"Strong inference" requires that we reject alternative hypotheses.

The 1st step to statistical inference is to test the 'it's-just-random-noise' null model. If we can reject it, we can infer that some non-random causal process is at play, and further investigation is warranted, to determine what that process is.

There are many different kinds of ‘random’ processes
(ex, flip coin, deal cards, radioactive decay, light bulb dies, mutations, meiosis, ….) so basic and all-pervasive that scientists sometimes take its importance for granted. At some level every discovery in biology and medicine rests on it.

Suppose we have two coins \( \bullet \bullet \); & we want to test whether they are “fair” (long-term ave = \( p(h) = 0.5 \)) & independent of each other.

Experiment: flip both, record combination, repeat \( n \) times (\( n \) is big) …

If the NULL MODEL (two “fair”=\( p(h)=0.5 \) & independent coins) is correct then the pattern/data we would expect to find (in a big sample) is:

- \( p(1h\&1t) = 0.5 \times 0.5 = 0.25 \)
- \( p(hh) = 0.5 \times 0.5 = 0.25 \)
- \( p(ht or th) = 0.5 \times 0.5 = 0.25 \)

More generally, if \( p(h) = p \) and \( p(t) = q = 1-p \), Then expect \( p(hh) + p(ht or th) + p(tt) = 1 \)

\[ p^2 + 2pq + q^2 = 1 \]

Suppose we find (data): \( p(h) = 0.5 \), but rel. frequency \( hh = p(hh) = 0.1 \) ?

This is why you need statistics!
Suppose we have two copies (diploid) of a locus \( \bigcirc \bigcirc \); w/ 2 alleles: \( \{ W, R \} \), & over the pop of \( n \) individuals (2n loci): \( p(W) = 0.8 \) & \( p(R) = 0.2 = 1 - p(W) \).

If we ignore the source of the \( (W_W) \) & \( (W_R) \) (parent: mom or dad), then \( p((W_W) \& (W_R)) = ? \)

- \( p(W_W) = 0.8 \times 0.8 = 0.64 \)
- \( p(W_R) = 0.2 \times 0.8 = 0.16 \)

\( p(W_W) \) or \( p(W_R) \) = \( 0.16 + 0.16 = 0.32 \)

\( p((W_W) \& (W_R)) = 0.2 \times 0.2 = 0.04 \)

This particular NULL MODEL (for ‘random’ sex & survival) is called: The Hardy–Weinberg equilibrium.

Suppose we find (data): \( p(W_W) = 0.2 \), but \( p(W_R) = 0.37 \) ?

Work through the PKU ex, C&R pg 474

---

Microevolution is a generation-to-generation change in a population’s allele or genotype frequencies. The H-W model is a random NULL MODEL for a population that is not evolving. What processes would cause a pop to evolve (and cause us to reject the H-W null model)?

1. Drift – random sampling error in small pops & accumulation of neutral diffs.
   - bottleneck-founder effects important in island speciation & genetic disease.
   - accumulation of neutral mutations over time since common ancestry.
2. Gene flow – immigrants (or gametes) with dif allele freq make the p & q of the breeding adult pop different from the p & q of the ‘native born’ pop.
   - Consequences can look similar to selection.
3. Non-random mating – assortative: \( \text{p(hh)} > \text{p(h)} \times \text{p(h)} \), ‘too many’ homozygotes disassortative: \( \text{p(hh)} > 2 \times \text{p(h)} \times \text{p(t)} \) ‘too many’ heterozygotes.
   - hyp: variable selection for mutation repair?
5. Natural selection – differences in p & q of breeders relative to initial whole pop.
   - resulting in change in p & q of next generation.

Impact of selection, mutation rate and genetic drift on human genetic variation.
Sunyaev et al. Human Molecular Genetics, 2003, 12, 3325-3330.
... in protein coding regions ... (stabilizing) selection lowers genetic variability ... and this effect is relatively strong in comparison to genetic drift.

---

Evolutionary genetics of the major histocompatibility complex.
HEDRICK PW. 1994. AM. NAT. 143: 945-964.

The major histocompatibility complex (MHC) was first studied because of its importance in tissue transplantation and the immune system in humans.
... over 80 genes in the MHC, ... and, in some populations, there is an observed deficiency of homozygotes.
... some type of selection is operating in this region.

(The MHC is out of H-W equilibrium) ... this selection is related to the basic role of the MHC as part of the immune system acting to suppress attack by viruses, bacteria, and other parasites.
... research suggests that there is selection [5] at the MHC involved with maternal-fetal interactions [selective spontaneous abortion-miscarriage] and nonrandom mating [3] ...


Americans who are heterozygous for HLA-A, -B or -C antigens are protected against rapid progression to AIDS {HLA is MHC in humans} after infection with the human immunodeficiency virus-1 (HIV-1).

for a good overview, see:
The evolutionary ecology of the major histocompatibility complex

---

MHC-dependent mate preferences in humans.

One ... benefit of sexual reproduction could be that it allows animals (including humans) to react rapidly to a continuously changing environmental selection pressure such as coevolving parasites.

This ... would be most efficient if the females were able to provide their progeny with certain allele combinations for loci which may be crucial in the parasite-host arms race, for example the MHC (major histocompatibility complex).

Female and male students were typed for their HLA-A, -B and -DR.
Each male student wore a T-shirt for two consecutive nights.
... each female student was asked to rate the odours of six T-shirts.

They scored male body odours as more pleasant when they differed from the men in their MHC than when they were more similar.

see http://www.pbs.org/wgbh/evolution/library/01/6/l_016_08.html

Wedekind & Furi (1997) had men sniff women’s T-shirts & found same pattern of preference for dissimilar MHC.

\{ Does this result in disassortative marriage? \}
The hunt for the genes that helped humans adapt to new climates, diseases, and diets is exposing how evolution works.

Rasmus Nielsen was analyzing the frequency of different mutations in the genomes of Tibetans living at high altitude, searching for adaptations that allow them to thrive in thin air.

... two stood apart ... they existed in almost all Tibetan highlanders but not in their close relatives, the Han Chinese.

... this was a radical example of rapid evolution, with strong (local) natural selection acting on a single gene (and limited migration) ...

... the EPAS1 gene that regulates oxygen sensing in humans. One of the mutations ... had spread to 90% of all Tibetans in just 4000 years ... the most rapid and strongest example of selection known in modern humans. See: http://www.sciencemag.org.proxy1.cl.msu.edu/cgi/content/full/329/5987/40

All living humans are remarkably similar genetically because we all descended from a small founder population that arose in Africa about 200,000 years ago. As these modern humans moved ... out of Africa in the past 80,000 years or so, they evolved genetic differences that helped them adapt to new climates, digest novel foods, and fight off new illnesses and parasites. { on the tree of life: deep, fundamental unity; superficial diversity }
When a single genetic mutation first let ancient Europeans drink milk, it set the stage for a continental upheaval.

During the most recent ice age, milk was essentially a toxin to adults because unlike children they could not produce the lactase enzyme required to break down lactose, the main sugar in milk.

As farming replaced hunting & gathering in the Middle East 11,000 years ago, cattle herders learned how to reduce lactose in dairy products by fermenting milk to make cheese or yogurt.

Several thousand years later, a genetic mutation spread through Europe that gave adults the ability to produce lactase & drink milk throughout lives.

This two-step milk revolution (1st “nurture” then “nature”) may have been a prime factor in allowing bands of farmers and herders from the south to sweep through Europe and displace the hunter-gatherer cultures that had lived there for millennia. (note: immigrants displaced locals) continued …

Most people who retain the ability to digest milk as adults can trace their ancestry to Europe… where the trait seems to be linked to a single nucleotide in which the DNA base cytosine changed to thymine … There are other pockets of lactase persistence in West Africa, the Middle East and south Asia that seem to be linked to separate mutations. (indep. convergence)

The single-nucleotide switch in Europe happened relatively recently. Researchers estimated the time by looking at genetic variation in modern populations and running computer simulations of how the mutation might have spread. … the lactase persistence LP allele, emerged about 7,500 years ago in Hungary. Once the LP allele appeared, it offered a major selective advantage. In a 2004 study, researchers … called that degree of selection “among the strongest yet seen for any gene in the genome”.

As agriculture spread from Anatolia to northern Europe over roughly two millennia… domesticated cattle at Neolithic sites in Europe were most closely related to cows from the Middle East, rather than indigenous wild aurochs. … incoming herders brought their cattle with them, rather than domesticating locally … (immigrants replaced locals)

A similar story is emerging from studies of ancient human DNA … which suggest that Neolithic farmers in Europe were not descended from the hunter-gatherers who lived there before. (immigrants replaced locals that’s how allele spread)

The milk revolution continued …

Most people who retain the ability to digest milk as adults can trace their ancestry to Europe… where the trait seems to be linked to a single nucleotide in which the DNA base cytosine changed to thymine … There are other pockets of lactase persistence in West Africa, the Middle East and south Asia that seem to be linked to separate mutations. (indep. convergence)

The single-nucleotide switch in Europe happened relatively recently. Researchers estimated the time by looking at genetic variation in modern populations and running computer simulations of how the mutation might have spread. … the lactase persistence LP allele, emerged about 7,500 years ago in Hungary. Once the LP allele appeared, it offered a major selective advantage. In a 2004 study, researchers … called that degree of selection “among the strongest yet seen for any gene in the genome”.

As agriculture spread from Anatolia to northern Europe over roughly two millennia, … domesticated cattle at Neolithic sites in Europe were most closely related to cows from the Middle East, rather than indigenous wild aurochs. … incoming herders brought their cattle with them, rather than domesticating locally … (immigrants replaced locals)

A similar story is emerging from studies of ancient human DNA … which suggest that Neolithic farmers in Europe were not descended from the hunter-gatherers who lived there before. (immigrants replaced locals that’s how allele spread)

Ancient urbanization predicts genetic resistance to tuberculosis

If the transition to urban living does result in an increase in disease-based mortality, then we might expect to see evidence of increased disease resistance in longer-term urbanized populations, as the result of natural selection.

To test this, we determined the frequency of an allele (SLC11A1 1729 + 55del4) associated with natural resistance to intracellular pathogens … tuberculosis. We found a highly significantly correlation with duration of urban settlement – populations with a long history of living in towns are better adapted to resisting these infections. This correlation remains strong when we correct for autocorrelation in allele frequencies due to shared population history.

Our results therefore support the interpretation that infectious disease loads became an increasingly important cause of human mortality after the advent of urbanization …

Genes and phenotypes vary within & between populations. Variation is the raw material (and product) of evolution.
news@nature.com
Published online: 10 February 2016; 1 doi:10.1038/news060214-17

Gene map opens up uncharted territory
News and Views - Evolutionary biology: Geography and skin colour

The most obvious aspect of human geographical variability is skin colour. Most people would say that skin colour is darker towards the Equator to give more protection against tropical sunlight. But the correlation of skin colour with latitude is riddled with exceptions ...

Jablonski & Chaplin have now brought order to this confused field, with quantitative measurements of skin colour and sunlight. Ultraviolet radiation (UVR) at the Earth’s surface does decrease with latitude, but the correlation is imperfect: UVR also increases with altitude & atmospheric water vapour. Variation in UVR proves to be the strongest predictor of skin reflectance, explaining 77% (Northern Hemisphere) or 70% (Southern Hemisphere) of its variation.

(is skin color closely correlated w/ ancestry? –does it predict genes at other loci?)
Implications of correlations between skin color and genetic ancestry for biomedical research

Skin pigmentation is a central element of "race". We studied the relationship between pigmentation and ancestry in (self-identified) populations of mixed ancestry with a wide range of pigmentation and ancestral proportions (○ African Americans from Washington, DC; African Caribbeans living in England; ● Puerto Ricans from New York; ▲ indigenous Mexicans from Guerrero; and Hispanics from San Luis Valley). The strength of the relationship between skin color and ancestry was quite variable …
These results … emphasize the need to be cautious when using pigmentation as a proxy of ancestry … we should not extrapolate findings based on superficial traits, strongly subject to selection, to the rest of the genome.

The success of pharmacogenomics in moving genetic association studies from bench to bedside: study design and implementation of precision medicine in the post-GWAS era (GWAS = Genome Wide Association Studies: correlates genetic markers, like SNPs, w/ phenotypes)
Marilyn D. Ritchie 2012, Vol 131, Pages 1615-1626

Personalized medicine, or more recently coined precision medicine, has advanced as one of the predominant strategic initiatives and goals …
... it is hypothesized that variability in drug response is due to underlying individual variation in genetic architecture.
The goal … treat patients with the correct dose of the appropriate medication based on their individual demographic (ethnicity) and genomic makeup.

Pharmacogenetics is the study of a single genetic variant (allele) with a drug response phenotype, such as treatment responders and non-responders (i.e. assessment of drug efficacy) or a serious adverse side effect (i.e. drug toxicity).
As molecular technologies to assay the entire genome have developed and genome-wide association studies (GWAS) emerged, so did pharmacogenomics (surveying the entire genome for associations w/ drug response phenotypes).

{lots of markers, like SNPs, allow us to calibrate genetic similarity between individuals; then if a drug works well or badly on others with similar genotypes, predict … }

Evolutionary Genetics of Coronary Heart Disease
Ding KY, Kullo IJ 2009. CIRCULATION 119:459-560

Susceptibility to common diseases such as coronary heart disease (CHD) may in part reflect historical or evolutionary legacies …
The evolutionary history of the human species may provide valuable insights into the origin of common diseases beyond what is possible by investigating only the most immediate or "proximal" causes of disease. … may help answer why … CHD has assumed epidemic proportions, and why …
beyond what is possible by investigating only the most immediate or "proximal" causes of disease. … may help answer why … CHD has assumed epidemic proportions, and why substantial differences in susceptibility to CHD are present between ethnic groups.

Selection will tend to maintain genes that increase reproductive success (early) even if these increase disease susceptibility in older age. (late)

For example, genetically determined high cytokine (immune sys) response levels may be associated with adverse cardiovascular outcomes in older individuals but may increase reproductive success in young age by conferring resistance to fatal infectious diseases ...
{more on "adaptive" theories for aging later}

The universal triplet genetic code (text Fig 17.5) has redundancies at 3rd base → code same AA; these are neutral: no effect on protein function.

Synonymous base substitutions (UUU → UUC) are neutral: no effect on protein function.
Non-synonymous base substitutions (UUU → UUA) change coded AA sequence & protein structure & maybe function;
Evolutionary Genetics of Coronary Heart Disease
Ding KY, Kullo IJ 2009. CIRCULATION 119:459-

Broadly, 2 types of selective forces have shaped the evolution of species:
- **purifying (stabilizing) selection**, favors conservation of existing phenotypes:
  - nonsynonymous (~functional) mutations will be selected against and eliminated (extremes are bad)
- **positive (directional) selection**, promotes new phenotypes:
  - manifests as rapid divergence of functional ~nonsynonymous sites between species
  - and a reduction in variation within species. (one extreme better)
  - (note text Fig 23.13 also includes disruptive selection)

Over a prolonged period, positive selection can increase the fixation rate of beneficial function-altering mutations (eg, substitutions that change amino acids [ie, nonsynonymous substitutions]).

A simple way of detecting selection from comparative genomic data is to calculate the ratio between the rate of nonsynonymous substitutions and the rate of synonymous substitutions (ie, dN/dS).

This ratio provides a means of detecting selective pressures:
- dN/dS = 1 for no selection, (random drift)
- dN/dS < 1 for purifying selection, and (stable phenotype)
- dN/dS > 1 for positive selection (changing phenotype)

Two major challenges in studying the genetic basis and evolutionary genetics of CHD are its phenotypic complexity and the presence of multiple causal factors.

- we enumerate genes in causal pathways of atherosclerosis, including blood pressure regulation, lipoprotein and glucose metabolism, coagulation, and inflammation, that may be subject to various degrees of selective pressures resulting from climatic and dietary changes and host response to pathogens.

A recent study suggested that evolutionary pressures resulting from climatic changes may have played an important role in shaping variation in genes in metabolic pathways.

Differential Susceptibility to Hypertension
Due to Selection during the Out-of-Africa Expansion.
JH Young et al. PLoS Genetics Vol. 1, No. 6 Full-text

The genetic basis of blood pressure variation is largely unknown but is likely to involve genes that influence renal salt handling & arterial vessel tone. Here we argue that differential susceptibility to hypertension is due to differential exposure to selection pressures during the out-of-Africa expansion.

The most important selection pressure was climate, which produced a latitudinal cline in heat adaptation and, therefore, hypertension susceptibility. (evol. adaptation to hot climate involves salt and water retention; ↑risk of hypertension)

Are we still evolving? The simple answer is yes.
A comparison of the sequences of over 11,000 genes reveals more than 1,139 that show evidence of either positive or weak negative selection.

Natural selection on protein-coding genes in the human genome

A team at Celera Genomics sequenced 20,362 loci in 20 European Americans, 19 African Americans and one male chimpanzee...

- comparing ... synonymous (S) versus non-synonymous (N) sites...
- (the universal triplet genetic code has redundancies at 3rd base → code same AA; non-synonymous DNA base substitutions change coded AA sequence & protein; synonymous subs have no effect on protein: serve as a null, background mutation rate)
- the parameter is <1 if a gene shows a paucity of amino acid divergence and >1 if a gene has an excess of amino acid divergence relative to the genomic average for synonymous sites.

304 (9.0%) loci showed evidence of rapid amino acid evolution.
813 (13.5%) show a paucity of amino acid differences between humans & chimps indicating (stabilizing) selection.

The results of analysis of human polymorphism at this scale may help to guide our thinking about human evolution...

More genes underwent positive selection in chimpanzee evolution than in human evolution !!!
MA Bakewell et al. 2007 PNAS 104:7489–7494