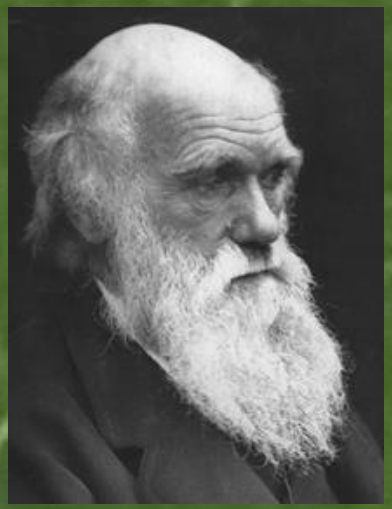
This course is about VARIATION: its causes, effects, and history.

For thousands of years, western thought had accepted the Platonic view that any object's ultimate reality was its *essence* or *ideal type*.

In biology, ***essentialism*** gave rise to the assumption that species are held together by their underlying, unchanging "types" or ideal forms.

On this view, individual variations are *departures* from the essence of a species; thus they are *imperfections* that make individuals less representative of the true nature of their species.



Darwin destroyed essentialism in biology and replaced it with a radical new idea: variationism.

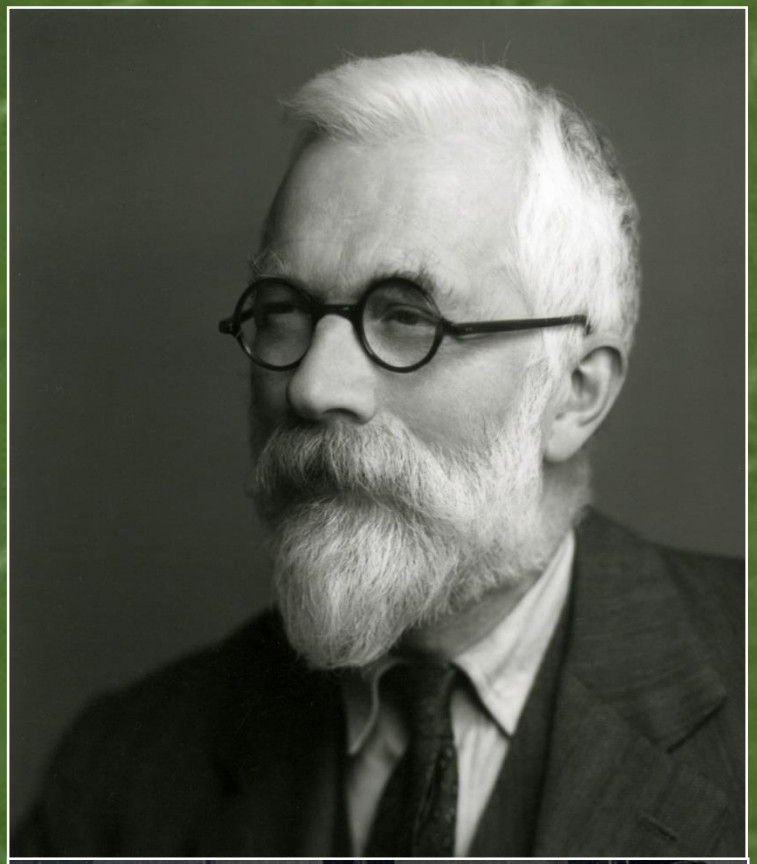
Variationism is the view that species are united only by *recent common ancestry*. Thus every individual is *equally* representative of the species. The *average* phenotype is just a statistical abstraction, *not* the reflection of some higher, more pure or more ultimate reality.

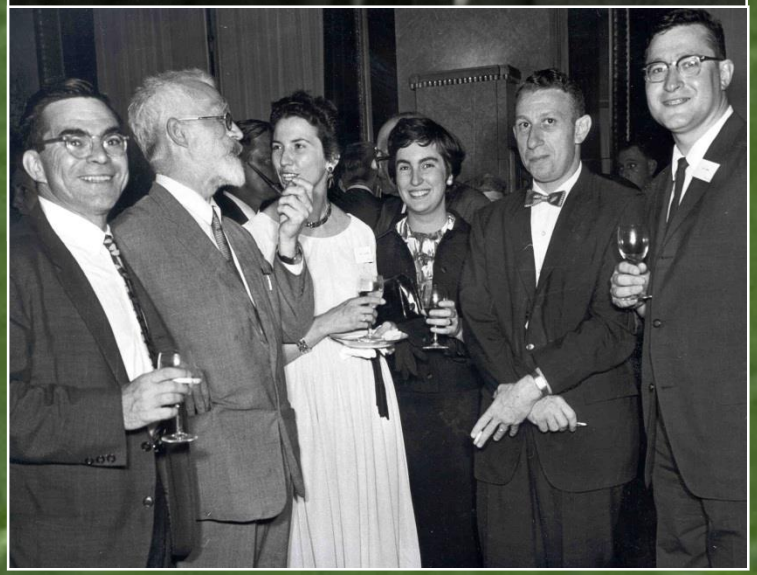
But Darwin was not a "radical" in the modern sense. He didn't set out to overturn essentialism. Facts and logic led him there.

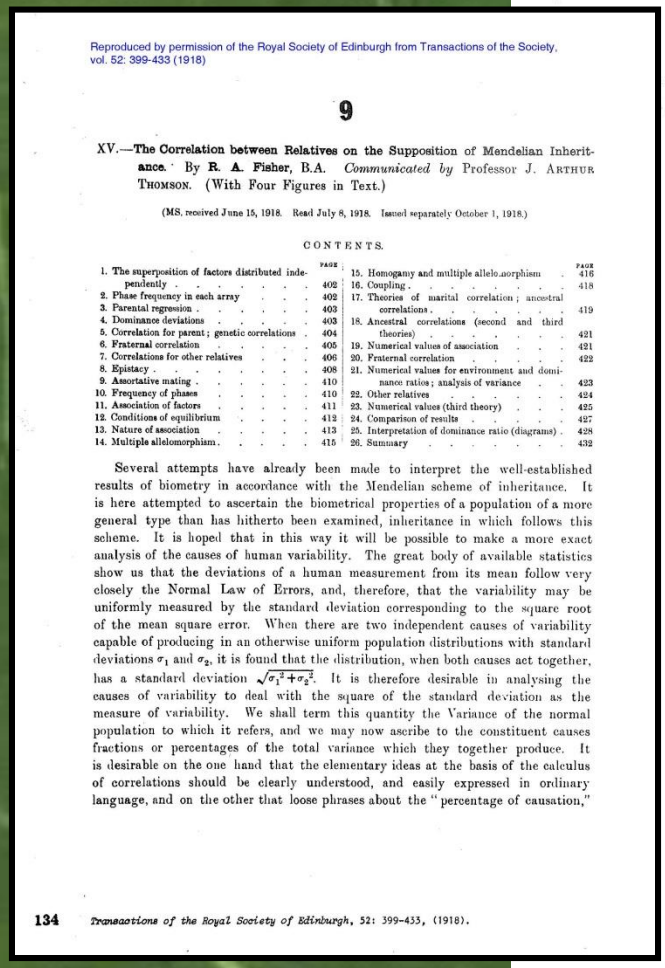
What causes variation, and why is some of it heritable (kids resemble parents?)

Darwin didn't know. Mendel's discovery of genes (1865) was rediscovered in 1900, and most of the early geneticists concluded that genes and Darwin were incompatible.

R.A. Fisher (1890-1962) invented the analysis of variance (ANOVA) in 1918 to show that Darwin's ideas about the inheritance of variation were consistent with Mendel's genetics. (He was 28!)







Reproduced by permission of the Royal Society of Edinburgh from Transactions of the Society, vol. 52: 399-433 (1918)

XV.—**The Correlation between Relatives on the Supposition** of Mendelian Inheritance. By **R. A. Fisher**, B.A. *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

CONTENTS.

Several attempts have already been made to interpret the well-established results of biometry in accordance with the Mendelian scheme of inheritance. It is here attempted to ascertain the biometrical properties of a population of a more general type than has hitherto been examined, inheritance in which follows this scheme. It is hoped that in this way it will be possible to make a more exact analysis of the causes of human variability. The great body of available statistics show us that the deviations of a human measurement from its mean follow very closely the Normal Law of Errors, and, therefore, that the variability may be uniformly measured by the standard deviation corresponding to the square root of the mean square error. When there are two independent causes of variability capable of producing in an otherwise uniform population distributions with standard deviations σ_1 and σ_2, it is found that the distribution, when both causes act together, has a standard deviation $\sqrt{\sigma_1^2+\sigma_2^2}$. It is therefore desirable in analysing the causes of variability to deal with the square of the standard deviation as the measure of variability. We shall term this quantity the Variance of the normal population to which it refers, and we may now ascribe to the constituent causes fractions or percentages of the total variance which they together produce. It is desirable on the one hand that the elementary ideas at the basis of the calculus of correlations should be clearly understood, and easily expressed in ordinary language, and on the other that loose phrases about the "percentage of causation,"

Transactions of the Royal Society of Edinburgh **52**, 399-433 (1918)

XV.—**The Correlation between Relatives on the Supposition** of Mendelian Inheritance. By **R. A. Fisher**, B.A. *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

CONTENTS.

Several attempts have already been made to interpret the well-established results of biometry in accordance with the Mendelian scheme of inheritance. It is here attempted to ascertain the biometrical properties of a population of a more general type than has hitherto been examined, inheritance in which follows this scheme. It is hoped that in this way it will be possible to make a more exact analysis of the causes of human variability. The great body of available statistics show us that the deviations of a human measurement from its mean follow very closely the Normal Law of Errors, and, therefore, that the variability may be uniformly measured by the standard deviation corresponding to the square root of the mean square error. When there are two independent causes of variability capable of producing in an otherwise uniform population distributions with standard deviations σ_1 and σ_2, it is found that the distribution, when both causes act together, has a standard deviation $\sqrt{\sigma_1^2+\sigma_2^2}$. It is therefore desirable in analysing the causes of variability to deal with the square of the standard deviation as the measure of variability. We shall term this quantity the Variance of the normal population to which it refers, and we may now ascribe to the constituent causes fractions or percentages of the total variance which they together produce. It is desirable on the one hand that the elementary ideas at the basis of the calculus

What causes variation, and why is some of it heritable (kids resemble parents?)

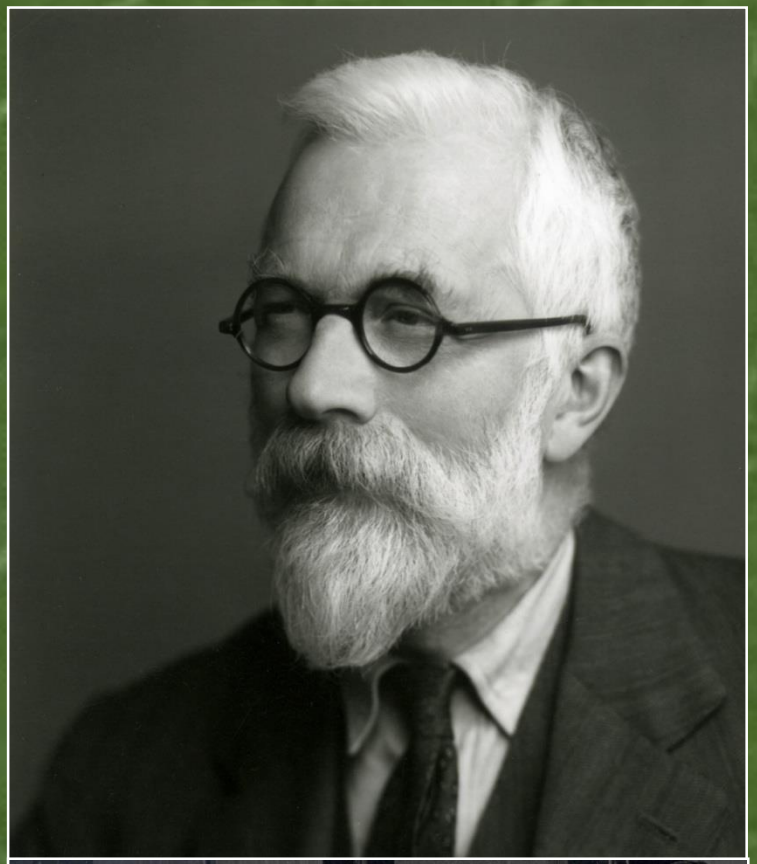
Darwin didn't know. Mendel's discovery of genes (1865) was rediscovered in 1900, and most of the early geneticists concluded that genes and Darwin were incompatible.

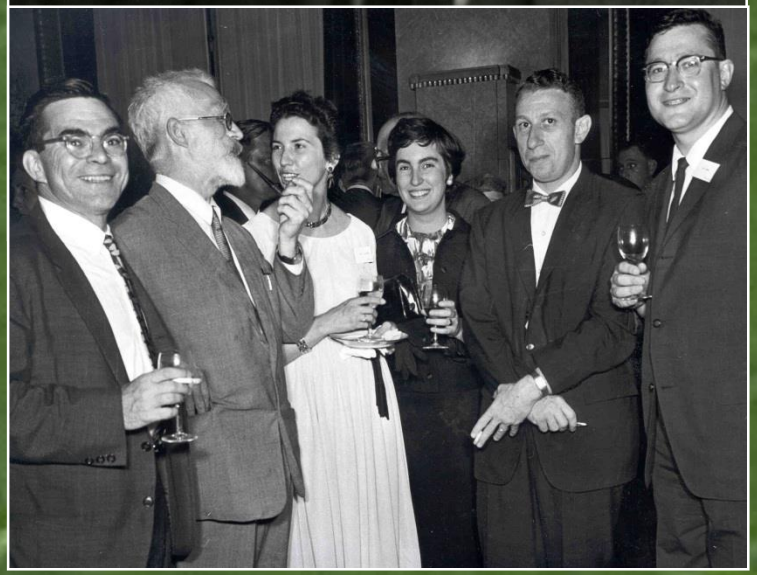
R.A. Fisher (1890-1962) invented the analysis of variance (ANOVA) in 1918 to show that Darwin's ideas about the inheritance of variation were consistent with Mendel's genetics. (He was 28!)

This is one of the deepest, most general and most transformative ideas in the history of human thought - and oddly, most invisible!

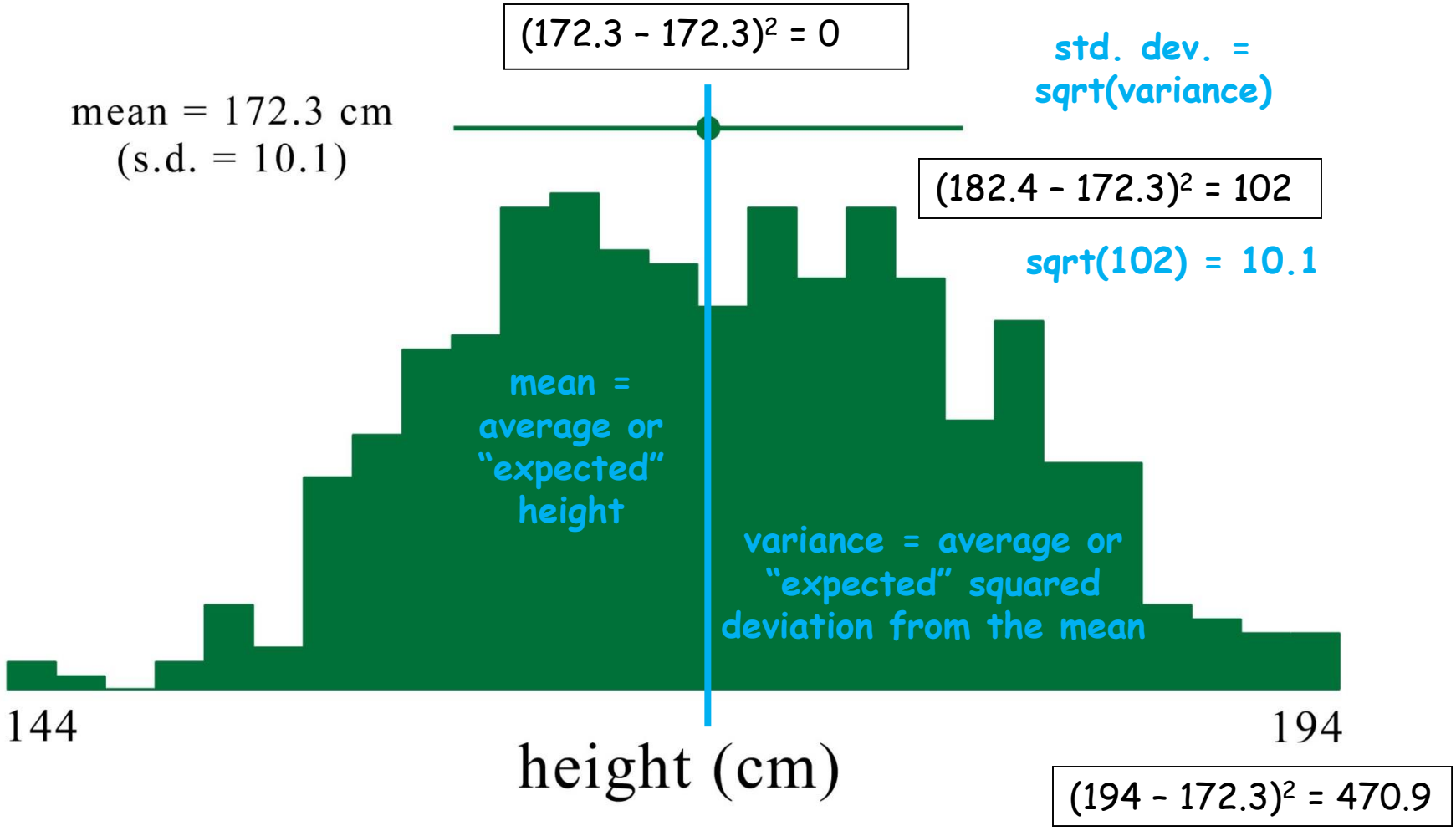
Fisher remains an obscure nerd celebrated by almost no one other than statisticians (right).

Meanwhile, ANOVA has become the foundation of statistical thinking and practice in industry, government and medicine as well as in science.

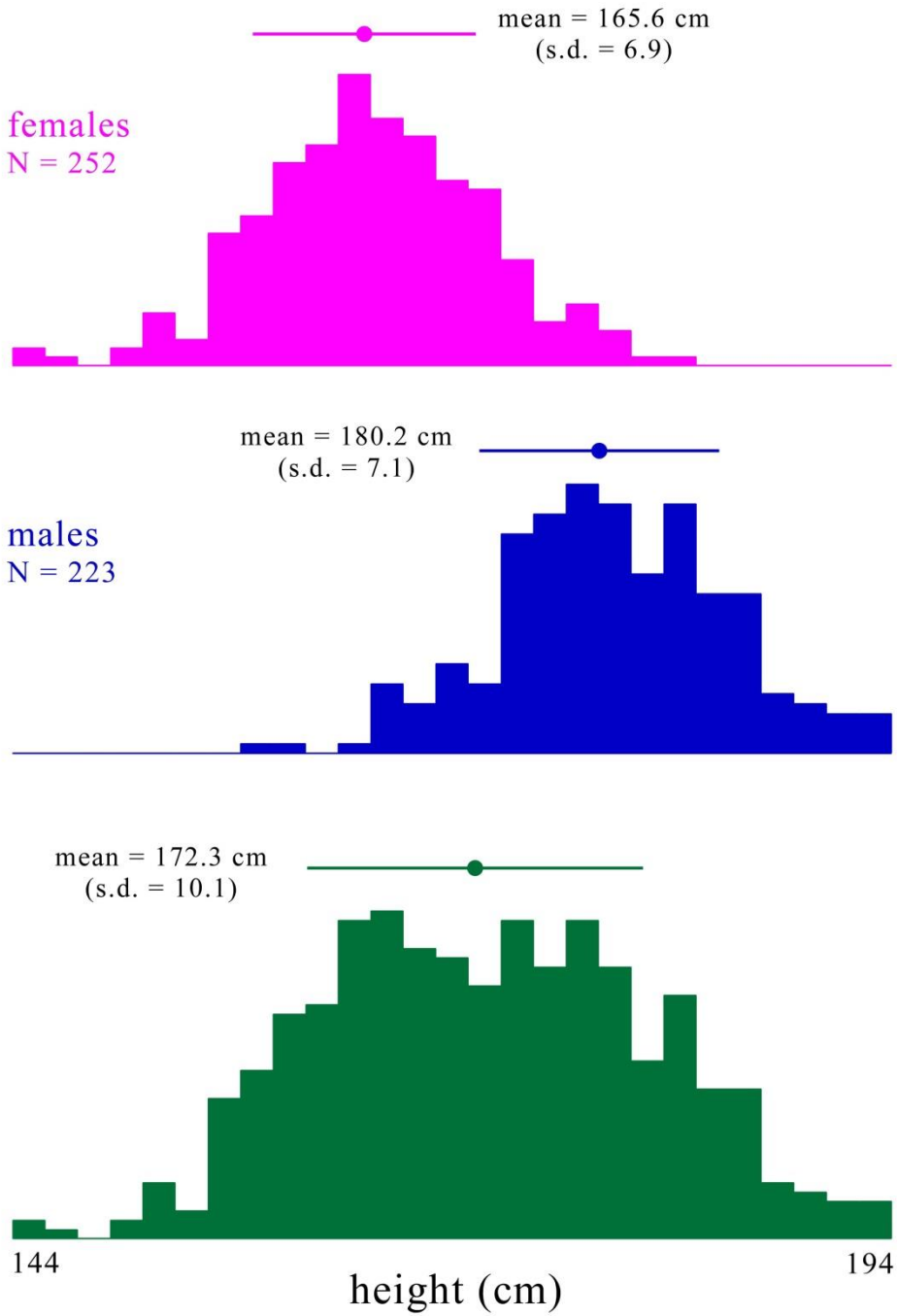




The classic quantitative trait, to which we will return …



Distribution of height for 475 individual adults participating in the Utah Genetic Reference Project (UGRP).



The female and male means differ greatly.

Each sex has a smaller variance (s.d.) than the population as a whole.

The central question answered by ANOVA:

What fraction of the total variance is "explained by" or "caused by" sex?

(In this case, by the difference between the female and male means?

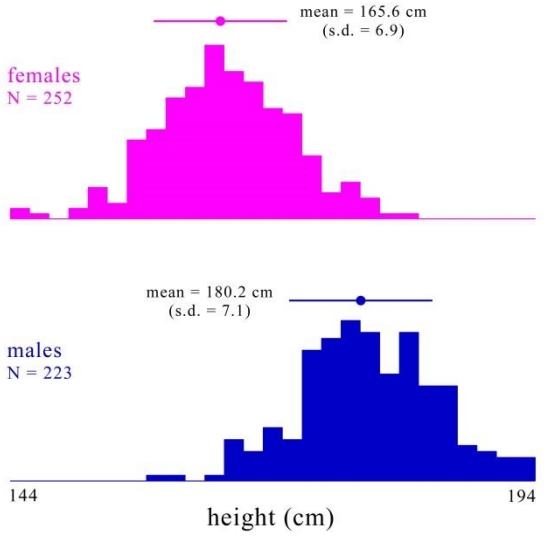
Heights of 252 women and 223 men in the Utah Genetic Reference Project

252 females : M = 165.6 V = 46.962
223 males   : M = 180.2 V = 49.782
475 total   : M = 172.5 V = 101.694

V(within) = 48.286 = 0.531*46.96 + 0.469*49.78
V(among) = 53.408 = 0.531*(165.6 - 172.5)^2 + 0.469*(180.2 - 172.5)^2
V(total) = 101.694

fraction "explained by sex" = 53.408/101.694 = 0.53

And notice that
V(within) + V(among)
is EXACTLY
V(total) !!



Distributions of height for individual adults participating in the Utah Genetic Reference Project (UGRP).

What about the effects of genes? (That is, of families?)

The subjects are all full siblings in 36 families with 1-12 sons and 1-12 daughters.

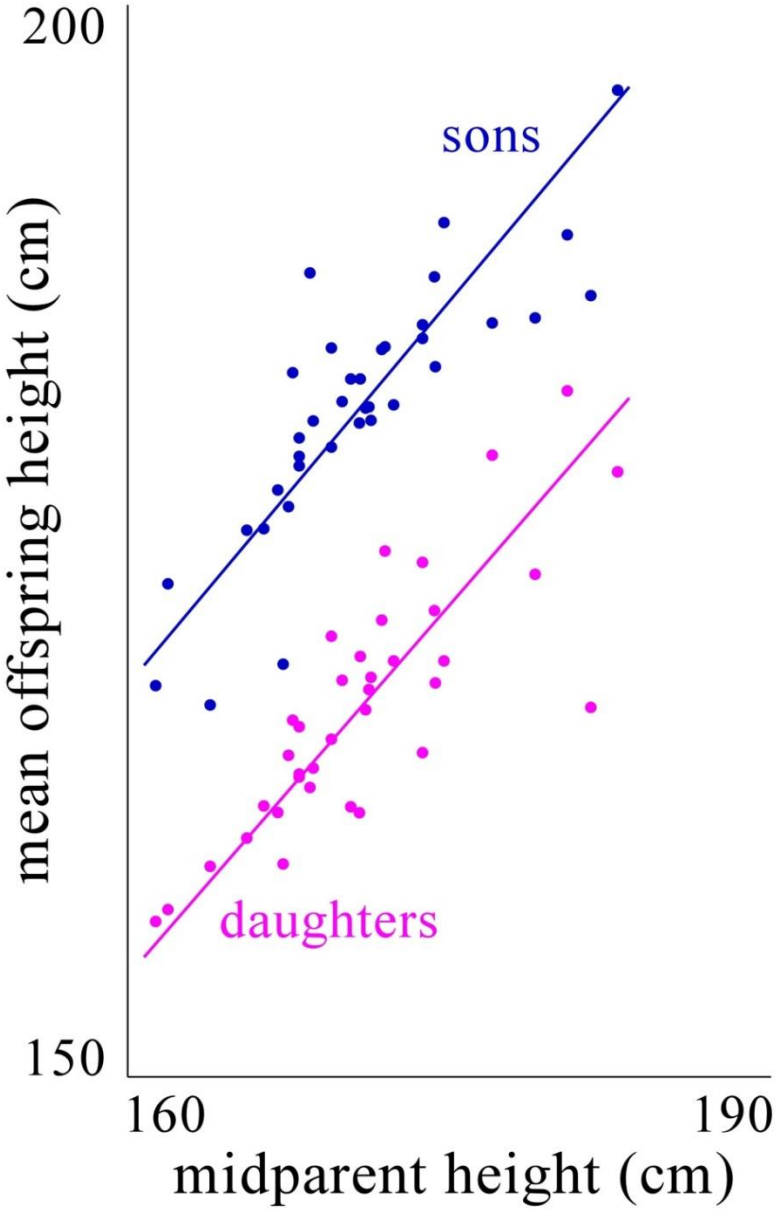
90% of the variance ***not*** explained by ***sex*** is explained by differences among the ***families***.

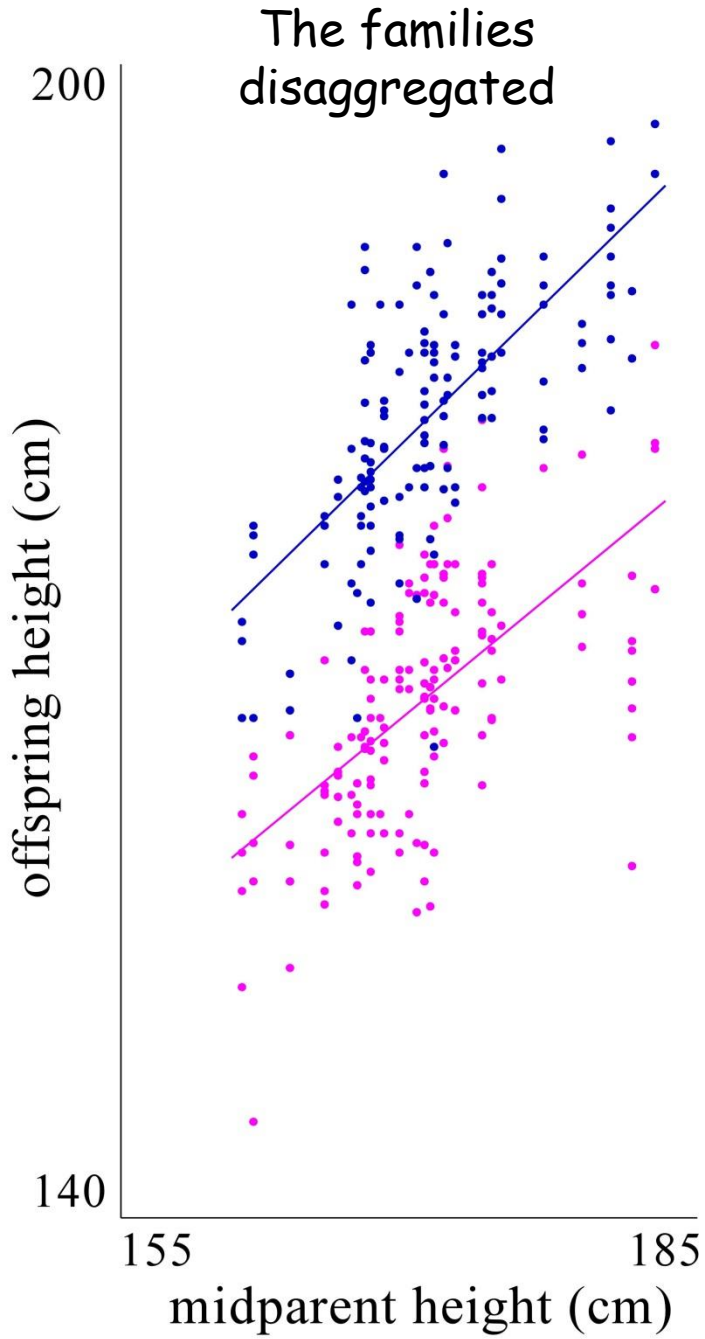
And 10% by ***effects of the environment*** (what remains after the effects of *sex* and of *genes* have been "removed" statistically).

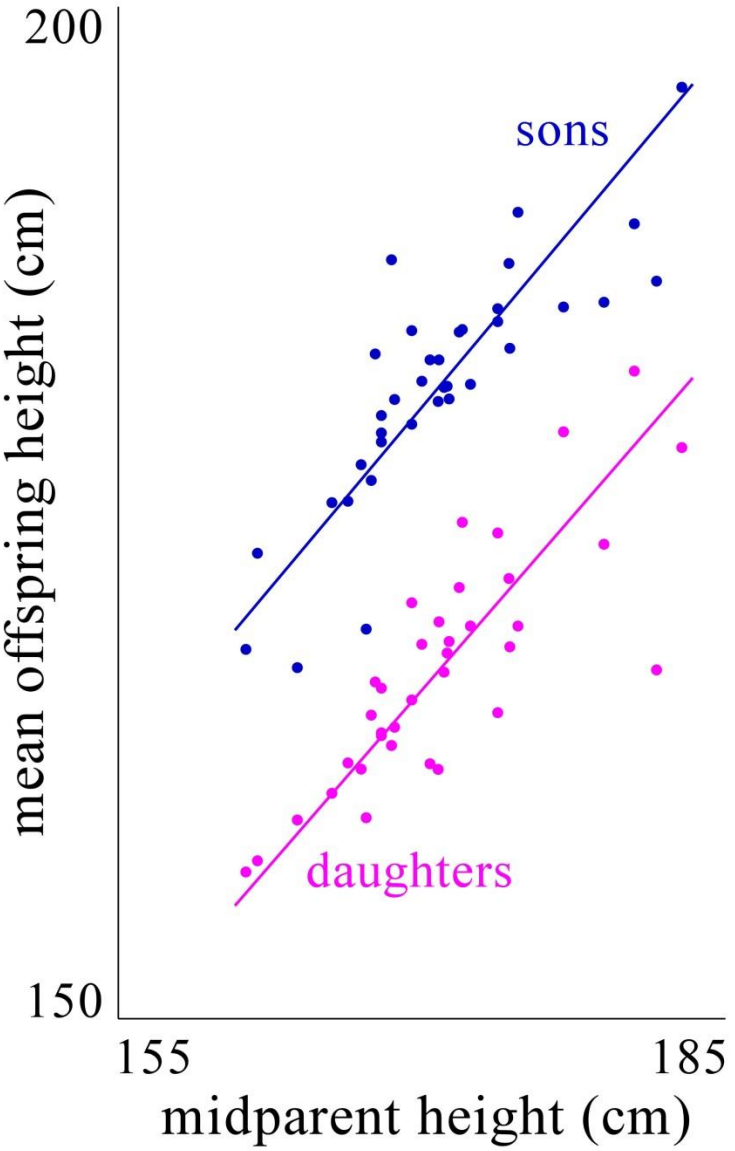
These people grew up in a very healthy and uniform environment (20th-century Utah).

In other times and places, the split tends to be 80/20 or even 70/30.

For other traits, in most species, it may be anywhere from 80/20 to 20/80.




The families
disaggregated
200
140
offspring height (cm)
155
185
midparent height (cm)


200
150
mean offspring height (cm)
sons
daughters
155
185
midparent height (cm)

Summary about phenotypic variation

Every quantitative phenotype you can think of ***varies***.

Often the distributions are roughly ***normal***.

If some of this variation is ***heritable***, then evolution by natural selection is ***inevitable*** (Darwin's world-changing insight).

Fisher invented **ANOVA** to show that ***darwinian*** evolution of quantitative traits is compatible with ***mendelian*** genetics.

(If ***many*** genetic loci make ***small***, independent contributions, and so does the environment).

The paper's **100th anniversary** was October 1, 2018.

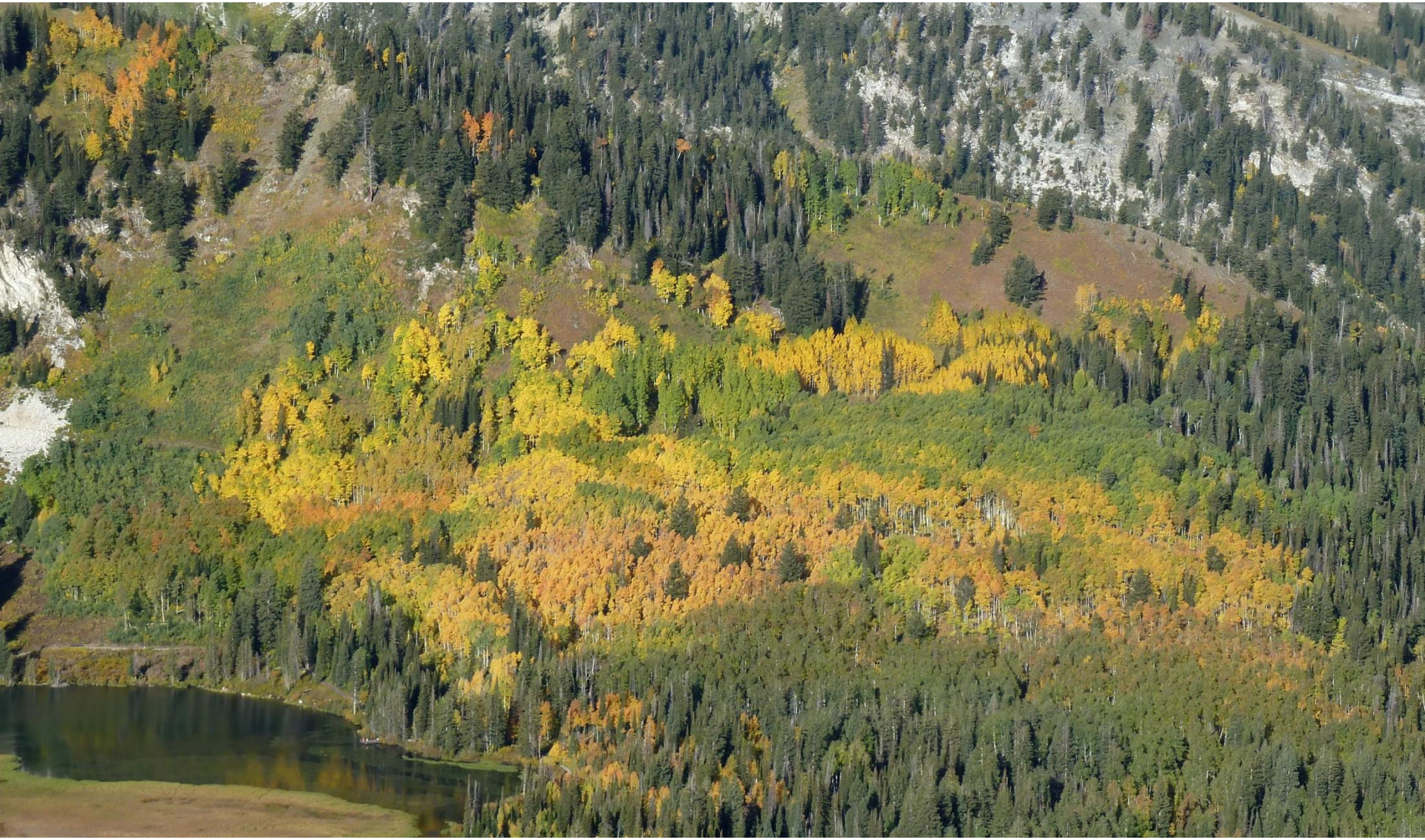
It + Darwin ***changed how we think*** about variation.

Now the variance can be "partitioned" into contributions associated with "factors" that "explain" the total.

Often (but not always) we can interpret the ***factors*** as ***causes***.

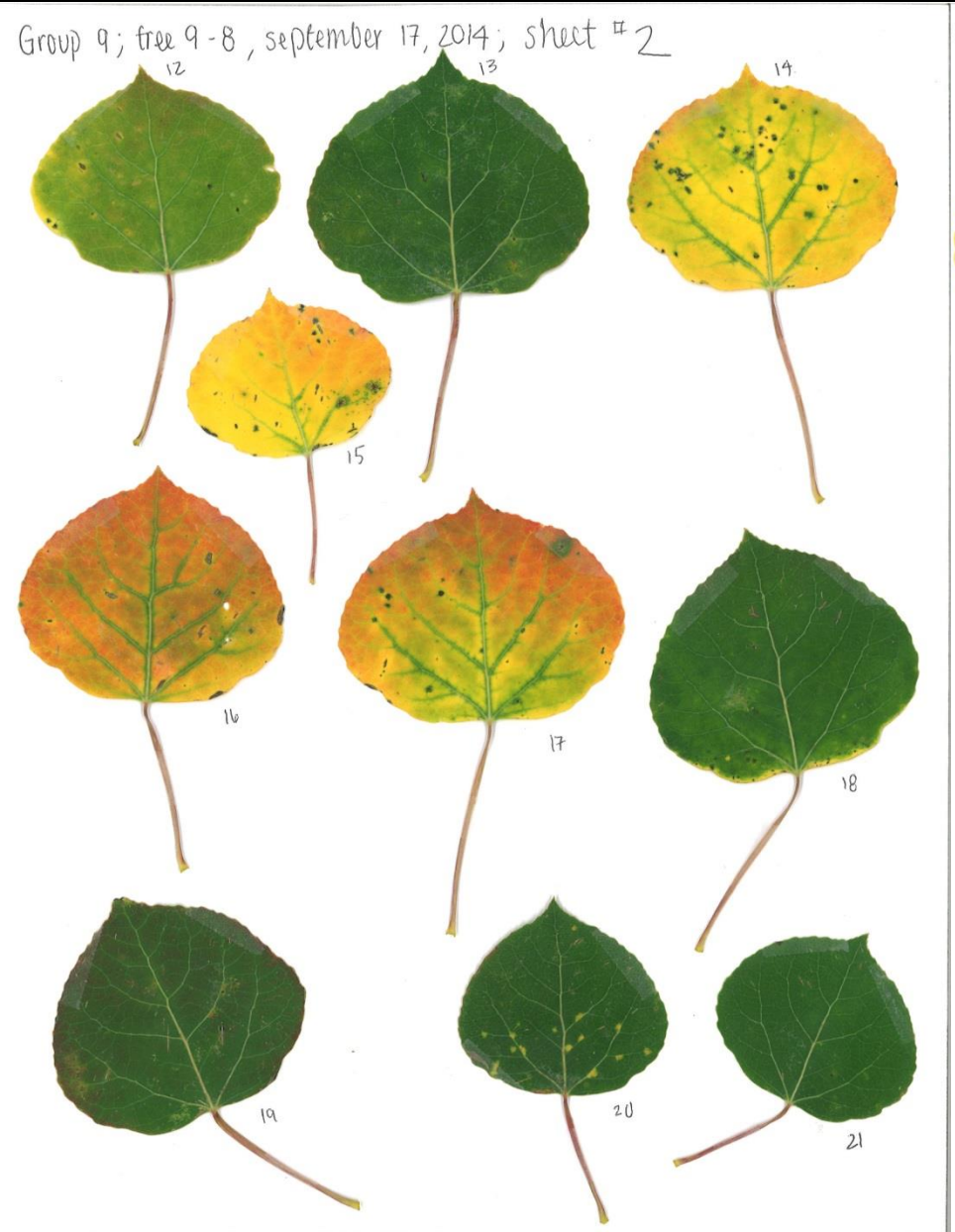
E.g., "***sex***" ~ 50%, "***genes***" ~ 40%, "***environment***" ~ 10% for height.

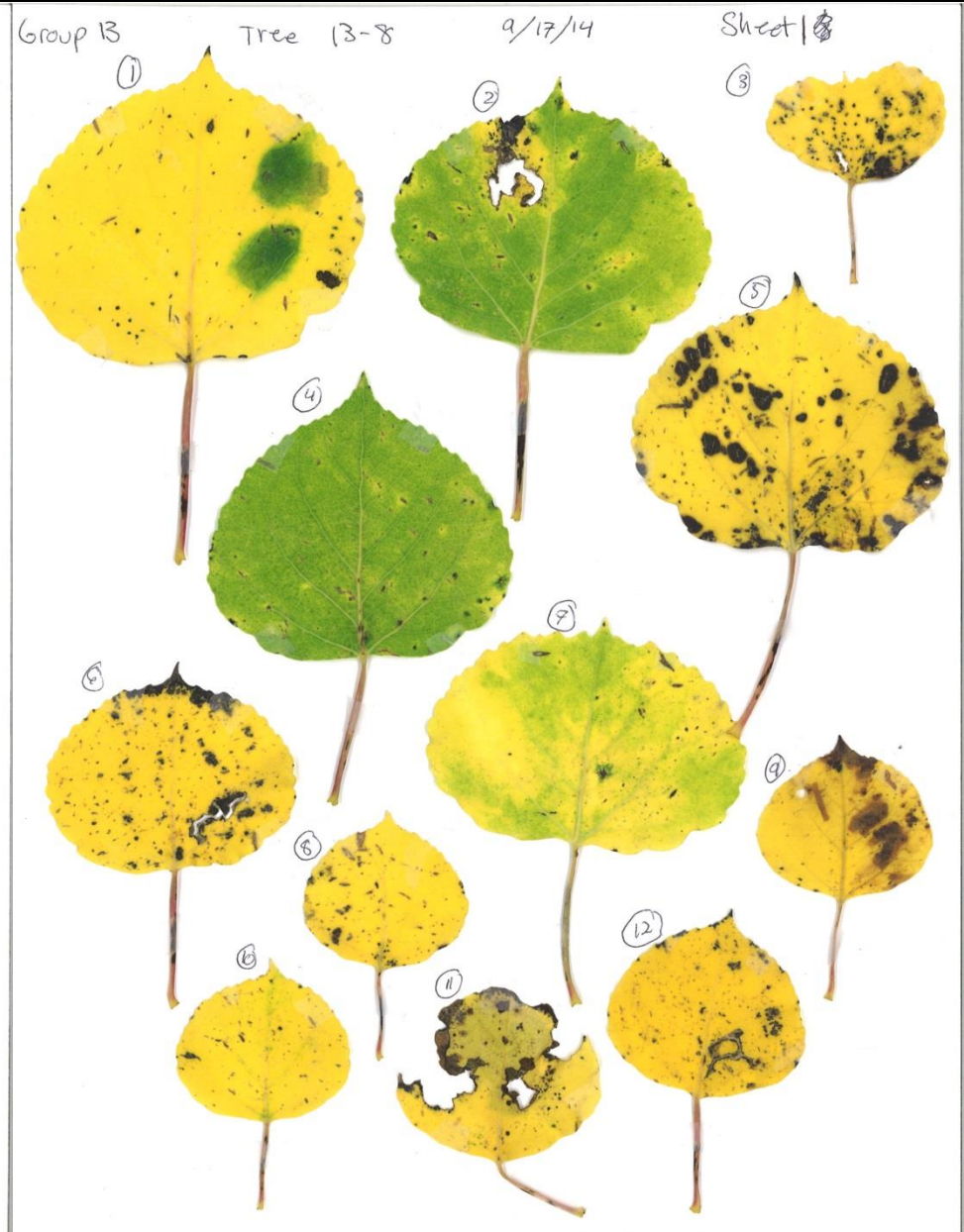
Why do some slides have this leaf as a background?

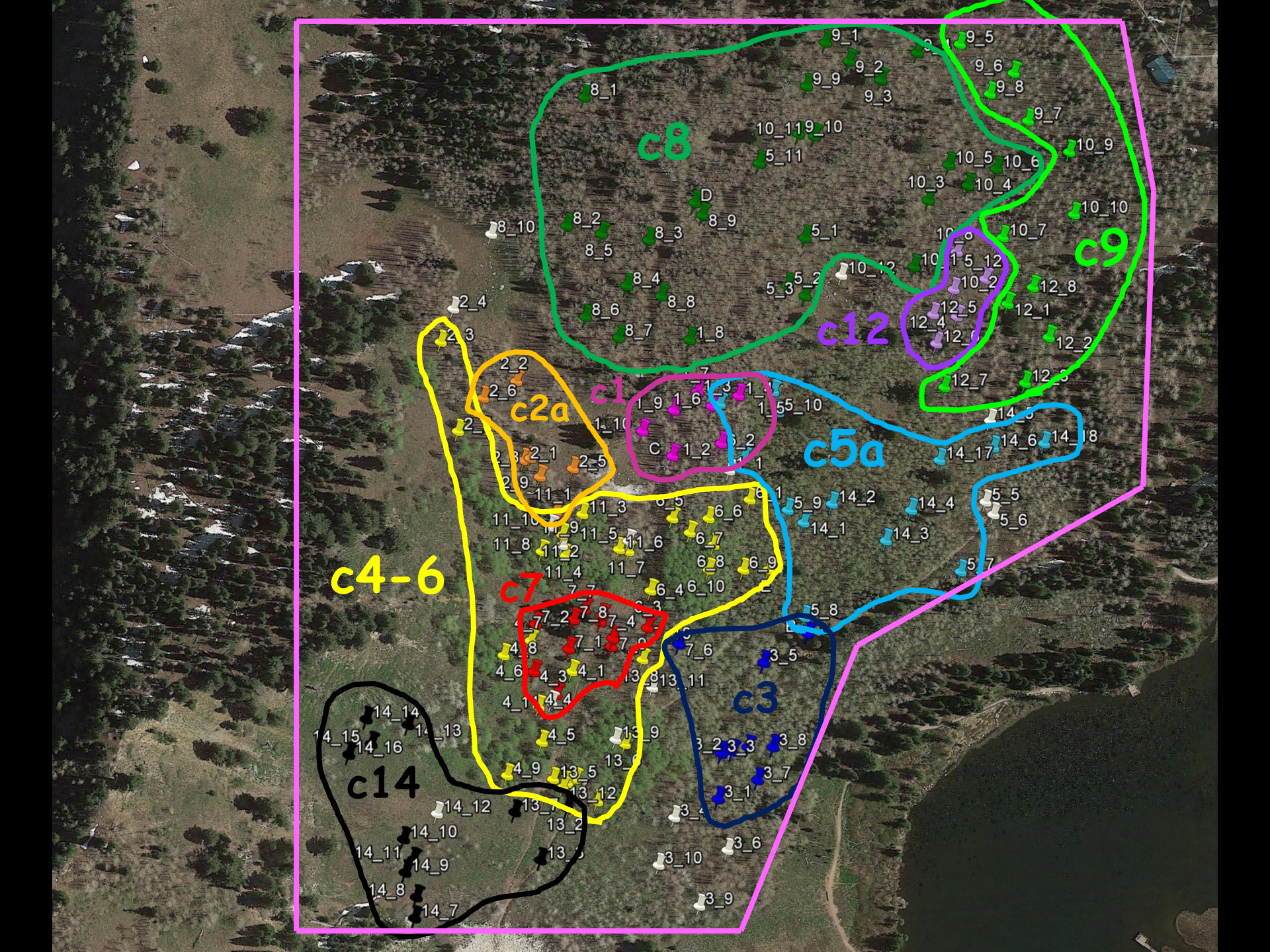


My then-other course studied aspens at Silver Lake!

Tree 9-8 belongs to Clone 9, and Tree 13-9 belongs to Clone 4-6






c8
c9
c12
c1
c2a
c5a
c4-6
c7
c3
c14