

1000 Genomes Project: Insights into Human Genetic Variation

COMPREHENSION • VOCABULARY • DISCUSSION

INTERACTIVE ONLINE TEST

Take the full interactive test with instant feedback →

<https://www.esl-tests.com/reading/c1/1000-genomes-project-insights-human-genetic-variation>



Reading Passage

Read the passage carefully. Each paragraph is labelled with a letter for easy reference.

A The 1000 Genomes Project aims to provide a deep characterization of human genome sequence variation as a foundation for investigating the relationship between genotype and phenotype. Here we present results of the pilot phase of the project, designed to develop and compare different strategies for genome-wide sequencing with high-throughput platforms. We undertook three projects: low-coverage whole-genome sequencing of 179 individuals from four populations; high-coverage sequencing of two mother-father-child trios; and exon-targeted sequencing of 697 individuals from seven populations. We describe the location, allele frequency and local haplotype structure of approximately 15 million single nucleotide polymorphisms, 1 million short insertions and deletions, and 20,000 structural variants, most of which were previously undescribed. We show that, because we have catalogued the vast majority of common variation, over 95% of the currently accessible variants found in any individual are present in this data set. On average, each person is found to carry approximately 250 to 300 loss-of-function variants in annotated genes and 50 to 100 variants previously implicated in inherited disorders. We demonstrate how these results can be used to inform association and functional studies. From the two trios, we directly estimate the rate of de novo germline base substitution mutations to be approximately 10^{-8} per base pair per generation. We explore the data with regard to signatures of natural selection, and identify a marked reduction of genetic variation in the neighbourhood of genes, due to selection at linked sites. These methods and public data will support the next phase of human genetic research.

Vocabulary Glossary

Key words and phrases from the passage. Study them before attempting the exercise below.

WORD / PHRASE	DEFINITION	EXAMPLE SENTENCE
genotype	the genetic makeup of an organism; what genes it has	<i>The genotype of a plant can affect how tall it grows.</i>
phenotype	the set of observable traits of an organism, like eye color	<i>The phenotype of the flower is red, but it can have different genotypes.</i>
haplotype	a group of genes that a child inherits from one parent	<i>Researches study haplotypes to understand how traits are passed down.</i>
polymorphism	a variation in the DNA sequence that is common in a population	<i>Polymorphisms can explain why some people are more resistant to diseases.</i>
allele	a different form of a gene that can cause different traits	<i>Having two different alleles for a gene can lead to mixed traits.</i>
inherited	passed from parents to children through genes	<i>Eye color is an inherited trait from a person's parents.</i>
mutation	a change in a DNA sequence that can affect a trait	<i>Some mutations can lead to diseases, while others have no effect.</i>
natural selection	the process by which organisms better suited to their environment survive and reproduce	<i>Natural selection can lead to evolution over many generations.</i>

Vocabulary Exercise — Fill in the Blank

Use one word or phrase from the glossary above to complete each sentence. Each item is used only once. Answers are on the final page.

1. The _____ of a person can affect physical traits like height and eye color.
2. A _____ can result in different traits like flower color within the same species.
3. Scientists study _____ to understand how certain traits are passed from parents to offspring.
4. _____ occur when there is a change in the DNA sequence.
5. _____ traits, such as blood type, are passed down from parents to their children.
6. _____ helps organisms adapt to their environment over many generations.

Comprehension Questions

These questions are different from the online practice test. Choose the best answer (A, B, C, or D). Answers and explanations are on the final page.

Question 1

What is one purpose of the pilot phase of the 1000 Genomes Project?

- A. To develop and compare different sequencing strategies
- B. To sequence the genomes of all living humans
- C. To create a new type of DNA
- D. To eliminate genetic disorders

Question 2

How many populations were involved in the low-coverage whole-genome sequencing?

- A. Two
- B. Four
- C. Seven
- D. Ten

Question 3

What does the passage imply about structural variants?

- A. Most were previously known
- B. They are the least common variation
- C. They were mostly undescribed before this project
- D. They are not important in genetic studies

Question 4

What finding about genetic variations does the passage highlight?

- A. Genetic variations are entirely random
- B. There is a reduction of variation near genes
- C. All variations are harmful
- D. Variations are more common in older individuals

Question 5

Why is the data from the 1000 Genomes Project significant for future research?

- A. It will allow the creation of new organisms
- B. It provides a catalog of common genetic variations
- C. It proves that all mutations are beneficial
- D. It eliminates the need for further genetic studies

Discussion & Writing Prompts

Each prompt references a specific detail from the passage above. Use for classroom discussion or a short written response (150–200 words).

1. The passage mentions that each person carries 250 to 300 loss-of-function variants. How might this information impact medical research and treatment options?

2. According to the passage, low-coverage whole-genome sequencing was performed on 179 individuals from four populations. What challenges might arise when expanding this research to include more diverse populations?

3. The passage states that the rate of de novo germline base substitution mutations is estimated at 10^{-8} per base pair per generation. How could this finding influence our understanding of genetic diseases?

Answer Key

COMPREHENSION QUESTIONS

Q1 A

Q2 B

Q3 C

Q4 B

Q5 B

VOCABULARY EXERCISE

FIB1 genotype

FIB2 polymorphism

FIB3 haplotypes

FIB4 Mutations

FIB5 Inherited

FIB6 Natural selection

Comprehension Question Explanations

Why the correct answer is right — and why each wrong option is incorrect.

1. What is one purpose of the pilot phase of the 1000 Genomes Project?

✓ A — To develop and compare different sequencing strategies

The passage states the pilot phase was designed to develop and compare sequencing strategies.

✗ B — To sequence the genomes of all living humans

The project focuses on sequencing diverse samples, not all humans.

✗ C — To create a new type of DNA

Creating new types of DNA is not mentioned in the passage.

✗ D — To eliminate genetic disorders

The project aims to study genetic variation, not eliminate disorders.

2. How many populations were involved in the low-coverage whole-genome sequencing?

✓ B — Four

The passage mentions low-coverage sequencing was done on individuals from four populations.

✗ A — Two

The passage does not mention two populations for this part of the project.

✗ C — Seven

Seven populations were involved in exon-targeted sequencing, not low-coverage.

✗ D — Ten

Ten populations are not mentioned in the passage.

3. What does the passage imply about structural variants?

✓ C — They were mostly undescribed before this project

The passage notes most structural variants were previously undescribed.

✗ A — Most were previously known

The passage states that most structural variants were previously undescribed.

✗ B — They are the least common variation

The passage does not specify the frequency of structural variants compared to others.

✗ D — They are not important in genetic studies

The passage does not claim they are unimportant.

4. What finding about genetic variations does the passage highlight?

✓ B — There is a reduction of variation near genes

The passage describes a reduction in genetic variation near genes due to selection.

✗ A — Genetic variations are entirely random

The passage does not state that genetic variations are random.

✗ C — All variations are harmful

The passage does not suggest all variations are harmful.

✗ D — Variations are more common in older individuals

The passage does not discuss variations being more common in older individuals.

5. Why is the data from the 1000 Genomes Project significant for future research?

✓ B — It provides a catalog of common genetic variations

The passage highlights that the project provides a catalog of common genetic variations.

✗ A — It will allow the creation of new organisms

The passage does not discuss creating new organisms.

✗ C — It proves that all mutations are beneficial

The passage does not claim all mutations are beneficial.

✗ D — It eliminates the need for further genetic studies

The passage suggests ongoing research will use this data, not eliminate further studies.

MORE C1 LEVEL READING PRACTICE

Sociology Trust Modern Societies — <https://www.esl-tests.com/reading/c1/sociology-trust-modern-societies>

Economics Global Art Market — <https://www.esl-tests.com/reading/c1/economics-global-art-market>

Digital Memory Technology Changing Human Recall — <https://www.esl-tests.com/reading/c1/digital-memory-technology-changing-human-recall>