Entering the CRISPR-Cas9 Era

Yangzhong Tang
Functional Genomics Platform@ Whitehead Institute
June 1st, 2015
Outline

• The Human Genome
  o Basics and timeline
  o Exponential increase in genomic data
  o How to use the big data for biological discoveries
    • Functional Genomics

• CRISPR-Cas systems
  o Discovery
  o Mechanisms
  o Applications

• Functional Genomics using CRISPR-Cas9
Basics of the Human Genome
Many Human Diseases Are Genetic Diseases

• Chromosome abnormalities
  o Aneuploidy, Deletion, Inversion, Translocation
    • Down syndrome (Trisomy 21, or three #21 chromosomes), Turner syndrome (single X chromosome)

• Single gene defects
  o Dominant, Recessive, X-linked
    • Cystic fibrosis, Sickle cell disease

• Multifactorial problems
  o Combined effect of gene and environment
    • Heart defect, cleft lip

• Common diseases
  o Inherited heart diseases, some type 2 diabetes, cancer
Timeline of the Human Genome
Exponential Increase in Genomic Data

- Complete cataloging of human genomes with decreasing cost

Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP)
Available at: [www.genome.gov/sequencingcosts](http://www.genome.gov/sequencingcosts)

Going Forward with the Big Data

• Some public databases of human genomic data
  o The Cancer Genome Atlas (TCGA)
  o Catalog of Somatic Mutations in Cancer (COSMIC)
  o International Cancer Genome Consortium (ICGC)
  o 1000 Genomes Project
  o HapMap Project
  o The Encyclopedia of DNA Elements (ENCODE) Project
  o …

• How can we gain meaningful information for biological discoveries and therapeutic use?
Gaining Insights from the Big Data

• Association studies
  o Genome-wide association studies (GWAS)
    • 1st success: Two SNPs on gene CFH found to associate with Age-related Macular Degeneration (AMD)

Haines, J., et. al., Science, 2005

• As of today, over 4000 SNPs are found to associate with 200 diseases and traits.
Gaining Insights from the Big Data

• Functional Genomics

  o “Specifically, functional genomics refers to the development and application of global (genome-wide or system-wide) experimental approaches to assess gene function by making use of the information and reagents provided by structural genomics. It is characterized by high throughput or large-scale experimental methodologies combined with statistical and computational analysis of the results.”

  o **Aim:** To understand the complex relationship between genotype and phenotype in a genome-wide scale.

  o **Approach:** Often uses high-throughput approach instead of “gene-by-gene” method.
Gaining Insights from the Big Data

- Functional Genomics
  - A powerful tool in target discovery in various disease areas.
  - A mock example:
    - Question: To identify gene targets whose mutations or overexpressions are critical for Triple-negative Breast Cancer.
    - Typical Approach:

  ![Diagram showing cell line selection, siRNA or shRNA treatment, measurement of viability, and data QC, normalization, analysis with secondary screen.]

  - TNBC lines
  - Control lines
  - Measurement of viability etc.
  - Data QC, normalization, analysis
  - Secondary screen
However, siRNA and shRNA have significant off-target effect, and inefficient knockdown level.

http://www.biospring.de/html/eng/rna1.html

• The Human Genome
  o Basics and timeline
  o Exponential genomic data generation
  o How to gain insights

• CRISPR-Cas systems
  o Discovery
  o Mechanisms
  o Applications

• Functional Genomics using CRISPR-Cas9
Discovery of the CRISPR-Cas System

- CRISPR (clustered regularly interspaced palindromic repeats), was first found in the genome of *Escherichia coli*, then later in many bacteria and archaea.
- Originally thought to have no significant biological functions.
- Key findings in 2005 by bioinformatics analysis that many spacer sequences in CRISPR match foreign plasmids, phages, and viruses.
- In 2007, Horvath and colleagues at Danisco (a food ingredient company), presented the 1st experimental evidence for the role of type II CRISPR system in bacteria adaptive immunity.


Three types of CRISPR-Cas Systems

Hsu, Lander, and Zhang, *Cell*, 2014
Mechanism of the type II CRISPR-Cas System

Use of CRISPR-Cas in Editing Human Genome

- In 2012, Jinek et al. proposed usage of a single synthetic guide RNA for gene editing.

- In 2013, Cong et al. and Mali et al. demonstrated usage of CRISPR-Cas system as a human genome-editing tool.


Applications of the CRISPR-Cas9 System

Hsu, Lander, and Zhang, *Cell*, 2014
A mock example:

- **Question**: To identify gene targets whose mutations or expressions are critical for Triple-negative Breast Cancer.
- **Typical Approach**:

  - **Cell line selection**
  - **siRNA or shRNA treatment**
  - **Measurement of viability etc. Data QC, normalization, analysis**
  - **Secondary screen**
A mock example:

- **Question:** To identify gene targets whose mutations or expressions are critical for Triple-negative Breast Cancer.
- **New Approach:**
  - Cell line selection
  - Lentiviral sgRNA delivery
  - Sequencing, data analysis
  - Secondary screen
## Functional Genomics using siRNA/shRNA or sgRNA

<table>
<thead>
<tr>
<th></th>
<th>siRNA or shRNA</th>
<th>CRISPR-Cas9 / sgRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-target effect</td>
<td>Significant</td>
<td>mild</td>
</tr>
<tr>
<td>Gene knockdown level</td>
<td>Variant</td>
<td>complete</td>
</tr>
<tr>
<td>Applications</td>
<td>Gene knockdown</td>
<td>Gene knockout</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transcriptional modulation (CRISPRi and CRISPRa)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect of epigenetic modulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Functions of noncoding elements</td>
</tr>
</tbody>
</table>
Functional Genomics Formats

Pooled Genome-wide CRISPR Screens

Functional Genomics Facility in Whitehead

- **Mission**: Teaming with the MIT and Whitehead scientific community, we strive to make novel biological discoveries with high therapeutic potentials for human diseases.

- **Main approach**: pooled lentiviral CRISPR-Cas9 screens
Since the invention of the next-generation sequencing (NGS), we have been generating unprecedented amount of genomic data.

How to gain insights from human genomic data will have huge impact on therapeutics.

Functional Genomics is an important tool in understanding gene functions in a systematic scale. However, it has been limited by the siRNA and shRNA tools.

The recent breakthrough in gene-editing technology, CRISPR-Cas, is accelerating the speed and increasing the accuracy in all aspects of biological research and discovery.

Equipped with CRISPR-Cas technology, Functional Genomics will help us make great discoveries with huge therapeutic potential.
Acknowledgment

Sabatini Lab
• David Sabatini
• Tim Wang
• Kevin Krupczack
• Yorick Post
• Edie Valeri
• Kathleen Ottina

Whitehead Institute