What’s in your genes?
Whole genome sequencing and its impact on personalized medicine
Personalized Medicine

Tailoring of medical care based on the genetic characteristics of each patient
The Story of Nicholas Volker

DNA sequencing and personalized medicine helped diagnose and treat his disease

- Unusual symptoms affecting bowel
- Endured over 100 surgeries
- No diagnosis or treatment found

More on this story later!
Tonight’s Talks

What is DNA and how does DNA sequencing work?

How can DNA sequencing improve diagnosis and prevention of disease?

How can personalized medicine help find the best treatments?
Outline

• What is DNA and why is it important?

• What are the consequences of alterations in our DNA sequence?

• How does “DNA sequencing” work?
Outline

• What is DNA and why is it important?

• What are the consequences of alterations in our DNA sequence?

• How does “DNA sequencing” work?
DNA: Instruction Manual for Life

Deoxyribonucleic acid = “DNA”
DNA is found in cells
DNA is divided into genes

“Instruction manual”  “Chapter”  “Machine”

DNA  Gene  Protein
Not all DNA codes for proteins

“Non-coding DNA”
“Junk DNA”

Gene A
Gene B

Less than 2% of all DNA encodes for protein
ATGC – The DNA Code

DNA is made up of building blocks called “nucleotides”

Human genome is made of 3 billion nucleotides

ATTGCAACACATTGTAGT
TAACGTTTGTAACATCA

Genetic sequence is the order of A’ s, T’ s, G’ s and C’ s that make up your DNA
DNA is divided into genes

“Instruction manual”

“Chapter”

“Machine”

DNA

Gene

Protein

Michale Strock, Andrzej Joachimiak and Argonne National Laboratory
DNA to RNA to Protein

“Chapter”

ATTGCAACATTGTAG

Gene (DNA)

“Photocopy”

AUUGCAACACAUUUGUAG

Message (RNA)

“Machine”

Amino acid

Protein

Michale Strock, Andrzej Joachimiak and Argonne National Laboratory
Questions?
Outline

• What is DNA and why is it important?

• What are the consequences of alterations in our DNA sequence?

• How does “DNA sequencing” work?
Changes in DNA can result in abnormal proteins that cause disease

Inherited

Mistakes during lifetime

Damage: UV Exposure, Smoking, etc.

Abnormal protein
A mutation is an accidental, permanent change in a DNA sequence.
Differences in DNA Sequence can alter proteins

Example: Sickle cell anemia – mutated hemoglobin
Not ALL differences in DNA sequence alter proteins

Gene (DNA)

Message (RNA)

Protein

Code is redundant: UUG and UUA make “Silent mutation”
Differences in DNA Sequence Can Be Good…

Differences contribute to variation in a population
“Single Nucleotide Polymorphisms”

A SNP is a one nucleotide difference in the DNA sequence compared to the reference genome.
Challenge of Sequencing:
What DNA changes are bad?

“Reference genome”

Thousands of SNPs can be found – but we don’t always know the consequences of each one.
Questions?
Outline

• What is DNA and why is it important?

• What are the consequences of alterations in our DNA sequence?

• How does “DNA sequencing” work?
DNA sequencing has come a long way…

<table>
<thead>
<tr>
<th>Year</th>
<th>Time to sequence 3 billion nucleotides</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>30 million years</td>
</tr>
<tr>
<td>2012</td>
<td>1 day</td>
</tr>
</tbody>
</table>
Sequencing Costs

Cost per Genome

Moore’s Law

National Human Genome Research Institute
genome.gov/sequencingcosts
Principles of DNA sequencing

• Use the DNA replication process that happens normally in cells.

• Spy on which nucleotide is added to the new strand in sequence.
DNA Replication

DNA Polymerase

Double-stranded DNA

ATTGCAACATTGTAG
TAACGTTGTAACATC
What is the process of DNA sequencing?

• Patient gives biological sample to be sequenced.

• DNA is isolated from the cells in the sample and cut into small pieces.
DNA is attached to a glass slide

ATTGCAACATTGTA

[Diagram of DNA sequence attached to a glass slide]
Detecting nucleotide sequence

Repeat cycle of A, T, G, C until sequencing complete
Aligning results to reference genome

Nucleotide peaks from the sequencer

ATTCGCAACCAATGTAAGT

Search:

Reference genome

“Did you mean: Reference genome?”

Tasks:
- Piece together DNA sequences
- Align to reference
Once aligned, compare all changes to reference genome

It’s up to doctors and scientists to figure out which changes are meaningful
Summary

• Pieces of DNA called genes encode for specific proteins, but most of the genome is “non-coding”

• Changes in DNA can alter proteins or be “silent”

• DNA sequencing uses the process of DNA replication to determine the order of nucleotides in a DNA sample
Tonight’s Talks

What is DNA and how does DNA sequencing work?

How can DNA sequencing improve diagnosis and prevention of disease?

How can personalized medicine help find the best treatments?
Part 2: How can DNA sequencing improve diagnosis and prevention of disease?

Adrianna San Roman
Clare Malone
Leah Liu
Outline

• Types of DNA sequencing used in the clinic

• Applying sequencing in diagnosis and detection of disease

• Using sequencing information to assess risk factors of disease
Outline

• Types of DNA sequencing used in the clinic

  • Applying sequencing in diagnosis and detection of disease

  • Using sequencing information to assess risk factors of disease
Single Gene Testing

Example: Hemophilia

Single gene testing relies on clues from symptoms or family history
Single Nucleotide Polymorphism analysis

- Only sequence single nucleotide polymorphisms (SNPs) which account for ~80% of genetic diversity
- Can look for a specific SNP or do genome wide SNP analysis (70,000+ SNPs)
- Often used for genealogical analysis
- Direct to consumer testing (23andMe, Genelex etc.)
“Exome” Sequencing

• Sequences every gene, but skips the “junk” DNA
• Don’t need to know which gene to look at
• Faster and cheaper than sequencing everything
Whole Genome Sequencing

• Sequence ALL the DNA – have all the information.

• Why do this?
  – “Junk” DNA may actually be important for some diseases
  – Know your sequence for the future – drug predictions, disease predictions, etc.
Questions?
Outline

• Types of DNA sequencing used in the clinic

• Applying sequencing in diagnosis and detection of disease

• Using sequencing information to assess risk factors of disease
Using genetic tools to diagnose a patient

Nic presents with severe bowel disease

Single gene tests were all negative

Exome sequencing—16,142 differences!

How do you figure out which change is the important one?

Ignore “silent” mutations

Select genes where both copies have alterations

Select changes predicted to be damaging to protein function

Exclude genes where changes are common

ONE G to A substitution!
National Institutes of Health Undiagnosed Diseases project: “The Clinic of Last Resort”

Using SNPs to diagnose (and discover) diseases

http://commonfund.nih.gov/diseases/
Whole genome sequencing in the Neonatal Intensive Care Unit

- ~1/3 of infants in NICU have genetic diseases
- Children’s Mercy Hospital and Clinics in Kansas City can now sequence the entire genome for ~$7,000
- From blood sample to diagnosis in 50 hours
- Faster diagnosis, and coverage of all genetic disorders

Early detection of some genetic disorders is critical.

Phenylketonuria occurs in 1 out of 15,000 births in the United States, but with early intervention patients can have normal lifespan and normal cognitive abilities.
Newborn screening

All newborns in the state of Massachusetts are screened for 30 common genetic disease where early intervention is critical
Some time soon…

It will be technologically and economically feasible to replace newborn screening with whole genome sequencing, which could detect all genetic disorders with one test.
Questions?
Outline

• Types of DNA sequencing used in the clinic

• Applying sequencing in diagnosis and detection of disease

• Using sequencing information to assess risk factors of disease
Genes and environment interact

“Nature”
Blood Type

“Nurture”
Language you speak

Many common diseases are a result of this interaction (ex. Diabetes, Heart Disease, Lung Cancer)
Identifying genetic “risk factors”

People with disease

People without disease
Single gene risk factors

BRCA1 mutations affect DNA repair

→ 12% risk of breast cancer

→ 60% risk of breast cancer

Interventions: Increased surveillance (earlier mammograms) or elective breast removal
Most diseases are the result of multiple genes in combination with environment

Type 2 Diabetes

~40 associated genetic variations
Each variation increases an individual’s lifetime risk by only a few percent

Diet, exercise, Body Mass Index, etc.

Clinical factors are currently better at assessing risk
The “heritability gap”

- Type 2 Diabetes is 20-40% heritable
- Known associated genes only account for 10% of the heritability
- There must be other risk factors that we haven’t found yet

![Genetic risk explained by known variations](chart.png)

- Environmental Risk
- Genetic Risk

Percent Risk

- 0
- 20
- 40
- 60
- 80
- 100
- 120
New genetic techniques may improve patient care

- Improved diagnosis of unusual symptoms or rare diseases
- Early detection of rare genetic diseases
- Risk assessment for complex diseases
Questions?
Part 3: How can personalized medicine help find the best treatments?

Adrianna San Roman
Clare Malone
Leah Liu
Outline

• Testing single genes can answer:
  – What treatment is most effective?
  – What dosage works the best?
  – Which treatment has the fewest side effects?

• Using whole genome and exome sequencing to find ways to treat patients
Medicines are not equally effective in all patients

Medicine **effectiveness** in patient populations

Adapted from Spear et al. 2001, “Clinical application of pharmacogenetics.”
Targeted therapies attack specific mutated proteins
Targeted therapies attack specific mutated proteins
Mutations in different genes can cause the same disease

When using targeted therapies, patients must be tested for their causative mutation before treatment.
B-raf mutation in melanoma

B-raf

Controlled cell division

Mutated B-raf

Uncontrollable cell division

Tumor formation
Vemurafenib inhibits mutated B-raf, but not normal B-raf
Vemurafenib inhibits mutated B-raf, but not normal B-raf

- Vemurafenib
- Normal B-raf
- Prevent overactive cell division
- Mutated B-Raf
- Tumors shrink
Vemurafenib effectiveness during clinical trials

Before

After

## Other examples of targeted therapies

<table>
<thead>
<tr>
<th>Disease</th>
<th>Mutation/Marker</th>
<th>Treatment name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer</td>
<td>Estrogen receptor</td>
<td>Aromasin (exemestane)</td>
</tr>
<tr>
<td>Chronic myelogenous leukemia</td>
<td>BCR-ABL</td>
<td>Gleevec (imatinib)</td>
</tr>
<tr>
<td>Colon Cancer</td>
<td>EGFR</td>
<td>Erbitux (cetuximab)</td>
</tr>
<tr>
<td>Cystic Fibrosis</td>
<td>CFTR (G551D)</td>
<td>Kalydeco (ivacaftor)</td>
</tr>
</tbody>
</table>
Questions?
Using genetics to find the best dose: Pharmacogenetics

- Warfarin (coumadin) is the most popular anticoagulant

Warfarin

VKORC1 promotes clotting

Clotting of blood
Gene variants impact warfarin function

VKORC1

Warfarin

CYP2C9 breaks down warfarin

Warfarin is excreted from body
Gene variants impact warfarin function

VKORC1 variants: VKORC1 levels are decreased

CYP2C9 variants: less warfarin excreted, stays in body longer

Warfarin

Patients require lower doses of Warfarin

May reduce thousands of strokes/year
Pharmacogenetics: predicting side effects of medicines

Two patients with epilepsy or bipolar disorder treated with Carbamazepine:

- HLA-B
  - Low risk of skin infection
- HLA-B*1502 variant
  - Severe skin reaction
Questions?
Whole genome and exome sequencing to find treatments

• Genome information can tell us:
  – New mutations/new diseases might be treated with existing treatments
  – Which genes should be targets for development of new treatments
Identification of Nic’s mutation: XIAP

- XIAP is involved in the immune system and preventing cell death, not associated with bowel disease
Confirmation that XIAP mutation leads to disease

Unaffected cells Stimulate specific immune response

Response

Nic’s cells No response

Conclusion: XIAP is defective in patient’s cells
XIAP deficiency is a serious blood disorder

- XIAP mutation can lead to life threatening imbalance in immune system cells
- Treatment: Transplant to replace blood/immune system

http://www.nlm.nih.gov/medlineplus/blooddisorders.html,
Advantages of Sequencing

• Advantages over single gene testing when
  – Many diseases share same symptoms
  – Unique presentation of disease
  – May be cheaper, may save time

• Learn about biology of new disease pathway
  – What is XIAP’s function in the bowel?
Summary

• **Targeted therapies** can treat diseases caused by specific mutated proteins

• **Pharmacogenetics** involves using genetics to find the best dose of a medicine and minimize side effects

• Whole genome and exome sequencing can also be used to find ways to treat patients
Future of Personalized Medicine: Prevention vs. Reaction

Today

- Few exome/whole genome sequencing examples
- Testing single genes is still cheaper than sequencing

Near Future

- $1000 genome
- Better diagnosis, detection, risk assessment
- More targeted therapies
- Increased use of Pharmacogenetics

Far Future

- Lower health care costs
- Understand junk DNA contribution to disease?
- Everyone has genome sequenced?
Lecture Summary

• Personalized medicine uses an individual’s genetic characteristics to inform medical care

• DNA sequencing is a technology that can be used to find alterations in our DNA that can impact health

• Different types of DNA sequencing can be used to diagnose and detect disease, to assess risk factors, and find ways to treat disease
Tour of the McCarroll Lab!
Thank you!

SITN would like to acknowledge the following organizations for their generous support.

Harvard Medical School
Office of Communications and External Relations
Division of Medical Sciences

The Harvard Graduate School of Arts and Sciences (GSAS)

The Harvard Graduate Student Council (GSC)

The Harvard Biomedical Graduate Students Organization (BGSO)

The Harvard/MIT COOP

https://sitn.hms.harvard.edu  SITNBoston@gmail.com  @SITNHarvard  Facebook.com/SITNBoston