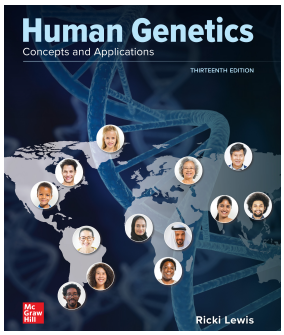
Because learning changes everything.®



Chapter 14

Genomes

Thirteenth Edition

Ricki Lewis

1

Learning Outcomes 1

1. Explain how linkage studies, positional cloning experiments, and gene mapping led to the idea to sequence human genomes.
2. Distinguish between the two approaches used to sequence the first human genomes.
3. Discuss the types of information that the first sequenced human genomes provided.
4. Define *reference genome*.
5. Explain why genome sequencing requires multiple copies of a genome.

2

Learning Outcomes 2

6. Identify limitations of exome and genome sequencing.
7. Discuss practical aspects of implementing exome and genome sequencing in health care.
8. Explain how a genome-level view of chromosomes reveals details of crossing over.
9. Discuss the possible consequences of sequencing the genomes of newborns.

3

Sequencing Genomes of the Deceased

- The Genomic Postmortem Research Project of the Marshfield Clinic Foundation in Wisconsin is sequencing the genomes of 300 deceased individuals and comparing the information to data in electronic medical records.
- The purpose of this project is to use the information gained from this undertaking to assist clinicians in preventing, detecting, and treating specific diseases.
- It would help to better identify specific genes involved in certain forms of cancer and genes that may affect human responses to certain medications, such as antidepressants, cancer drugs, painkillers, and others.

4

From Genetics To Genomics

Genetics is a young science, genomics is younger still

Term **genome** was coined in 1920, to refer to a complete set of chromosomes and its genes

- Now it refers to all the DNA in a haploid set of chromosomes

Term **genomics** was coined in 1986

- Indicates the study of genomes

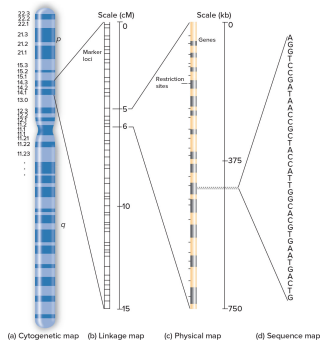
5

Linkage Studies 1

- Genetic maps have increased in detail and resolution
- **Cytogenetic map:** Distinguishes DNA sequences that are at least 5,000 kilobases apart
- **Linkage map:** Distinguishes genes hundreds of kilobases apart
- **Physical map:** Distinguishes genes tens of kilobases apart
- **Sequence map:** Depicts the order of bases

6

Linkage Studies 2



Access the text alternative for slide images.

7

Positional Cloning

- Gene-by-gene approach that matches single genes to specific diseases
- Begins with a phenotype, and gradually identifies a causative gene, localizing it to a part of a chromosome
- Yielded discoveries of genes that cause diseases as Duchenne muscular dystrophy, cystic fibrosis and Huntington disease

8

The Human Genome Project 1

Idea to sequence the human genome emerged in the 1980s with several goals

Officially started in 1990

- $3 billion, 15-year project
- Under the direction of the DOE and NIH

Draft of the human genome in 2001

Finished sequence in 2003

Represents the work of thousands of researchers in an international collaboration

9

The Human Genome Project 2

The project set aside 3% of its budget for the Ethical, Legal and Social Implications (ELSI) Research Program.

- Recognizing the impact of the human genome project on public policy
- Prevents misuse of information and genotypic discrimination

Eventually, an international consortium as well as a private company, Celera Genomics, sequenced the first human genomes.

10

Key Inventions

Expressed Sequence Tag (EST) technology

- Enables researches to find protein-encoding genes
- cDNAs that are expressed in a particular cell type

DNA microarrays

- Display short DNA molecules
- Important in sequencing and assessing gene expression

11

Sequencing the Human Genome 1

- Researchers cut several genomes-worth of DNA into overlapping pieces of about 40 Kilobases
- Followed by randomly cutting into small fragments and sequencing
- Computer algorithms were used to search for overlaps
- By overlapping the pieces, the software derives the overall DNA sequence

12

Sequencing the Human Genome 2

The U.S. government-funded international consortium used a clone-by-clone approach

- Aligned pieces one chromosome at a time

Celera Genomics, a private company, used a whole genome shotgun approach

- Shattered the entire genome; then rebuilt it

 13

13

Deriving a DNA Sequence

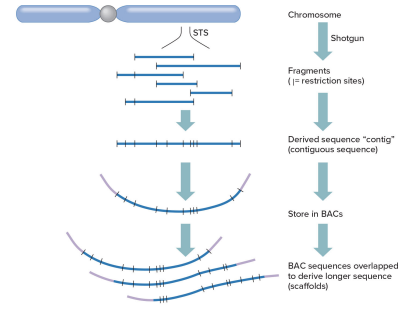
Random fragments:	AGTCCT CTAG AGCTA CTACT TAGAGT CCTAGC
Alignment:	CTAG TAGAGT AGTCCT CCTAGC AGCTA CTACT
Sequence:	CTAGAGTCCTAGCTACT

Access the text alternative for slide images.

 14

14

Route 1 to Sequencing the Human Genome



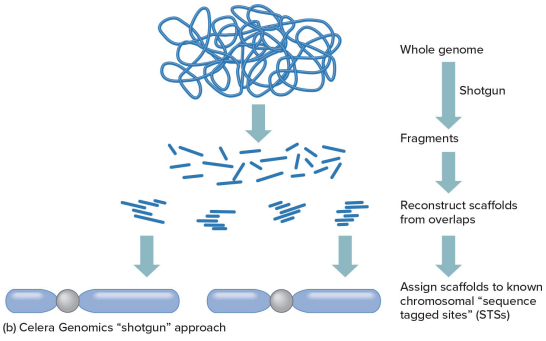
(a) International Human Genome Mapping Consortium "BAC by BAC"
(BAC = bacterial artificial chromosome)

Access the text alternative for slide images.

 15

15

Route 2 to Sequencing the Human Genome



(b) Celera Genomics "shotgun" approach

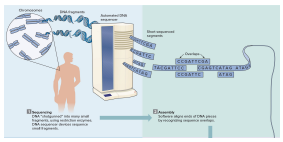
Access the text alternative for slide images.

 16

16

Sequencing Genomes 1

1. **Sequencing**
   DNA "shotgunned" into many small fragments, using restriction enzymes. DNA sequencer devices sequence small fragments.
2. **Assembly**
   Software aligns ends of DNA pieces by recognizing sequence overlaps.

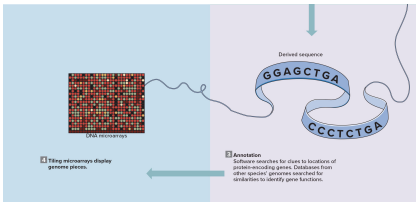


Access the text alternative for slide images.

 17

17

Sequencing Genomes 2



3. **Annotation**
   Software searches for clues to locations of protein-encoding genes. Databases from other species' genomes searched for similarities to identify gene functions.
4. **Tilling microarrays display genome pieces.**

 18

18

Types of Information in Human Genomes

The first two human genomes sequenced were from genetic pioneers J. Craig Venter and James Watson

- They showed that the numbers of copies of short sequences—copy number variants, or CNVs—contribute significantly to genetic variation.

The third person to have his genome sequenced was called, simply, "YH."

- He is Han Chinese, an East Asian population that accounts for 30% of modern humanity

Each of the three men has about 1.2 million SNPs, but a unique collection.

- About 0.07% of our SNPs may affect our phenotypes.

 19

19

Analysis of Human Genome Content

After sequencing genomes became possible, attention turned to refining the process, cataloging human variation, and discovering gene functions

To ease sequence comparisons and interpretations, researchers derive a **reference genome** sequence

- A digital DNA sequence assembled from the most common base at each point in many sequenced genomes.
- It is haploid

 20

20

Improving Speed and Coverage

Sequencing the first human genomes took 6 to 8 years; today it can be done in hours.

Improvements in sequencing technology enabled researchers to work with many more copies of an individual's genome, which is termed *coverage*.

- At least 28 human genome copies are necessary to ensure that most sequences are represented in the final derived sequence.

Sequencing genomes provides much more information than sequencing exomes

 21

21

The Ongoing Goal: Annotation

Description and significance of a particular gene variant

Includes

- Normal function of the gene
- Mode of inheritance
- Genotype (heterozygote, homozygote, compound heterozygote);
- Frequency of a variant in a particular population
- Classification as benign, likely benign, variant of uncertain significance, likely pathogenic, or pathogenic.

 22

22

Limitations of Genome Sequencing

Genome sequencing does not provide a complete picture of health.

Technically, it will not detect

- Copy number variants
- Mitochondrial DNA
- Uniparental disomy
- Gene-gene and gene-environment interactions

In a conceptual sense, genome information must be interpreted to be useful.

 23

23

Practical Medical Matters 1

With the introduction of consumer genetic testing in 2008, the possibility of testing genes for variants that indicate only risk, was suddenly available to all

Genetic and genomic testing as part of health care must meet certain practical criteria.

- The most important requirement for is clinical utility.

A DNA test result alone is not sufficient to diagnose a disease, but may support a clinical diagnosis based on symptoms and the results of other types of tests.

- Knowing that a mutation is present can motivate a person to seek further testing

 24

24

Practical Medical Matters 2

The nuances of Mendel's laws affect the interpretation of genome sequence data.

A DNA test alone is not sufficient for diagnosis because of

- incomplete penetrance (genotype does not always foretell phenotype),
- variable expressivity (different severities in different individuals),
- epistasis (gene-gene interactions),
- genetic heterogeneity (mutation in more than one gene causing a phenotype), and
- environmental influences (epigenetics).

25

A Genomic View Expands Knowledge

Different investigations have sequenced genomes to differing degrees.

- *Exome sequencing* refers to the exons (the protein-encoding parts)
- *Genomic sequencing* to more than an exome but less than a full genome
- *Whole genome sequencing* to entire genomes derived from many overlapped copies

Genome-wide association studies

- Consult sets of single nucleotide polymorphisms (SNPs) dispersed among the chromosomes

26

A New View of Crossing Over 1

Icelandic genomes enabled researchers an unprecedented view of two processes that foster human genetic diversity

- They uncovered the sites of 4.5 million crossover events that recombined parental chromosomes.
- They identified more than 200,000 *de novo* mutations.

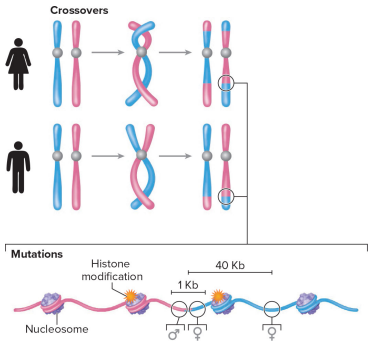
The findings reveal that both crossing over and generation of *de novo* mutations are not as random as had been thought

- At least 35 genes affect crossover frequency and location.

In addition, the two processes—crossing over and *de novo* mutations—are connected

27

A New View of Crossing Over 2



28

Sequencing the Genomes of Newborns

Knowing the genome sequence of an individual from birth can have medical advantages

- However, obtaining genome information from birth also raises issues of use or abuse in a variety of scenarios.

The US government funded a pilot study called BabySeq in two Boston-area hospitals

- The results attest to the potential value of newborn genome sequencing
  - The intervention detected childhood-onset diseases that hadn't yet caused symptoms in 15 (9.4%) of the babies.

29

Mc Graw Hill

Because learning changes everything.®

www.mheducation.com

30