

Next-generation human genomics

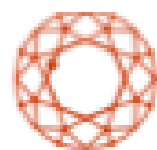
Tim Yu, MD, PhD

Division of Genetics & Genomics, Boston Children's Hospital

Claritas Genomics (\$, stock)

Dept of Neurology, MGH

Harvard Medical School & the Broad Institute



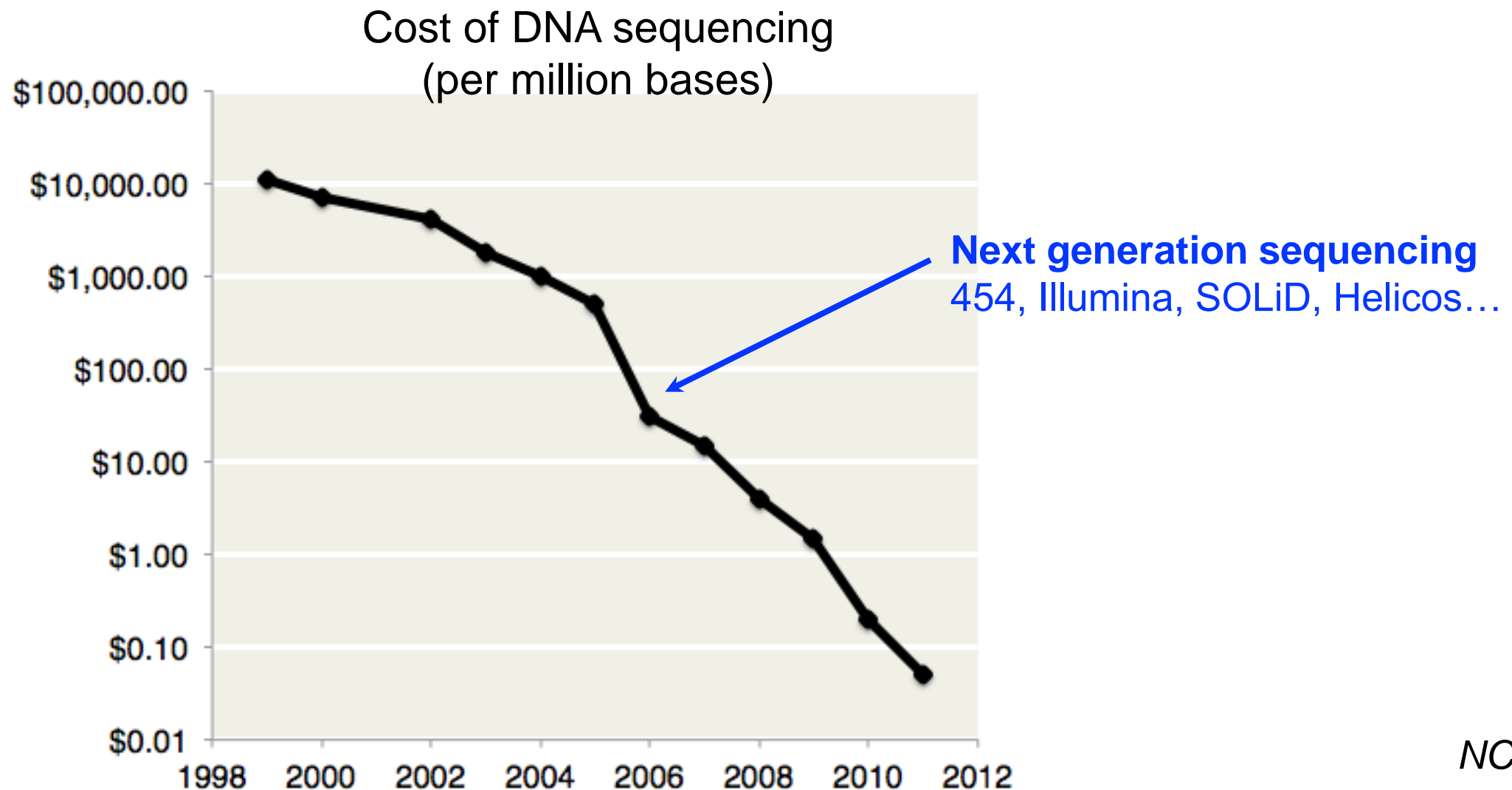
CLARITAS
GENOMICS



The personalized genome

- Amy, age 21 years, visits with her physician and elects to have **complete genome sequencing**.
- At a follow-up visit, Amy chooses to learn of her genetic risk factors for heart disease, diabetes, breast cancer, and colon cancer. Amy's physician provides her with risk scores for those disorders, and with suggestions for lifestyle modifications. Specifically, Amy is alerted to her particularly **high risk of developing type 2 diabetes, and her physician recommends a rigorous program of diet and exercise** that had been shown in a controlled study to delay or prevent disease onset.
- The next year, Amy develops mild asthma and **her physician selects an optimal therapy based on Amy's genetic profile**.
- Five years later, Amy informs her physician that she and her husband are planning to start a family, and they request information regarding the risk of having a child affected by a serious genetic disease, based on their genome sequence data. She learns that both she and her husband are **carriers for the recessive lethal childhood disorder spinal muscular atrophy, and they seek further counseling**.
- When Amy turns 40, she begins **colorectal cancer screening** based on her higher-than-average risk factors, and at age 45 a precancerous polyp is detected in her colon and is successfully removed.

Sequencing costs



| | |
|---|---------------|
| Cost of the first human genome (1990-2003): | \$2.7 Billion |
|---|---------------|

| | |
|--------------------------------|----------|
| Cost of a human genome in 2009 | \$20,000 |
|--------------------------------|----------|

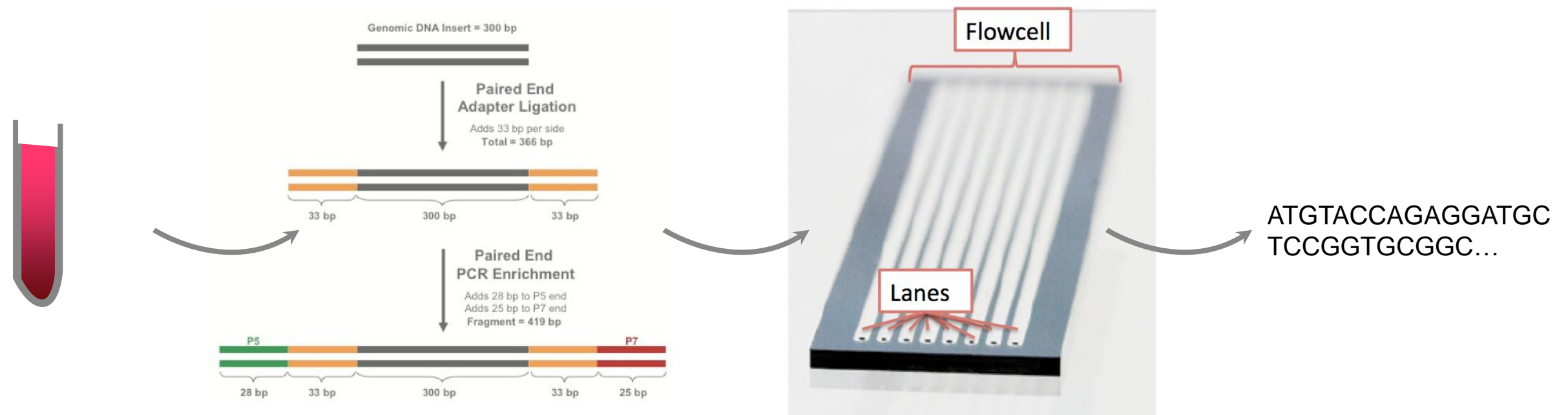
| | |
|--------------------------------|---------|
| Cost of a human genome in 2013 | \$2,500 |
|--------------------------------|---------|

| | |
|---------------------------------|---------------------|
| Cost of a human “exome” in 2013 | \$800 (< brain MRI) |
|---------------------------------|---------------------|

Advances in the last 5 years have made it easy to generate whole genome sequencing data.

The challenge is interpretation.

Next-generation sequencing



Blood samples

DNA Libraries

Flowcells

Sequence

Millions of reads are mapped en masse to a reference genome

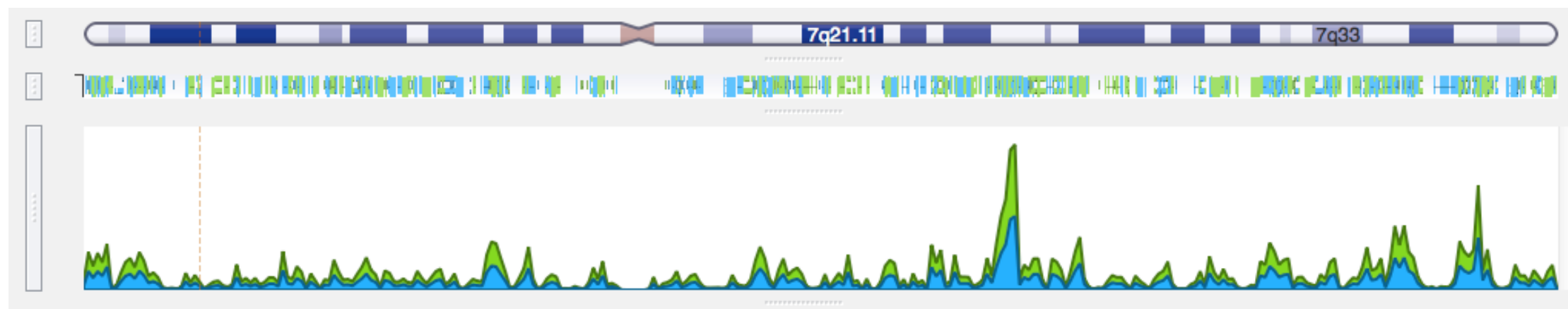
```
40421551 40421561 40421571 40421581 40421591 40421601 40421611 40421621 40421631 40421641 40421651 40421661 40421671 40421681 40421691 40421701 40421711 40421
721tttgagcagacctatataagatgggttatgaagattcacacagcggctcatgctgtgatcccagcactttgggaggctgaggcaagtggagcacctgagatcatgagttcaagaccagcctggccaacatggtgaaaccccatctctactaaagatacaaaaattatccaggtgtggtg
.....A.....T.....
A tgaacagacctatataagatgggt tgaagattcacacagtggtctcatgctgtgatcccagcac tgggaggctgagtcagtgaggcacctgagatcatgagtt ACCAGCCTGGCCAACATGGTGAACCCCATCTCTACTAAA ATACAAAAATTATCCAGGTGTGGTG
aca cagacctatataagatgggt aagatacacacagtggtctcatgctgtgatcccagcactt GGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTC cagcctggccaacatggtgaaaccccatctctactaaaga ACAAAAAATTATCCAGGTGTGGTG
acatt GACCTATATAAGATGGTTATGAAGATTACACAGTGGCTC CCTGTGATCCCAGCAGCTTTGGGAGGCTGAGGCAAGTGGAG ACCTGAGATCATGAGTTCAAGACCAGCCTGGACAACATGG AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT G
ACATTGAACAG ATATAAGATGGTTATGAAGATTACACAGTGGCTCATGCC tgatcccagcactttgggagg TGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGACCA GCCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA ATCCAGGTGTGGTG
ACATTGAACAGAC TCAGATGGTTATGAAGATTACACAGTGGCTCATGCCGTGT ATCCCAGCAGCTTTGGGAGGCTGAGGCAAGGGGAGCACCTG ATGAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCA CTCTACTAAAGATACAAAAATTA aggtgtggtg
acatttgaacagacctata aagatgggttatgaagattcacacagtggtctcatgctgtg TCCCAGCAGCTTTGGGAGGCTGAGGCAAGTGGAGCACCTGA ATGAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCA TACTACTAAAGATACAAAAATTATCCAGGTGTGGTG
acatttgaacagacctatata gatgggttatgaagattcacacagtagctcatgctgtgat AGCACTTTGGGAGGCTGAGGCAAGGGGAGCACCTGA GAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCATC CTACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACAT gacctatataagatgggttatgaagattcacacagtggtctc CTGTAATCCCACACTTTGGGAGGCTGAGGCAAGTGGAGC CCTGAGATCATGAGTTCAAGA AGCCTGGCCAACATCGTGAACCCCATATCTACTAAAGAT caaaattatccaggtgtggtg
ACATTGAACAGACCTATATAA TGGTTATGAAGATTACACAGTGGCTCATGCCGTGTATCC cactttgggatgctgaggcaagtggagcacctgagatcat CAAGACCAGCCTGGCCAACATGGTGAACCCCATCTCTAC AGAAATACAAAAATTATCCATGTGTGGTG
ACATTG ACCTATATAAGATGGTTATGAAGATTACACAGTGGCTCA TGTGATCCCAGCAGCTTTGGGAGGCTGAGGCAAGTGGAGCA CTGAGATCACGAGTTCAAGACCAGCCTGCCAACATGGTC AACCCCATCTCTACTAAAGATACAAAAATTACCCAGGTGT GG
ACATTGAACAGACCTATATAAGA GGTTACGAAGATTACACAGTGGCTCATGCCGTGTATCCC cacattgggaggctgaggcaagtggagcacctgagatcat AAGACCAGCCTGGCCAACATGGTGAACCCCATCTCTACT AAGATACAAAAATTATCCAGGTGTGGTG
acatttgaacagacctatataagat ttatgaagattcacacagtggtctcatgctgtgatcccag CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA agcctggccaacatggtgaaaccccatctctactaaagat AAAATTATCCAGGTGTGGTG
acatttgaacagacctatataagatgggt aagattcacacagtggtctcatgctgtgatcccagcactt GGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTC GCCTGGCCAACATGGTGAAC CCCATCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACATTGAACAGACCTATATAAGATGGTTA agattcacacagaggctcatgctgtgatcccagcacttt AGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAG CCTGGCCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA TTATCCAGGTGTGGTG
acatttgaacagacctatataagatggtta TTCACAGTGGCTCATGCCGTGTATCCCAGCAGCTTTGGG GCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGAC CCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA atccaggtgtggtg
ACATTGAACAGACCTATATAAGATGGTTAT CAGTGGCTCATGCCGTGTAT ACTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG CAACATGGTGAACCCCATCTCTACTAAAGATACAAAAAT TCCAGGTGTGGTG
ACATTGAACAGACCTATATAAGATGGTTATGAAG CAGTGGCTCATGCCGTGTATC ACTTCGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG AACATGGTGAACCCCATCTCTACTAAAGATACAGAAAT aggtgtggtg
ACATTGAACAGACCTATATAAGATGGTTATGAAGAT CAGTGGCTCATGCCGTGTATCC CCTCTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG ACATGGTGAACCCCATCTCTACTAAAGATACAAAAATTA GTGTGGTG
ACATTGAACAGACCTATCTAAGATGGTTATGAAGATT GCGGCTCATGCCGTGTATC CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA ACATGGTGAACCCCATCTATACTAAAGATACAAAAATTA TGTGGTG
ACATTGAACAGACCTATATAAGATGGTTATGAAGATTG CTCTTGCTGTGATCCCAGCAGCTTTGGGAGGCTGAGGCAA TGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCA TGGTGAACCCCATCTCTACTAAAGATACAAAAATTATCC ggtg
ACATTGAACAGACCTATATAAGATGGTTATGAAGATTG CTCATGCCGTGTGATCCCAGCAGCTTTGGGAGGCTGAGGCAA TGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCA GGTGAACCCCATCGCTACTAAAGATACAAAAATTATCCA GG
ACATTGAACAGACCTATATAAGATGGTTATGAAGATTCA GTGATCCCAGCAGCTTTGGGAGGCTGAGGCAAGTGGAGCAC GATCATGAGTTCAAGACCCGCTGGCCAACATGGTGAAC ccctctctactaaagatacaaaaattatccaggtgtggtg
AGATGGTTATGAAGATTACACAGTGGCTCATGCCGTGA CCAGCAGCTTTGGGAGGCTGAGGCAAGTGGAGTACCTGAGA GAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCATC TACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACATGGTTATGAAGATTACACAGTGGCTCATGCCGTGA CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA CATGGTGAACCCCATCTCTACTAAAGATACAAAAATTAT g
GGTTATGAAGATTACACAGTGGCTCATGCCGTGTATCCC CTCTGGGAGGCTGAGGCAAGTG agcacctgagatcatgagttcaagaccagcctg*caacat tgaaccccatctctactaaagatacaaaaattatccagg
TATGAAGATTACACAGTGGCTCA gatcccagcactttgggaggctgaggcaagtggagcacct agttcaagaccagcctggccaacatggtgaaaccccatct TACTAAAGATACAAAAATTATCCAGGTGTGGTG
ATGAAGATTACACAGTGGCTCATGCCGTGTATCCCAGCA TCTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAG gatcccagctattttgggaggctgaggcaagtggagcacct CATGGTGAACCCCATCTCTACTAAAGATACAAAAATTAT
atcccagcactttgggaggctgaggcaagtggagcacct CATGGTGAACCCCATCTCT TAAAGATACAAAAATTATCCAGGTGTGGTG
TCTGAGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAG GTGAAACCCCATCTCTACTAAAGATACAAAAATTATCCAG
GGGATGCTTAGTCAATTGTAGCACCTGAGATCATGAGTTC GTGAAACCCCATCTCTACTAAAGATACAAAAATTATCCAG
aggctgaggcaagtggagcacctgagatcatgagttcaag gtgaaaccccatctctactaaagatacaaaaattatccag
tggggcaagtggagcacctgagatcatgagttcaagacca gtgaaaccgtgtctctac aaagatacaaaaattatccaggtgtggtg
tgaggcaagtggagcacctgagatcatgagttcaagacca GAAATCCCATCTCTACTAAAGATACAAAAATTATCCAGGT
GAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGACCAG GAAACCCCATCTCTACTAAAGATACAAAAATTATCCAGGT
AGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGACCAG GAAACCCCATCTCTACTAAATAACA atccaggtgtggtg
aggcaatttgagctctgagatcatgagttcaagaccagc gaaaccccatctctgtgagatgcaaaaatta
GCAAGTGGAGCACCTGAGATCA AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT
CAAGTGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTG AATCCCATCTCTACTAAATATACAAAAATTATCCAGGTGT
caagtggagcacctgagatcatgagttcaagaccagcctg aaccccatctctactaaagatccaaaaattatccaggggt
AAGTGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGG AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT
AGTGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGC ACCCCGTTCTACTAAAGATACAAAAATTATCCAGGTGTG
AGTGCAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGC accccatctctactaaagatacaaaaattatccaggtgtg
GTGGAGCACCTGAGATCATGAGTTCAAGACCAGCATGGCC CCCCATCTCTACTAAAGATAC atccaggtgtggtg
GGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCA CATCTCTAATAAAGATACAAAAATTATCCAGGTGTGGTG
ggagcacctgagatcatgagttcaagaccaggggtggcaa CATCTCTACTAAAGATACAAAAATTATCCAGGTGTGCTG
ggagcacctgagatcatgagttcaagaccagcctggcaa CGTCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG
GAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCA CATCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG
```

Reads can effectively cover 95% of the genome or exome

Coverage of chr7 in a typical whole exome sequencing experiment

Genes

Read depth



Variants are detected when enough reads disagree with reference

```
40421551 40421561 40421571 40421581 40421591 40421601 40421611 40421621 40421631 40421641 40421651 40421661 40421671 40421681 40421691 40421701 40421711 40421
721tttgagcagacctatataagatggttatgaagattcacacagcggctcatgctgtgatcccagcactttgggaggctgaggcaagtggagcacctgagatcatgagttcaagaccagcctggccaacatggtgaaccccatctctactaaagatacaaaaattatccagggtgtggtg
.....A.....T.....
A tgaacagacctatataagatggtt tgaagattcacacagtggctcatgctgtgatcccagcac tgggaggctgagtcagtgaggcacctgagatcatgagtt ACCAGCCTGGCCAACATGGTGAACCCCATCTCTACTAAA ATACAAAAATTATCCAGGTGTGGTG
aca cagacctatataagatggtt aagatacacacagtggctcatgctgtgatcccagcactt GGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTC cagcctggccaacatggtgaaccccatctctactaaaga ACAAAAAATTATCCAGGTGTGGTG
acatt GACCTATATAAGATGGTTATGAAGATTACACAGTGGCTC CCTGTGATCCCAGCACTTTGGGAGGCTGAGGCAAGTGGAG ACCTGAGATCATGAGTTCAAGACCAGCCTGGACAACATGG AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT G
ACATTGAACAG ATATAAGATGGTTATGAAGATTACACAGTGGCTCATGCC tgatcccagcactttgggagg TGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGACCA GCCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA ATCCAGGTGTGGTG
ACATTGAACAGAC TCAGATGGTTATGAAGATTACACAGTGGCTCATGCCGTGT ATCCCAGCACTTTGGGAGGCTGAGGCAAGGGGAGCACCTG ATGAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCA CTCTACTAAAGATACAAAAATTA aggtgtggtg
acatttgaacagacctata aagatggttatgaagattcacacagtggctcatgctgtg TCCCAGCACTTTGGGAGGCTGAGGCAAGTGGAGCACCTGA ATGAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCA TACTACTAAAGATACAAAAATTATCCAGGTGTGGTG
acatttgaacagacctatata gatggttatgaagattcacacagtggctcatgctgtgat AGCACTTTGGGAGGCTGAGGCAAGGGGAGCACCTGA GAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCATC CTACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACAT gacctatataagatggttatgaagattcacacagtggctc CTGTAATCCCATCACTTTGGGAGGCTGAGGCAAGTGGAGC CCTGAGATCATGAGTTCAAGA AGCCTGGCCAACATCGTGAACCCCATATCTACTAAAGAT caaaattatccagggtgtggtg
ACATTGAACAGACCTATATAA TGGTTATGAAGATTACACAGTGGCTCATGCCGTGTGATCC cactttgggagtgctgaggcaagtggagcacctgagatcat CAAGACCAGCCTGGCCAACATGGTGAACCCCATCTCTAC AGAAATACAAAAATTATCCATGTGTGGTG
ACATTG ACCTATATAAGATGGTTATGAAGATTACACAGTGGCTCA TGTGATCCCAGCACTTTGGGAGGCTGAGGCAAGTGGAGCA CTGAGATCACGAGTTCAAGACCAGCCTGGCCAACATGGTC AACCCCATCTCTACTAAAGATACAAAAATTACCCAGGTGT GG
ACATTGAACAGACCTATATAAGA GGTTACGAAGATTACACAGTGGCTCATGCCGTGTGATCCC cacattgggaggctgaggcaagtggagcacctgagatcat AAGACCAGCCTGGCCAACATGGTGAACCCCATCTCTACT AAGATACAAAAATTATCCAGGTGTGGTG
acatttgaacagacctatataagat ttatgaagattcacacagtggctcatgctgtgatcccag CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA agcctggccaacatggtgaaccccatctctactaaagat AAAATTATCCAGGTGTGGTG
acatttgaacagacctatataagatggtt aagattcacacagtggctcatgctgtgatcccagcactt GGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTC GCCTGGCCAACATGGTGAAC CCCATCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACATTGAACAGACCTATATAAGATGGTTA agattcacacagggctcatgctgtgatcccagcacttt AGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAG CCTGGCCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA TTATCCAGGTGTGGTG
acatttgaacagacctatataagatggtta TTCACAGTGGCTCATGCCGTGTGATCCCAGCACTTTGGG GCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGAC CCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA atccagggtgtggtg
ACATTGAACAGACCTATATAAGATGGTTAT CAGTGGCTCATGCCGTGTGAT ACTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG CAACATGGTGAACCCCATCTCTACTAAAGATACAAAAAT TCCAGGTGTGGTG
ACATTGAACAGACCTATATAAGATGGTTATGAAG CAGTGGCTCATGCCGTGTGATC ACTTCGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG AACATGGTGAACCCCATCTCTACTAAAGATACAGAAAT aggtgtggtg
ACATTGAACAGACCTATATAAGATGGTTATGAAGAT CAGTGGCTCATGCCGTGTGATCC CCTCTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG ACATGGTGAACCCCATCTCTACTAAAGATACAAAAATTA GTGTGGTG
ACATTGAACAGACCTATCTAAGATGGTTATGAAGATT GCGGCTCATGCCGTGTATC CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA ACATGGTGAACCCCATCTATACTAAAGATACAAAAATTA TGTGGTG
ACATTGAACAGACCTATATAAGATGGTTATGAAGATTG CTCTTGCTGTGATCCCAGCACTTTGGGAGGCTGAGGCAA TGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCA TGGTGAACCCCATCTCTACTAAAGATACAAAAATTATCC ggtg
ACATTGAACAGACCTATATAAGATGGTTATGAAGATTG CTCATGCCGTGTGATCCCAGCACTTTGGGAGGCTGAGGCAA TGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCA GGTGAACCCCATCGCTACTAAAGATACAAAAATTATCCA GG
ACATTGAACAGACCTATATAAGATGGTTATGAAGATTCA GTGATCCCAGCACTTTGGGAGGCTGAGGCAAGTGGAGCAC GATCATGAGTTCAAGACCCGCTGGCCAACATGGTGAAC ccctctctactaaagatacaaaaattatccagggtgtggtg
AGATGGTTATGAAGATTACACAGTGGCTCATGCCGTGTA CCAGCACTTTGGGAGGCTGAGGCAAGTGGAGTACCTGAGA GAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCATC TACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACATGGTTATGAAGATTACACAGTGGCTCATGCCGTGTA CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA CATGGTGAACCCCATCTCTACTAAAGATACAAAAATTAT g
GGTTATGAAGATTACACAGTGGCTCATGCCGTGTGATCCC CTCTGGGAGGCTGAGGCAAGTG agcacctgagatcatgagttcaagaccagcctg*caacat tgaaccccatctctactaaagatacaaaaattatccagg
TATGAAGATTACACAGTGGCTCA gatcccagcactttgggaggctgaggcaagtggagcacct agttcaagaccagcctggccaacatggtgaaccccatct TACTAAAGATACAAAAATTATCCAGGTGTGGTG
ATGAAGATTACACAGTGGCTCATGCCGTGTGATCCCAGCA TCTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAG gatcccagctatttgggaggctgaggcaagtggagcacct CATGGTGAACCCCATCTCTACTAAAGATACAAAAATTAT
atcccagcactttgggaggctgaggcaagtggagcacct CATGGTGAACCCCATCTCT TAAAGATACAAAAATTATCCAGGTGTGGTG
TCTGAGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAG GTGAAACCCCATCTCTACTAAAGATACAAAAATTATCCAG
GGGATGCTTAGTCAATTGTAGCACCTGAGATCATGAGTTC GTGAAACCCCATCTCTACTAAAGATACAAAAATTATCCAG
aggctgaggcaagtggagcacctgagatcatgagttcaag gtgaaccccatctctactaaagatacaaaaattatccag
tggggcaagtggagcacctgagatcatgagttcaagacca gtgaaccgtgtctctac aaagatacaaaaattatccagggtgtggtg
tgaggcaagtggagcacctgagatcatgagttcaagacca GAAATCCCATCTCTACTAAAGATACAAAAATTATCCAGGT
GAGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGACCAG GAAACCCCATCTCTACTAAAGATACAAAAATTATCCAGGT
AGGCAAGTGGAGCACCTGAGATCATGAGTTCAAGACCAG GAAACCCCATCTCTACTAAATAACA atccagggtgtggtg
aggcaatttggctctgagatcatgagttcaagaccagc gaaccccatctctgtgaagatgcaaaaatta
GCAAGTGGAGCACCTGAGATCA AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT
CAAGTGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTG AATCCCATCTCTACTAAATATACAAAAATTATCCAGGTGT
caagtggagcacctgagatcatgagttcaagaccagcctg aaccccatctctactaaagatccaaaaattatccaggggt
AAGTGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGG AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT
AGTGGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGC ACCCCGTTCTACTAAAGATACAAAAATTATCCAGGTGTG
AGTGCAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGC accccatctctactaaagatacaaaaattatccagggtgtg
GTGGAGCACCTGAGATCATGAGTTCAAGACCAGCATGGCC CCCCATCTCTACTAAAGATAC atccagggtgtggtg
GGAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCA CATCTCTAATAAAGATACAAAAATTATCCAGGTGTGGTG
ggagcacctgagatcatgagttcaagaccaggggtggcaa CATCTCTACTAAAGATACAAAAATTATCCAGGTGTGCTG
ggagcacctgagatcatgagttcaagaccagcctggcaa CGTCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG
GAGCACCTGAGATCATGAGTTCAAGACCAGCCTGGCCAAC CATCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG
```


Variants are detected when enough reads disagree with reference

```
40421551 40421561 40421571 40421581 40421591 40421601 40421611 40421621 40421631 40421641 40421651 40421661 40421671 40421681 40421691 40421701 40421711 40421
721tttgagcagacctatataagatgttatgaagattcacacagcggctcatgcctgtgatccacacactttgggaggctgaggcaagtggagcacctgagatcatgagttcaagaccagcctggccaacatggtgaaccccatctctactaaagatacaaaaattatccagggtggtg
.....A.....T.....
A tgaacagacctatataagatgttt tgaagattcacacagtggctcatgcctgtgatccacacac tgggaggctgagtcagtgaggcacctgagatcatgagtt ACCAGCCTGGCCAACATGGTGAACCCCATCTCTACTAAA ATACAAAAATTATCCAGGTGTGGTG
aca cagacctatataagatgttt aagatacacacagtggctcatgcctgtgatccacacactt GGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTC cagcctggccaacatggtgaaccccatctctactaaaga ACAAAAAATTATCCAGGTGTGGTG
acatt GACCTATATAAGATGTTATGAAGATTCACACAGTGGCTC CCTGTGATCCCACTTTGGGAGGCTGAGGCAAGTGGAG ACCTGAGATCATGAGTTCAGACCAGCCTGGACAACATGG AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT G
ACATTTGAACAG ATATAAGATGTTATGAAGATTCACACAGTGGCTCATGCC tgaatccacacactttgggagg TGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAGACCA GCCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA ATCCAGGTGTGGTG
ACATTTGAACAGAC TCAGATGTTATGAAGATTCACACAGTGGCTCATGCCGT ATCCCACTTTGGGAGGCTGAGGCAAGGGGAGCACCTG ATGAGTTCAGACCAGCCTGGCCAACATGGTGAACCCCA CTCTACTAAAGATACAAAAATTA aggtgtggtg
acatttgaacagacccata aagatgtttatgaagattcacacagtggctcatgcctgtg TCCCACTTTGGGAGGCTGAGGCAAGTGGAGCACCTGA ATGAGTTCAGACCAGCCTGGCCAACATGGTGAACCCCA TATACTAAAGATACAAAAATTATCCAGGTGTGGTG
acatttgaacagacctatata gatgtttatgaagattcacacagtggctcatgcctgtgat ACCTTTGGGAGGCTGAGGCAAGGGGAGCACGTGA GAGTTCAGACCAGCCTGGCCAACATGGTGAACCCCATC CTACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACAT gacctatataagatgtttatgaagattcacacagtggctc CTGTAATCCCACTTTGGGAGGCTGAGGCAAGTGGAGC CCTGAGATCATGAGTTCAGGA AGCCTGGCCAACATCGTGAACCCCATATCTACTAAAGAT caaaaattatccagggtggtg
ACATTTGAACAGACCTATATAA TGTTATGAAGATTCACACAGTGGCTCATGCCGTGTGATCC cactttgggatgctgaggcaagtggagcacctgagatcat CAAGACCAGCCTGGCCAACATGGTGAACCCCATCTCTAC AGAAATACAAAAATTATCCATGTGTGGTG
ACATTTG ACCTATATAAGATGTTATGAAGATTCACACAGTGGCTCA TGTGATCCCACTTTGGGAGGCTGAGGCAAGTGGAGCA CTGAGATCACGAGTTCAGACCAGCCTGGCCAACATGGTC AACCCCATCTCTACTAAAGATACAAAAATTACCCAGGTGT GG
ACATTTGAACAGACCTATATAAGA GTTACGAAGATTCACACAGTGGCTCATGCCGTGTGATCCC cacattgggaggctgaggcaagtggagcacctgagatcat AAGACCAGCCTGGCCAACATGGTGAACCCCATCTCTACT AAGATACAAAAATTATCCAGGTGTGGTG
acatttgaacagacctatataagat tttatgaagattcacacagtggctcatgcctgtgatccac CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA agcctggccaacatggtgaaccccatctctactaaagat AAAATTATCCAGGTGTGGTG
acatttgaacagacctatataagatgtt aagattcacacagtggctcatgcctgtgatccacactt GGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTC GCCTGGCCAACATGGTGAAC CCCATCTCTACTAAAGATACAAAAATTATCCAGGTGTGGT G
ACATTTGAACAGACCTATATAAGATGTTA agattcacacagaggctcatgcctgtgatccacacttt AGGCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAG CTTGGCCAACATGGTGAACCCCATCTCTACTAAAGATAC TTATCCAGGTGTGGTG
acatttgaacagacctatataagatgtta TTACACAGTGGCTCATGCCGTGTGATCCCACTTTGGG GCTGAGGCAAGTGGAGCACCTGAGATCATGAGTTCAGAC CCAACATGGTGAACCCCATCTCTACTAAAGATACAAAA atccagggtgtggtgG
ACATTTGAACAGACCTATATAAGATGTTAT CAGTGGCTCATGCCGTGTGAT ACTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG CAACATGGTGAACCCCATCTCTACTAAAGATACAAAAAT TCCAGGTGTGGTG
ACATTTGAACAGACCTATATAAGATGTTATGAG CAGTGGCTCATGCCGTGTGATC ACTTCGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG AACATGGTGAACCCCATCTCTACTAAAGATACAGAAATT aggtgtggtgG
ACATTTGAACAGACCTATATAAGATGGTTATGAAGAT CAGTGGCTCATGCCGTGTGATCC CCTCTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATG ACATGGTGAACCCCATCTCTACTAAAGATACAAAAATTA GTGTGGTG
ACATTTGAACAGACCTATCTAAGATGGTTATGAAGATT GCGGCTCATGCCGTGTATC CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA ACATGGTGAACCCCATCTATACTAAAGATACAAAAATTA TGTGGTG
ACATTTGAACAGACCTATATAAGATGGTTATGAAGATTG CTCTTGCTGTGATCCAGCACCTTTGGGAGGCTGACGCAA TGGAGCACCTGAGATCATGAGTTCAGACCAGCCTGGCCA TGGTGAACCCCATCTCTACTAAAGATACAAAAATTATCC ggtgG
ACATTTGAACAGACCTATATAAGATGGTTATGAAGATTG CTCATGCCGTGTGATCCAGCACCTTTGGGAGGCTGAGGCAA TGGAGCACCTGAGATCATGAGTTCAGACCAGCCTGGCCA GGTGAACCCCATCGCTACTAAAGATACAAAAATTATCCA GG
ACATTTGAACAGACCTATATAAGATGGTTATGAAGATTCA GTGATCCAGCACCTTTGGGAGGCTGAGGCAAGTGGAGCAC GATCATGAGTTCAGACCAGCCTGGCCAACATGGTGAAC ccatctctactaaagatacaaaaattatccagggtgtggtgG
AGATGGTTATGAAGATTCACACAGTGGCTCATGCCGTGA CCAGCACCTTTGGGAGGCTGAGGCAAGTGGAGTACCTGAGA GAGTTCAGACCAGCCTGGCCAACATGGTGAACCCCATC TACTAAAGATACAAAAATTATCCAGGTGTGGTG
ACATGGTTATGAAGATTCACACAGTGGCTCATGCCGTGA CTTTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGA CATGGTGTAAACCCCATCTCTACTAAAGATACAAAAATTAT gG
GGTTATGAAGATTCACACAGTGGCTCATGCCGTGTGATCCC CTCTGGGAGGCTGAGGCAAGTG agcacctgagatcatgagttcaagaccagcctg*caacat tgaaccccatctctactaaagatacaaaaattatccagg G
TATGAAGATTCACACAGTGGCTCA gatccagcactttgggaggctgaggcaagtggagcacct agttcaagaccagcctggccaacatggtgaaccccatct TACTAAAGATACAAAAATTATCCAGGTGTGGTG
ATGAAGATTCACACAGTGGCTCATGCCGTGTGATCCAGCA TCTGGGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAG gatccagcctttgggaggctgaggcaagtggagcacct CATGGTGAACCCCATCTCTACTAAAGATACAAAAATTAT G
gatccagcctttgggaggctgaggcaagtggagcacct CATGGTGAACCCCATCTCT CTAAAGATACAAAAATTATCCAGGTGTGGTG
atccagcactttgggaggctgaggcaagtggagcacct CATGGTGAACCCCATCTCTACTAAAGATACAAAAATGAT G
TCTGAGAGGCTGAGGCAAGTGGAGCACCTGAGATCATGAG GTGAAACCCCATCTCTACTAAAGATACAAAAATTATCCAG G
GGGATGCTTAGTCAATTGTAGCACCTGAGATCATGAGTTC GTGAAACCCCATCTCTACTAAAGATACAAAAATTATCCAG G
aggctgaggcaagtggagcacctgagatcatgagttcaag gtgaaccccatctctactaaagatacaaaaattatccag G
tggggcaagtggagcacctgagatcatgagttcaagacca gtgaaccgtgtctctac aaagatacaaaaattatccagggtgtggtgG
tgaggcaagtggagcacctgagatcatgagttcaagacca GAAATCCCATCTCTACTAAAGATACAAAAATTATCCAGGT G
GAGGCAAGTGGAGCACCTGAGATCATGAGTTCAGACCAG GAAACCCCATCTCTACTAAAGATACAAAAATTATCCAGGT G
AGGCAAGTGGAGCACCTGAGATCATGAGTTCAGACCAGC GAAACCCCATCTCTACTAAATAACA atccagggtgtggtgG
aggcaatttgagctctgagatcatgagttcaagaccagc gaaaccccatctctgctgaagatgcaaaaatta g
GCAAGTGGAGCACCTGAGATCA AACCCCATCTCTACTAAAGATACAAAAATTATCCAGGTGT G
CAAGTGGAGCACCTGAGATCATGAGTTCAGACCAGCCTG AATCCCATCTCTACTAAATATACAAAAATTATCCAGGTGT G
caagtggagcacctgagatcatgagttcaagaccagcctg aaccccatctctactaaagatccaaaattatccaggggt G
AAGTGGAGCACCTGAGATCATGAGTTCAGACCAGCCTGG AACCCTCTCTACTAAAGATACAAAAATTATCCAGGTGT G
AGTGGAGCACCTGAGATCATGAGTTCAGACCAGCCTGGC ACCCCGTTTCTACTAAAGATACAAAAATTATCCAGGTGT G
AGTGCAGCACCTGAGATCATGAGTTCAGACCAGCCTGGC accccatctctactaaagatacaaaaattatccagggtgtg G
GTGGAGCACCTGAGATCATGAGTTCAGACCAGCATGGCC CCCCATCTCTACTAAAGATAC atccagggtgtggtgG
GGAGCACCTGAGATCATGAGTTCAGACCAGCCTGGCCA CATCTCTAATAAAGATACAAAAATTATCCAGGTGTGGTG G
ggagcacctgagatgtgagttcaagaccaggggtggcaa CATCTCTACTAAAGATACAAAAATTATCCAGGTGTGCTGG G
ggagcacctgagatcatgagttcaagaccagcctggcaa CGTCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG G
GAGCACCTGAGATCATGAGTTCAGACCAGCCTGGCCAAC CATCTCTACTAAAGATACAAAAATTATCCAGGTGTGGTG
```

Variants are detected when enough reads disagree with reference

```
40421571 40421581 40421591 40421601
ttatgaagattcacacagcggctcatgcctgtgatccca
.....T.....
tt tgaagattcacacagtggtcatgcctgtgatccca
tt  aagatacacacagtggtcatgcctgtgatccca
TTATGAAGATTACACAGTGGCTC CCTGTGATCCCA
TTATGAAGATTACACAGTGGCTCATGCC tgaatccca
TTATGAAGATTACACAGTGGCTCATGCCCTGT ATCCCA
ttatgaagattcacacagtggtcatgcctgtg TCCCA
ttatgaagattcacacagtagctcatgcctgtgat
ttatgaagattcacacagtggtc CTGTAATCCCA
TTATGAAGATTACACAGTGGCTCATGCCCTGTGATCC
TTATGAAGATTACACAGTGGCTCA TGTGATCCCA
TTACGAAGATTACACAGTGGCTCATGCCCTGTGATCCC
ttatgaagattcacacagtggtcatgcctgtgatccca
tt  aagattcacacagtggtcatgccagtgaatccca
TTA  agattcacacagaggctcatgcctgtgatccca
tta  TTCACACAGTGGCTCATGCCCTGTGATCCCA
TTAT  CAGTGGCTCATGCCCTGTGAT
```

← Position

← Reference sequence

← Inferred patient sequence

← Raw sequence reads from patient

C > T variant

The human genome is big

3 billion basepairs

$$X \quad \quad \quad 0.1\%$$

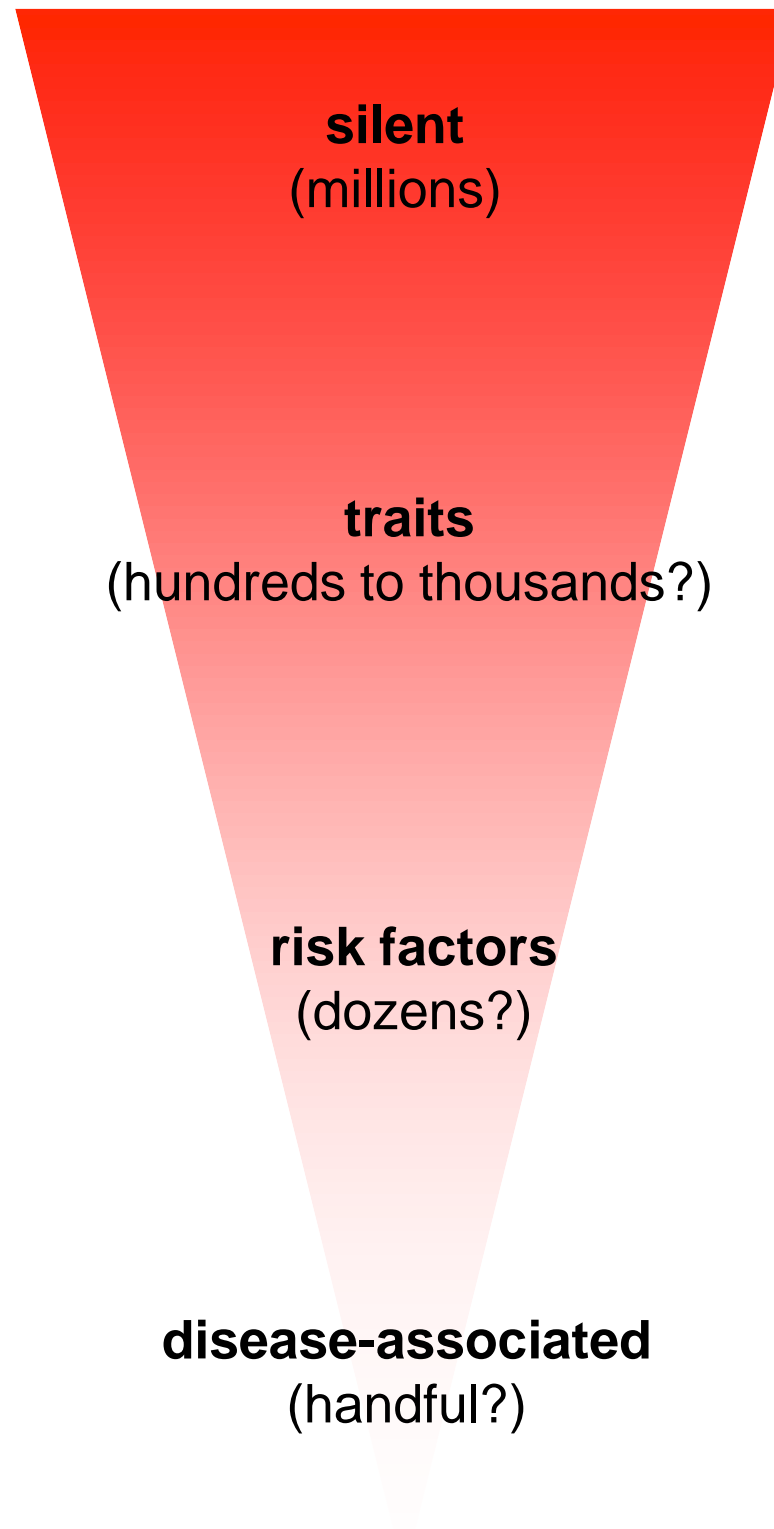
3 million sites of variation
between any two individuals

A parts list of variants from one individual

| Table 2. SNPs Identified through Whole-Genome Sequencing of DNA from the Proband.* | |
|---|--------------------|
| SNP Type | No. of SNPs |
| Nongene | 2,255,102 |
| Gene | 1,165,204 |
| Intron | 1,064,655 |
| Promoter | 60,075 |
| 3' UTR | 16,350 |
| 5' UTR | 3,517 |
| Splice regulatory site | 2,089 |
| Splice site | 112 |
| Synonymous | 9,337 |
| Stop→stop | 17 |
| Nonsynonymous | 9,069 |
| Stop→gain | 121 |
| Stop→loss | 27 |
| Total | 3,420,306 |

Lupski et al, *NEJM* 2010

Of these 3 million, which are medically relevant?



How do you interpret 3,000,000 variants?

3,000,000

Commonly encountered?

5-10% of variants have not been seen frequently in controls

In a gene coding region?

0.25% of rare variants lie within a gene coding region

Alter protein (or affect splicing)?

~50% of variants in genes alter protein or affect splicing

Affect a known disease gene?

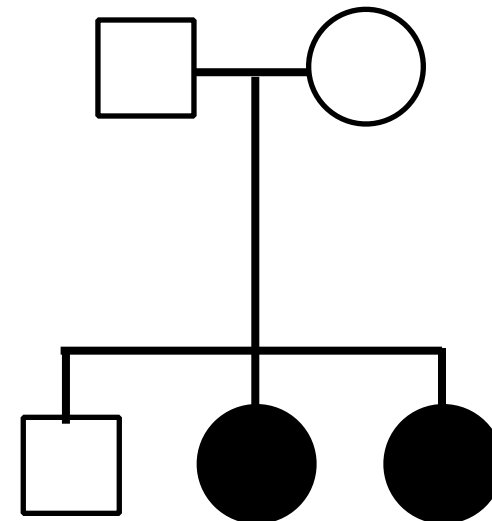
~10% of genes have been associated with human disease

~20-40

WGS has revolutionized Mendelian genetics

~20-40

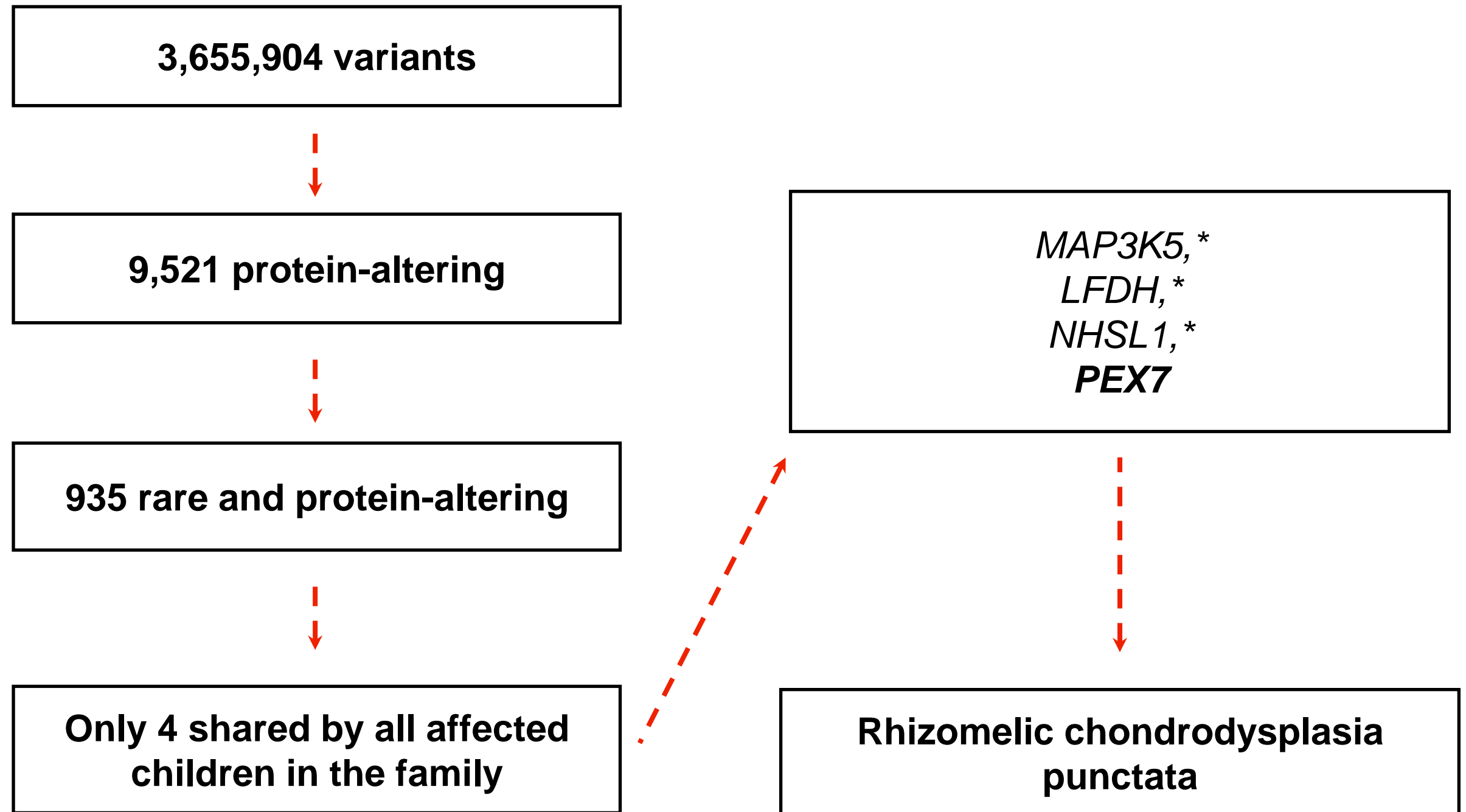
**Inheritance
pattern**



Sequencing parents
& available siblings

Single disease genes!

Three siblings w/intellectual disability, cataracts and seizures



Challenges to clinical adoption

- A “data deluge”
- Many results are of uncertain medical significance
- Insufficient numbers of geneticists and genetic counselors to handle the flood of clinical data
- Clinicians and patients will need education and training
- Collaborative efforts will be required to amass and organize data
- Birth of a new specialty?

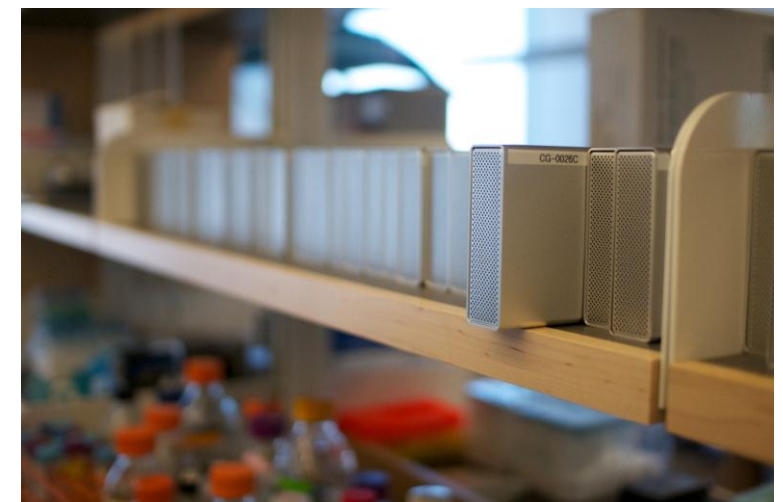
Challenges to clinical adoption

- Genomic medicine requires cross-disciplinary skills
 - Genetics
 - Genomics
 - Lab Medicine
 - Informatics
- Example:
 - Processing time for 1 human genome: **one week**
 - Requires highly specialized computer systems and expertise that **stresses most hospital infrastructures**

Challenges to clinical adoption

- 3 billion bases & 3 million variants **don't fit in a patient chart**
- 1 patient's whole genome data = **1 terabyte hard drive**
- Costs of secure storage!
- **Electronic Medical Records** are not yet equipped to handle this

40 patient whole genomes, on a shelf



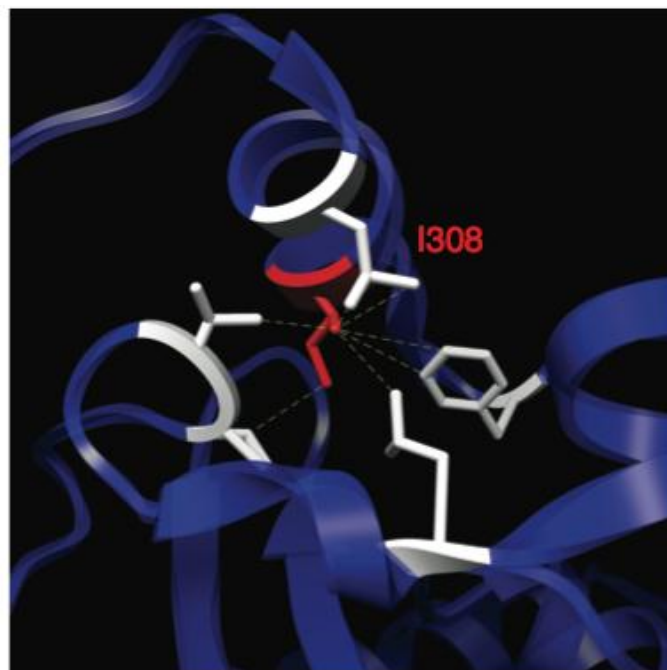
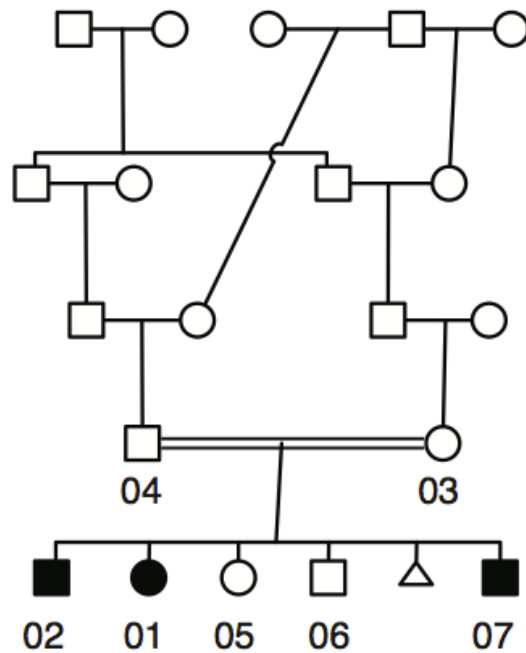
Ethical, legal, & social issues

- Privacy and confidentiality
- Stigmatization, and discrimination
- Results of uncertain medical significance, emotional distress, follow-up tests, and cost
- Genetic paternalism vs. right to control your genomic information

Standards of evidence

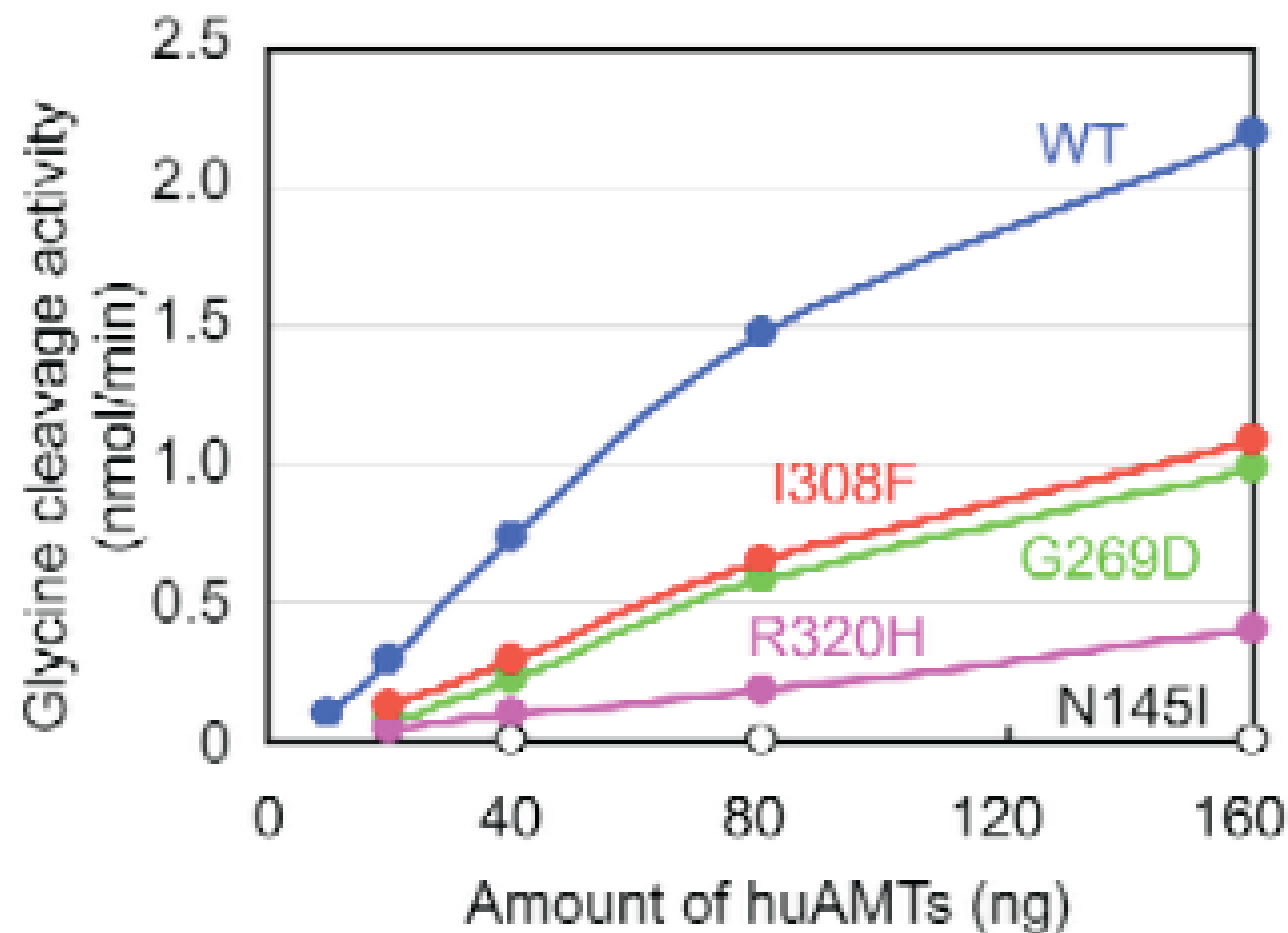
- Need to be careful applying what we think we know
- Many currently reported disease associations are based on old studies of with small sample sizes, or using outdated technologies
- These will need to be re-evaluated as we get better at interpreting genetic results
- We also need to be cost-effective, and demonstrate that this leads to better care

The importance of functional follow-up



- ✓ Family with three children with autism, intellectual disability, and seizures
- ✓ WES performed in 3 children + parents
- ✓ Homozygous I308F mutation in **AMT**, a cause of **glycine encephalopathy**
- ✓ Rare (never before seen in controls)
- ✓ Alters a highly conserved residue, “looks deleterious”
- ✓ But patients with glycine encephalopathy often die in the neonatal period

AMT p.I308F is a hypomorphic LOF mutation



- **I308F: defects in protein folding and enzyme function, but retains some residual activity**

- ✓ These children had a **mild version** of glycine encephalopathy with autism and epilepsy
- ✓ Had been **previously undiagnosed** in the family (requires lumbar puncture, liver biopsy)

Grappling with “incidental findings”

The Incidentalome A Threat to Genomic Medicine

Isaac S. Kohane, MD, PhD

Daniel R. Masys, MD

Russ B. Altman, MD, PhD

GENOMIC MEDICINE IS POISED TO OFFER A BROAD ARRAY of new genome-scale screening tests. However, these tests may lead to a phenomenon in which multiple abnormal genomic findings are discovered, analogous to the “incidentalomas” that are often discovered in radiological studies. If practitioners pursue these unexpected genomic findings without thought, there may be disastrous consequences. First, physicians will

There is a rich literature in radiology on the “incidentaloma,” which is a finding (most commonly a mass) found on computed tomography or magnetic resonance imaging studies ordered for symptoms or concerns totally unrelated to the gland in which the mass is found. The workup of an incidentaloma is complicated by concerns that it may be associated with malignant disease and, at least initially, the lack of good data on the prevalence of malignant disease in the general population. Incidentalomas occur because imaging modes do not only report on the areas of direct clinical concern but, incidentally, report on all organs in the field of view.¹

This phenomenon of possible incidental genomic find-

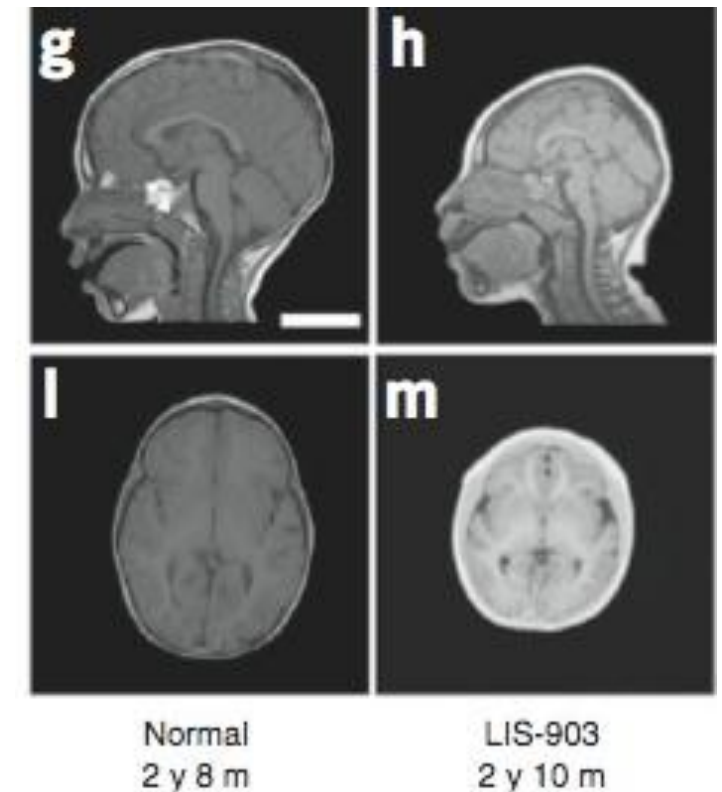
Other open issues in genomic medicine

- Primary vs. Secondary Findings
- Genes of Uncertain Significance
- Research vs. Clinical boundaries
- Stretching our definitions of disease
- Collaborative resources and standards

If you're interested in this, you're in a pretty good place

Mutations in *WDR62*, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture

Timothy W Yu¹⁻⁷, Ganeshwaran H Mochida¹⁻⁷, David J Tischfield¹⁻⁵, Sema K Sgaier^{1-5,8}, Laura Flores-Sarnat⁹, Consolato M Sergi^{10,11}, Meral Topçu¹², Marie T McDonald¹³, Brenda J Barry¹⁻⁵, Jillian M Felie¹⁻⁵, Christine Sunu¹⁻⁵, William B Dobyns¹⁴, Rebecca D Folkerth¹⁵, A James Barkovich¹⁶ & Christopher A Walsh¹⁻⁶



One of the first demonstrations of using NGS in humans
Identification of a new gene for human microcephaly from 6 families

[Yu et al, *Nat Genetics*, 2010]

If you're interested in this, you're in a pretty good place

Genetic Defect in *CYP24A1*, the Vitamin D 24-Hydroxylase Gene, in a Patient with Severe Infantile Hypercalcemia

Andrew Dauber, Thutrang T. Nguyen, Etienne Sochett, David E. C. Cole, Ronald Horst, Steven A. Abrams, Thomas O. Carpenter, and Joel N. Hirschhorn

Novel Microcephalic Primordial Dwarfism Disorder Associated with Variants in the Centrosomal Protein Ninein

Andrew Dauber, Stephen H. LaFranchi, Zoltan Maliga, Julian C. Lui, Jennifer E. Moon, Cailin McDeed, Katrin Henke, Jonathan Zonana, Garrett A. Kingman, Tune H. Pers, Jeffrey Baron, Ron G. Rosenfeld, Joel N. Hirschhorn, Matthew P. Harris, and Vivian Hwa

**Whole exome sequencing of children with endocrine disorders:
new genes for pediatric hypercalcemia and dwarfism**

[Dauber et al, *JCEM* 2012]

If you're interested in this, you're in a pretty good place

REPORT

Exome Sequencing and Functional Validation in Zebrafish Identify *GTDC2* Mutations as a Cause of Walker-Warburg Syndrome

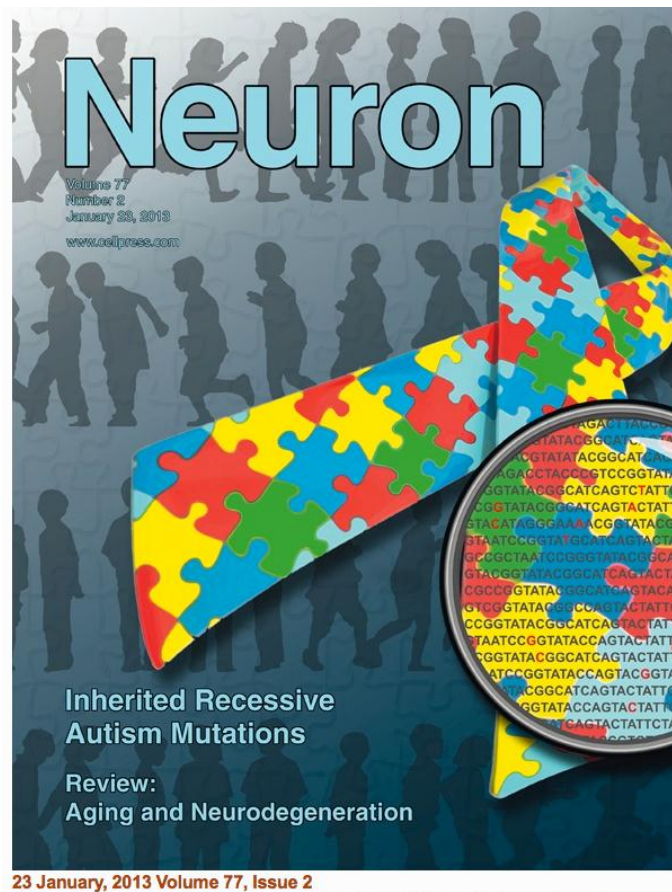
M. Chiara Manzini,^{1,2} Dimira E. Tambunan,^{1,2} R. Sean Hill,^{1,2} Tim W. Yu,^{1,2} Thomas M. Maynard,³ Erin L. Heinzen,⁴ Kevin V. Shianna,⁴ Christine R. Stevens,⁵ Jennifer N. Partlow,^{1,2} Brenda J. Barry,^{1,2} Jacqueline Rodriguez,^{1,2} Vandana A. Gupta,^{1,6} Abdel-Karim Al-Qudah,⁷ Wafaa M. Eyaid,⁸ Jan M. Friedman,^{9,10} Mustafa A. Salih,¹¹ Robin Clark,¹² Isabella Moroni,¹³ Marina Mora,¹⁴ Alan H. Beggs,^{1,6} Stacey B. Gabriel,⁵ and Christopher A. Walsh^{1,2,5,*}

Larger cohorts:

**WES in 19 families identifies a novel gene responsible for
a well-known neuromuscular disorder**

[Manzini et al, *AJHG*, 2012]

If you're interested in this, you're in a pretty good place



Neuron
Article

Using Whole-Exome Sequencing to Identify Inherited Causes of Autism

Timothy W. Yu,^{1,2,3,4,5,6,7,32,*} Maria H. Chahrour,^{1,2,3,4,5,7,32} Michael E. Coulter,^{1,2,3,5} Sarn Jiralerspong,⁸ Kazuko Okamura-Ikeda,⁹ Bulent Ataman,¹⁰ Klaus Schmitz-Abe,^{1,2,5} David A. Harmin,¹⁰ Mazhar Adli,¹¹ Athar N. Malik,¹⁰ Alissa M. D'Gama,⁵ Elaine T. Lim,¹² Stephan J. Sanders,¹³ Ganesh H. Mochida,^{1,2,3,5,6} Jennifer N. Partlow,^{1,2,3} Christine M. Sunu,^{1,2,3} Jillian M. Felie,^{1,2,3} Jacqueline Rodriguez,^{1,2,3} Ramzi H. Nasir,^{5,14} Janice Ware,^{5,14} Robert M. Joseph,^{4,15} R. Sean Hill,^{1,2,3,5} Benjamin Y. Kwan,¹⁶ Muna Al-Saffar,^{1,2,17} Nahit M. Mukaddes,¹⁸ Asif Hashmi,¹⁹ Soher Balkhy,²⁰ Generoso G. Gascon,^{6,18,21} Fuki M. Hisama,²² Elaine LeClair,^{5,14} Annapurna Poduri,^{5,23} Ozgur Oner,²⁴ Samira Al-Saad,²⁵ Sadika A. Al-Awadi,²⁶ Laila Bastaki,²⁶ Tawfeg Ben-Omran,^{27,28} Ahmad S. Teebi,^{27,28} Lihadh Al-Gazali,¹⁷ Valsamma Eapen,²⁹ Christine R. Stevens,⁷ Leonard Rappaport,^{4,5,14} Stacey B. Gabriel,⁷ Kyriacos Markianos,^{1,2,5} Matthew W. State,¹³ Michael E. Greenberg,¹⁰ Hisaaki Taniguchi,⁹ Nancy E. Braverman,⁸ Eric M. Morrow,^{4,30,31} and Christopher A. Walsh^{1,2,3,4,5,7,*}

**Larger cohorts:
WES in >150 consanguineous families
to find new recessive autism genes**

[Yu et al, *Neuron* 2013]

If you're interested in this, you're in a pretty good place

The Boston Globe

Health & wellness

Search

Q

NEWSMETROARTSBUSINESSSPORTSOPINIONLIFESTYLEMAGAZINEINSIDERS

TODAY'S PAPERMY SAVED

FOOD & DININGHEALTH & WELLNESSSTYLETRAVELNAMESCOMICSCROSSWORDSPRING STYLE

In this section

Health & wellness

US policy clouds approvals of medical marijuana

Survivors of pediatric brain tumors share a special bond

Q&A with Dr. Paul Offit

Finding drug-free techniques to treat chronic pain

Shape of wine glass determines how much we pour, study finds

Missed diagnosis most common malpractice accusation

'Expecting Better' from your doctors during pregnancy

Boston-area team to study DNA sequencing in newborns

By Carolyn Y. Johnson | GLOBE STAFF | SEPTEMBER 04, 2013

ARTICLE

COMMENTS

SUBSCRIBE

PRINT

REPRINTS

E-MAIL

SHARE

By early next year, parents of newborns at two Boston hospitals will have the chance to participate in the first randomized study of the medical and ethical repercussions of sequencing the DNA of babies. The research is part of a major federal effort to finally settle a debate that has raged for years about the possible benefits and harms of finding out such information.

The five-year study, a joint effort of Boston Children's Hospital and Brigham and Women's Hospital, was one of four proposals selected for funding, federal health officials announced Wednesday in a press conference. The National Institutes of Health will spend \$25 million over five years to support the program, \$6 million of which will support the Boston-based study.

BabySeq:
WES and WGS on 120 normal newborns and 120 NICU patients
-> Clinical outcomes, healthcare utilization, and safety

If you're interested in this, you're in a pretty good place

What is Claritas Genomics?



A Diagnostic Testing Company

Arose out of Boston Children's Hospital

- CLIA licensed molecular lab for genomic medicine
- Specialty testing based on Boston Children's research and clinical knowledge

Partnership with major pediatric hospitals in the US

Partnering with industry (Life Technologies, Cerner)

Partnering with country health systems (Saudi Genome Project, US Million Veterans Project)

**Starting an Interpretive Genomics Service
at Boston Children's Hospital**

```
40421571 40421581 40421591 40421601
ttatgaagattcacacagcggtcatgcctgtgatccca
.....T.....
tt tgaagattcacacagtggctcatgcctgtgatccca
tt aagatacacacagtggctcatgcctgtgatccca
TTATGAAGATTCACACAGTGGCTC CCTGTGATCCCA
TTATGAAGATTCACACAGTGGCTCATGCC tgatccca
TTATGAAGATTCACACAGTGGCTCATGCCTGT ATCCCA
ttatgaagattcacacagtggctcatgcctgtg TCCCA
ttatgaagattcacacagttagctcatgcctgtgat A
ttatgaagattcacacagtggctc CTGTAATCCCA
TTATGAAGATTCACACAGTGGCTCATGCCTGTGATCC
TTATGAAGATTCACACAGTGGCTCA TGTGATCCCA
TTACGAAGATTCACACAGTGGCTCATGCCTGTGATCCC
ttatgaagattcacacagtggctcatgcctgtgatccca
tt aagattcacacagtggctcatgccagtgatccca
TTA agattcacacagaggctcatgcctgtgatccca
tta TTCACACAGTGGCTCATGCCTGTGATCCCA
TTAT CAGTGGCTCATGCCTGTGAT
```

Questions!

