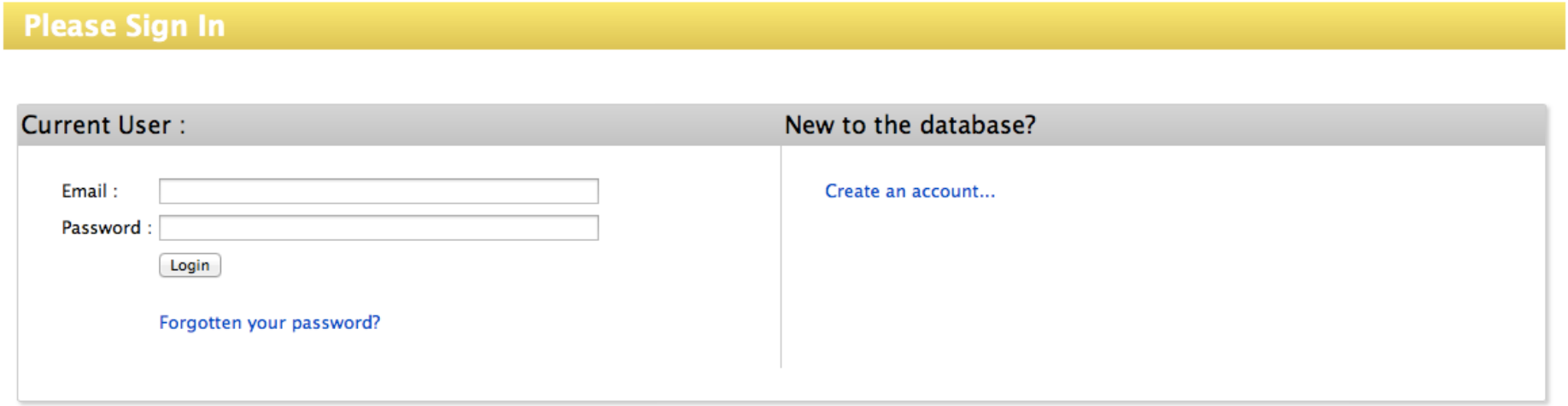
- Can store phenotypic features and final results
  - ✓ VCF is immediately converted to Annovar file
  - ✓ Annovar file is available for 24 hours
  - ✓ No images, etc can be uploaded to this system
  - ✓ Useful for those in resource limited environment
    - ✓ 133 accounts created
- Use this to try out the system
  - ✓ especially the analysis tool

PhenoDB.org

- Download the tool for use in a clinic or lab
- Can be toggled to fully identified or deidentified
- Both tools have been downloaded >200 times

Welcome to GeneMatcher

(GM)



Centers for Mendelian Genomics

GeneMatcher (GM) was created by Nara Sobreira, François Schiettecatte, Ada Hamosh and members of the BHCMG of the Centers for Mendelian Genomics network to promote productive interactions among investigators around the world who are trying to identify and understand the genes and variants responsible for Mendelian disease.

http://genematcher.org

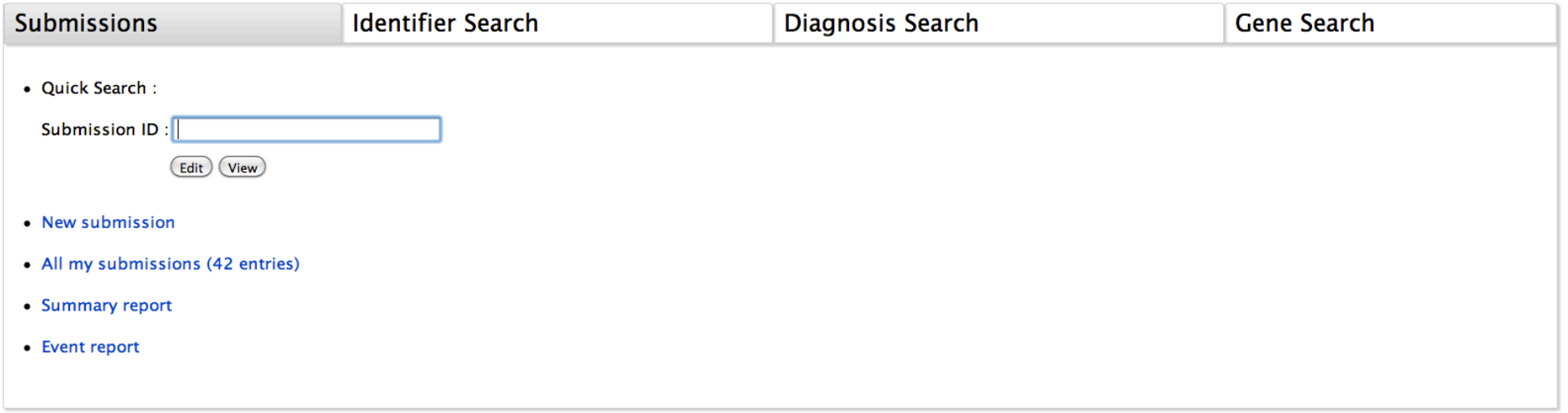
Reasons for Development of GeneMatcher

- Unsolved exomes from probands/families with a rare/unique phenotype
- Small number of candidate genes/variants from research or clinical exomes, need more probands/families with mutation in the same gene to prove causality
- Little is known about the biology of the candidates gene(s)
- Unpublished animal models with mutations in the same candidate genes and their phenotype

Features of GeneMatcher

- Quick input of data
- De-identified
- No consent requirements
- Not searchable
- Only one gene name is required (variant level is possible if desired)
- Can find not only another patient, but potentially model organism

GeneMatcher Home Page



GeneMatcher Submission Page

View

Submission Identification :

Submission ID: Required...

You need to provide an identifier that uniquely identifies this entry as it will be included in emails that are sent out when a match is made.
You should not use any information that could be used to identify the patient or their family.

Organism Model :

Organism Model: Human (Homo sapiens)

Select the organism model.
You should only enter Human genes in the Results section.

Genetic Disorder *(optional)* :

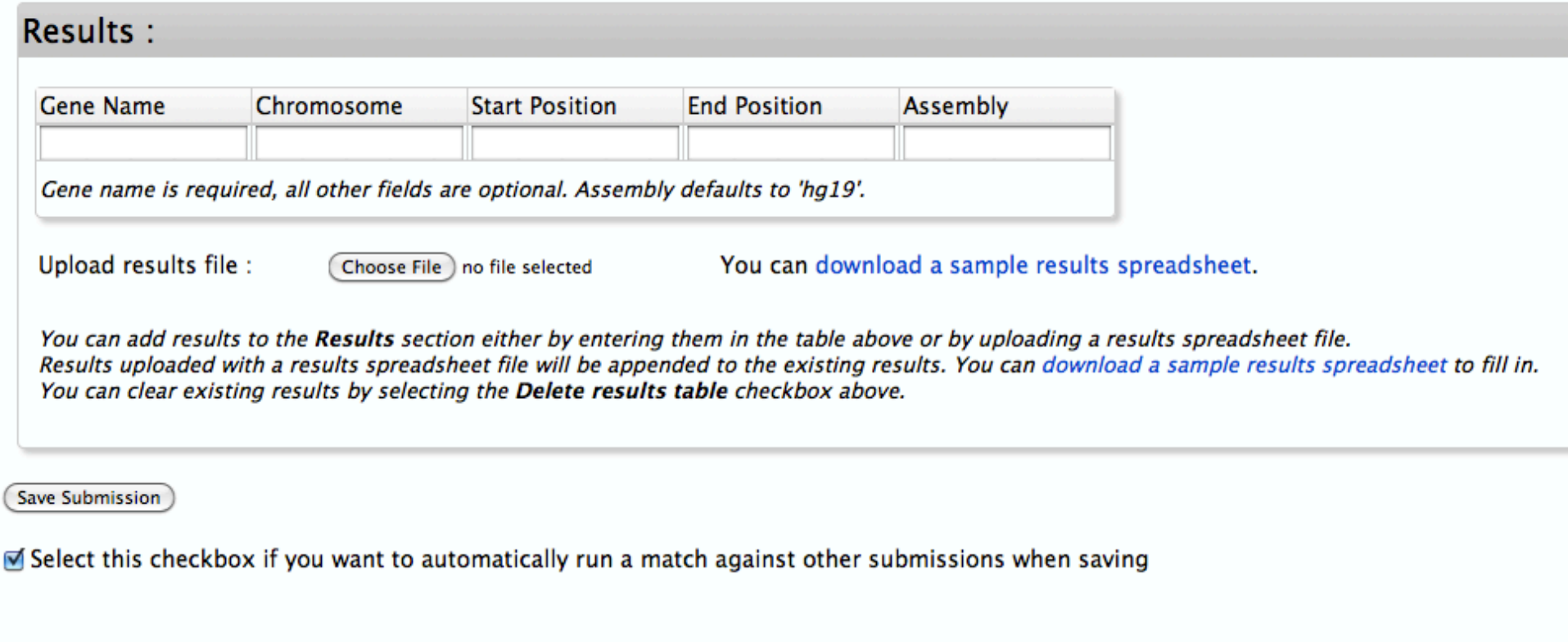
Diagnosis Category:

- ☐ A Mendelian disorder described in OMIM for which the responsible gene has not been identified (example: 223370, Dubowitz syndrome)
- ☐ 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)
- ☐ An unknown disorder (not described in OMIM) but with segregation in your family consistent with Mendelian Inheritance

Inheritance pattern : Pick best fit...

*The **Genetic Disorder** section is optional, and all the fields in this section are optional.*

It only requires the name of at least one gene



Matching notification email

Submitted for matching:

Nara Sobreira - narasobreira@ig.com.br - Johns Hopkins University - test

Matching submissions :

Ada Hamosh - ahamosh@jhmi.edu - - - BH1010 - SPATA5
David Valle - dvalle@jhmi.edu - Johns Hopkins - test - SPATA5

Please do not reply to this email, it was sent from an unattended email address.

GeneMatcher (GM).
https://genematcher.org/

- Seizures
- Microcephaly
- Profound bilateral sensorineural hearing loss

Welcome Back Nara Sobreira

Submissions | Identifier Search | Diagnosis Search | Gene Search

Matcher ID :

Submission ID :

Search

Welcome Back Nara Sobreira

Submissions | Identifier Search | Diagnosis Search | Gene Search

MIM number :

Disorder type : Select optional disorder type...

Presumed inheritance pattern : Select optional inheritance pattern...

Features :

Submissions must match all features

Search

Welcome Back Nara Sobreira

Submissions | Identifier Search | Diagnosis Search | Gene Search

Gene names :

Submissions must match all gene names

Genomic locations :

Submissions must match all genomic locations

Assembly: (defaults to hg19)

Gene name :

Genomic location :

Assembly :

Optional match on genomic location

Search

GeneMatcher Statistics

- As of April 1st 2015:
- 1923 genes submitted by
- 400 submitters from
- 38 countries
- 201 matches !!!

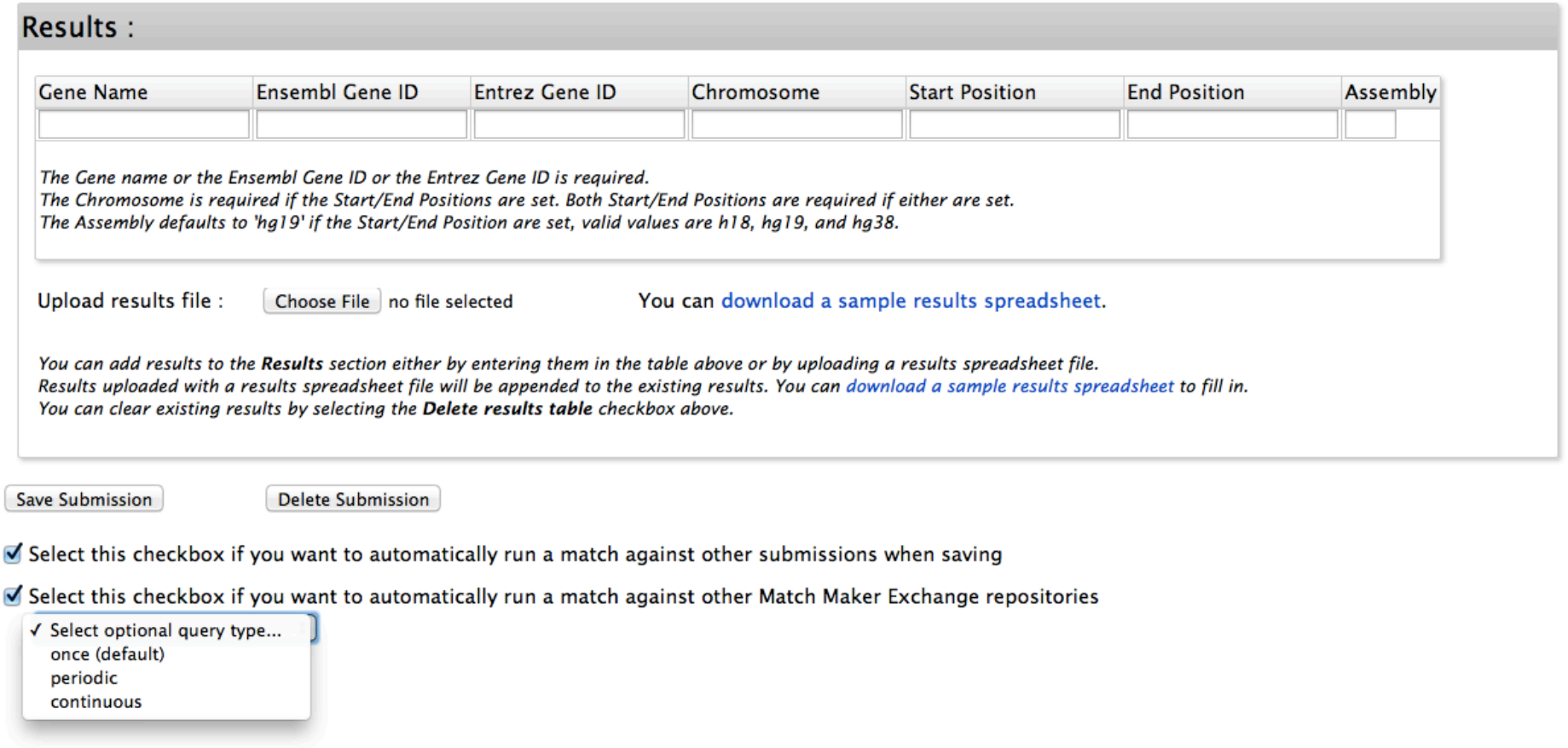
Matchmaker Exchange

Genomic discovery through the exchange of phenotypic & genotypic profiles

Matchmaker Exchange

Participants & Supporters

Query other matchmaking databases



Coming Soon

- Phenotypic features matching (for presumed novel disorders)

Thanks for your attention!

Acknowledgements

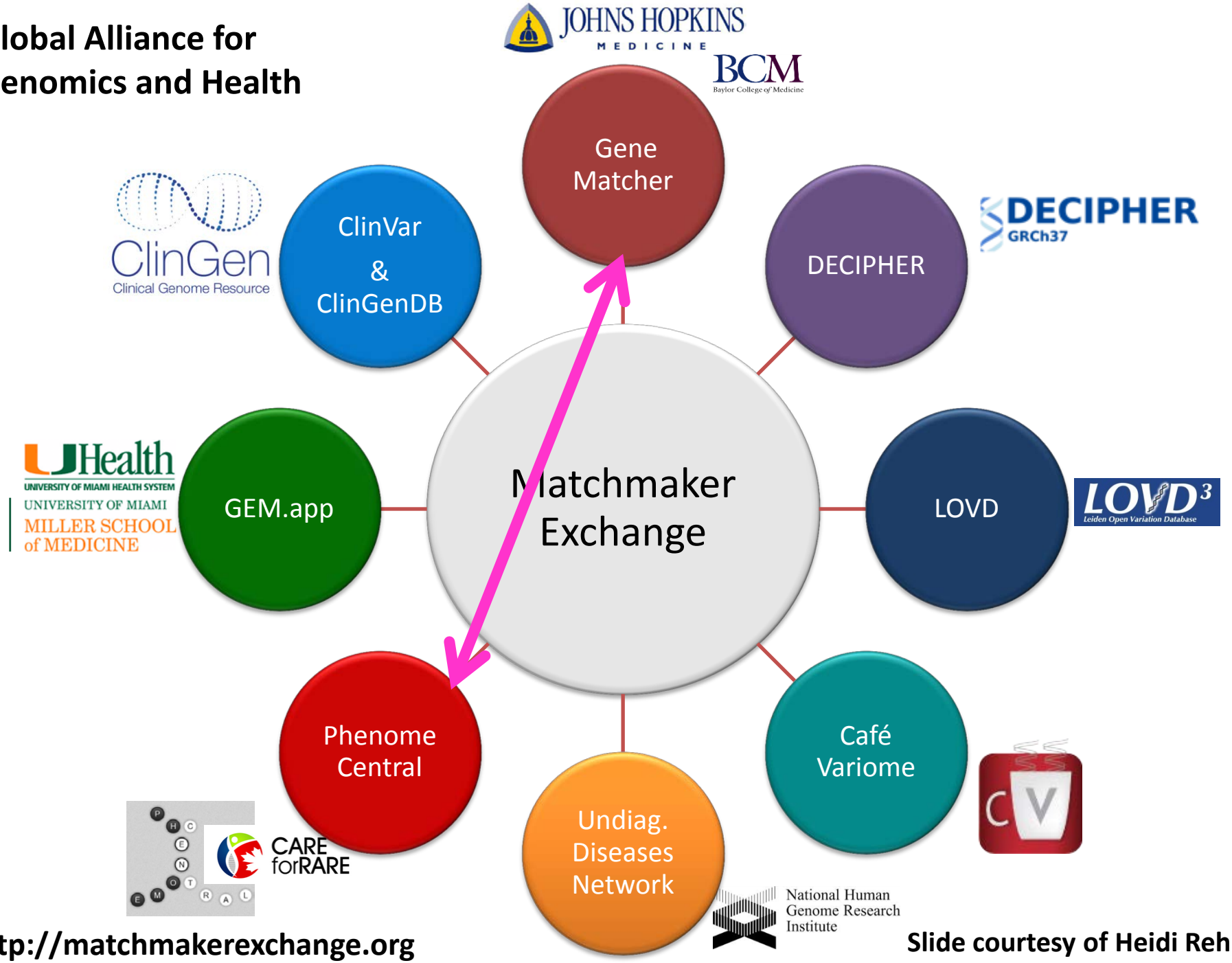
- François Schiettecatte, Corinne Boehm, Julie Hoover-Fong, Reid Sutton, Jim Lupski, David Valle & others for PhenoDB
- The CMGs and especially the Baylor-Hopkins CMG team

Thank you!

ahamosh@jhmi.edu

nsobrei@jhmi.edu

Global Alliance for Genomics and Health



http://matchmakerexchange.org

Slide courtesy of Heidi Rehm