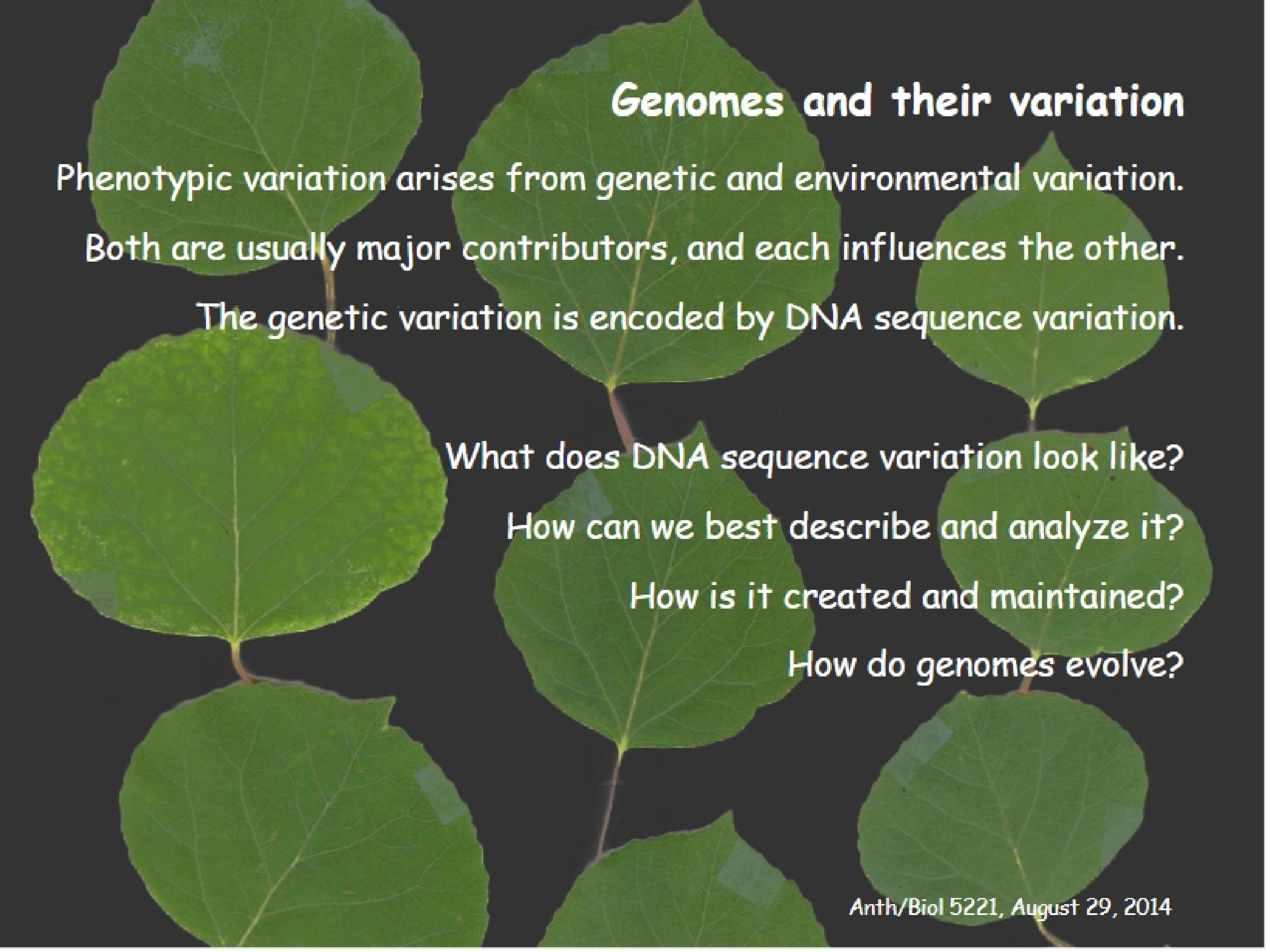




Genomes and their variation

Phenotypic variation arises from genetic and environmental variation.

Both are usually major contributors, and each influences the other.

The genetic variation is encoded by DNA sequence variation.

What does DNA sequence variation look like?

How can we best describe and analyze it?

How is it created and maintained?

How do genomes evolve?

How does a chimp differ
from a human?

Better hair!

Fewer self-doubts!

40,000 amino-acid substitutions!
(~20,000 on each line of descent)

Assuming 22,000 proteins, the
difference is less than 2/protein,
or roughly 1/250 amino acids.

~18,000,000 nucleotide
substitutions!

Given 3×10^9 base pairs per haploid
genome, that's 1/167 base pairs, or
0.6% different (99.4% the same).

How did these nt and aa
substitutions (and other genomic
changes) occur on the chimp and
human lineages, while *billions* of
others did *not* occur?

And why did they take 5 myr?

1 September 2005 | www.nature.com/nature | \$10

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

nature



STAR FORMATION

A massive protostar unveiled

CANCER IMMUNOLOGY

How tumours dupe T cells

AIR POLLUTION

China's NO₂ build-up
seen from space

NATUREJOBS

Membrane proteomics

THE CHIMPANZEE GENOME



Substitution is a *population process*

An *evolutionary substitution* occurs when:

- (1) **Mutation creates a new allele.** (Initially there is only one copy of this allele in the population.)
- (2) **The new allele *increases in frequency*** (briefly increasing genetic *variation* or "*polymorphism*").
- (3) **The old alleles are *lost*.** (The population then contains *only descendants of the new allele*.)

Whether the new mutation is *lost* from the population, or eventually *fixed*, is determined during the *polymorphic phase* (2), which may last a long time.

Thus present-day *variation* and long-term *evolution* are connected through the *population processes* that *change allele frequencies*.

The most important of these processes or "forces" are *genetic drift* and *natural selection*.

During the next few weeks we will develop a general quantitative framework for thinking about genetic variation at the DNA sequence level, and about rates of allele frequency change under the combined influences of drift and selection.

