National Human Genome Research Institute

Exome Sequencing to Identify Variants Associated with Moebius Syndrome

Rachel L. Goldfeder[1], Peter S. Chines[1], Lori Bonnycastle[1], Amy Swift[1], Narisu Narisu[1], Irini Manoli[2], and Francis S. Collins[1]

[1]Molecular Genetics Section, Genome Technology Branch, NHGRI, Bethesda, MD 20892
[2]Organic Acid Research Section, Genetics and Molecular Biology Branch, NHGRI, Bethesda, MD 20892

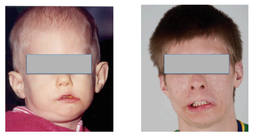
Molecular Genetics Section

Abstract

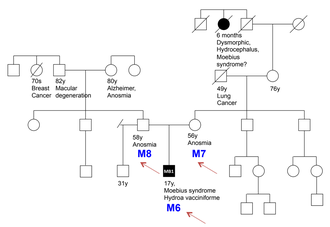
Moebius Syndrome is a rare, congenital neurological condition. The primary symptom is a lack of facial expressions and lateral eye movements due to paralysis of facial muscles. Most cases are sporadic; the disease is believed to result from genetic and prenatal environmental factors. Our group is using whole-exome DNA sequencing to identify genetic variations that are associated with the disease. We sequenced the DNA of seven affected individuals and their parents (and siblings, where available).

Preliminary data from exome sequencing has identified over 290,000 variants in the affected individuals. We have implemented filters based on data quality and predicted function to filter and prioritize variants. We have found several interesting variants that we plan to validate using mass spectrometry and eventually follow up with functional studies.

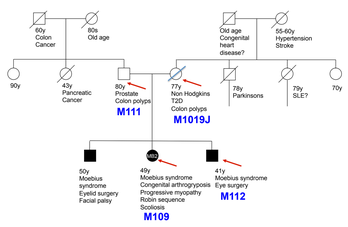
Introduction



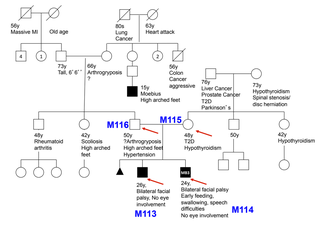
Family 1



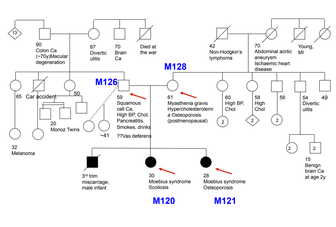
Family 2



Family 3

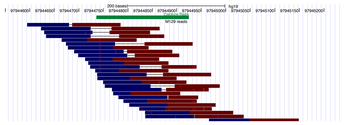


Family 4

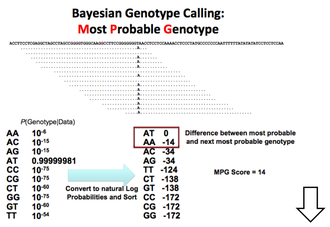


Processing

Sequence Generation
- Whole Exome Sequencing



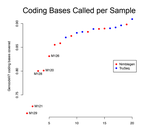
Align Sequence Reads
- Novoalign
- hg19

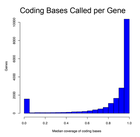


Call Genotypes
- MPG

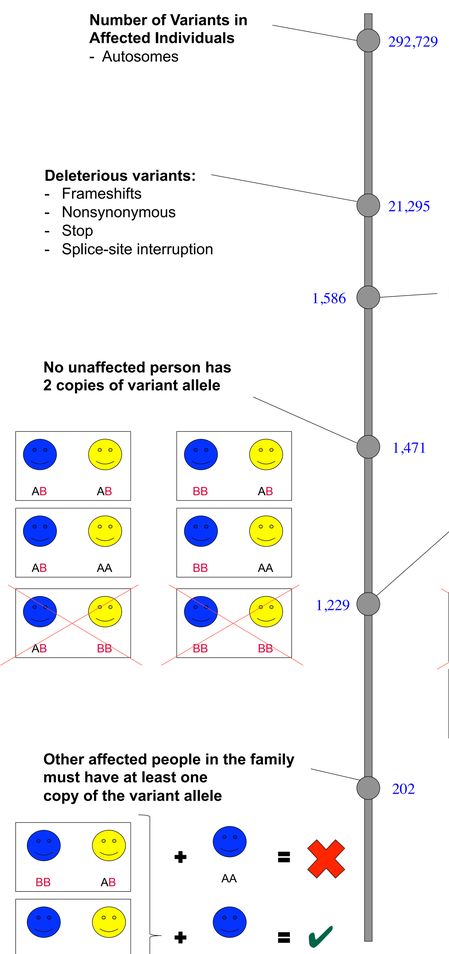
Quality Control Measurements

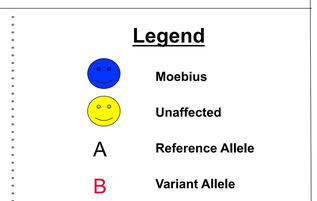
Family	ID	Pedigree	Status	Exome Capture	Number of Reads	Percent Duplicates	Coding Bases with Confident Genotype Calls
MB1	**M6**	**Son**	**Affected**	**Nimblegen**	**54,483,705**	**6.81**	**31,570,288**
MB1	M7	Mother	Unaffected	Nimblegen	56,457,736	10.92	31,043,370
MB1	M8	Father	Unaffected	Nimblegen	56,305,891	6.42	31,350,519
MB2	**M109**	**Daughter**	**Affected**	**TruSeq**	**89,882,299**	**8.45**	**31,273,435**
MB2	M111	Father	Unaffected	TruSeq	189,991,208	28.3	32,301,599
MB2	**M112**	**Son**	**Affected**	**TruSeq**	**105,298,661**	**31.22**	**31,349,399**
MB2	M1019J	Mother	Unaffected	TruSeq	99,493,405	39.44	30,931,287
MB3	**M113**	**Son**	**Affected**	**TruSeq**	**115,305,493**	**27.66**	**31,613,741**
MB3	**M114**	**Son**	**Affected**	**TruSeq**	**145,212,828**	**32.47**	**31,832,480**
MB3	M115	Mother	Unaffected	TruSeq	117,996,745	31.49	31,563,831
MB3	M116	Father	Unaffected	TruSeq	124,040,031	31.75	31,678,005
MB4	**M120**	**Daughter**	**Affected**	**Nimblegen**	**26,719,631**	**7.09**	**28,459,431**
MB4	**M121**	**Daughter**	**Affected**	**Nimblegen**	**20,228,821**	**7.31**	**25,788,919**
MB4	M126	Father	Unaffected	Nimblegen	30,418,194	7.36	29,532,103
MB4	M128	Mother	Unaffected	Nimblegen	26,850,846	6.99	28,426,448





Filtering





Results

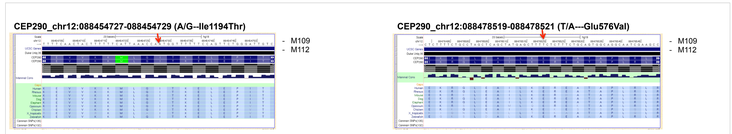
Single Variants

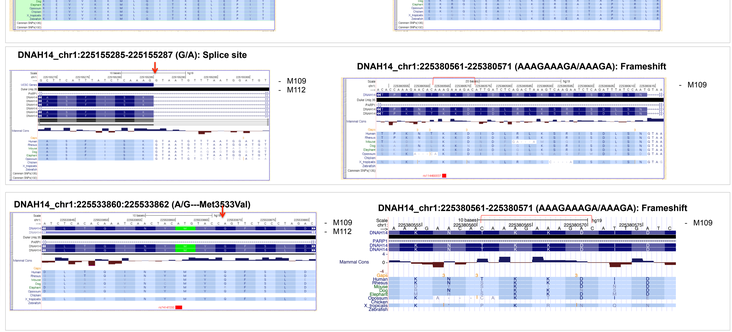
Affected Sample(s)	Model	Gene ID	Ref allele	Var allele	AA Change	Type	Freq	Gene Decription
M6	Recessive	AKAP3	G	A	S700F	NonSyn	0	A-kinase anchor protein 3
			A	G	S700P	NonSyn	0	
	Dominant (NMI)	DNAH17	G	A	R711W	NonSyn	0	Dynein heavy chain 17, axonemal
	Dominant (NMI)	ZFP64	AAAAAAAAAA	AAAAAAAAAAA	---	Frameshift	0	Zinc finger protein 64
M109, M112	Recessive	PRICKLE4	CT	CTTCT	---	Frameshift	0	Over-expressed breast tumor protein
M113, M114	Recessive	TAC C1	TGACTG	TG	---	Frameshift	0	Transforming, acidic coiled-coil containing
M113	Recessive	HNF1A	TCATTCATTCATTT CAT	TCATTCATTCATTC ATTCATTCAT	---	Frameshift	0	Transcription factor 1, hepatic
M121, M120	Recessive	KRTAP4-6	G	A	S202F	NonSyn	0.00366	Keratin Associated Protein 4-6

Compound heterozygotes

Affected Sample(s)	Gene	Freq.A	Freq.B	Type.A	Type.B	AA Change.A	AA Change.B	Gene Description
M6	SLC7A2	0.02542	0.00014	NonSyn	NonSyn	V545M	R517W	Solute carrier family 7, member 2
M109, M112	CEP290 (CPR8)	0.00634	0	NonSyn	NonSyn	I1194T	E576V	Centrosomal protein 290kDa;Homo sapiens monoclonal antibody 3H11 antigen mRNA, complete cds.
M109	DNAH14	0.00229	0	Splicing	Frameshift	---	---	Dynein, axonemal, heavy polypeptide 14 isoform;Dynein heavy chain 14, axonemal (Axonemal beta dynein heavy chain 14) (Ciliary dynein heavy chain 14).
		0.01315	0	NonSyn	Frameshift	M3533V	---	

Example: Compound heterozygotes in M109 and M112





Future Directions

- Call genotypes with Genome Analysis ToolKit
  - Validate discrepancies between GATK and MPG
- Use MODY samples as additional controls
- Call and filter non-autosomes (chrX, chrY, chrM)
- Better annotation:
  - Nearby/overlapping variants
  - Predict effect of nonsynonymous variants
- Validation of variants
- Functional follow-up

Acknowledgements

- **Collins Lab**
- **Larry Brody**
- **Mullikin Lab**
  - Jim Mullikin
  - Jamie Teer
  - Nancy Hansen
- **IT Support**
  - Jesse Becker
- **Bill Gahl Lab**
- **NIH Intramural Sequencing Center**
  - Alice Young
- **Mount Sinai School of Medicine**
  - Ethylin Jabs
  - Bryn Webb
- **Moebius Syndrome Foundation**