Evolutionary Anthropology 18:125–130 (2009)

Pieces of Eight!

Like centuries of pursuit of the Spanish treasure fleet, gene hunters seek elusive treasures in the human genome. But are they there?

KENNETH M. WEISS

In 1565, a Spanish galleon crossed the Pacific Ocean. The galleons were the first Hispanic high-riders, topheavy and cumbersome (Fig. 1A). This voyage inaugurated 250 years of annual shipments of spices and silk collected at Manila. The goods were carried east in a scurvy-ridden nonstop crossing of 8,000 miles of open ocean to Acapulco, where they were sold for silver pieces of eight, stamped coins worth 8 *reales* (Fig. 2) that became accepted tokens of exchange around the world.[1] The silver treasure carried by these galleons was mined and minted in Peru and Mexico, then called New Spain. The proceeds from the year's sales were taken on a return voyage to Manila (Fig. 1B). On the Atlantic side of the New World, whole convoys of galleons and their accompaniment sailed regularly from the Caribbean, twice yearly, when war, pirates, wealth, and logistics would allow, to haul the glittering bounty of New Spain to the mother country.[2]

At that time, ships sailing the eastbound Pacific route embarked on the world's most dangerous sea voyage, through unpredictable weather in unstably overloaded galleons run by crews with deficient sailing skill. The length of the voyage could vary greatly, and depended heavily on luck. Ships with skeleton crews—that is, crews that were literally skeletons because of starvation and disease—were occasionally found. Some of them never made port. Even after months of sailing, and with the destination at hand, a galleon might be attacked by predatory buccaneers waiting to board these ships and heist their treasure.[1–3]

On both sides of the New World, many galleons foundered on rocks or sank due to weather, poor navigation, or fights with buccaneers.[3] The Spanish quickly salvaged much of the treasure from these ships, but enough went to the bottom to stimulate vigorous hunts for sunken treasure even today.

Spain was impatient to use its Indies treasure for short-term gain. A land power in Europe, she developed neither first-rate maritime nor mercantile traditions. The Spanish silver largely flowed immediately through Spain, to repay Italian or Dutch lenders for debts recklessly accumulated in waging endless wars for European domination. For nearly three centuries, Spain, rather than building a commercial base at home or in its colonies, depended on milking the colonies for silver. Meanwhile, the Spanish treasure fleets were good for many businesses at the time: those that provided equipment and supplies to mines and built the ships to transport or pursue the treasure did very well.

In the end, an exhausted, backward, and impoverished Spain gave up both its colonies and its European power. Such was the price of greed, that a hunger for immediate payoff replaced a long-term plan. A treasure that would support the society was elusive and never realized. The English, in contrast, invested in settling their colonies as sources of renewable goods and trade rather than relying on shiny booty, and that led to long-term prosperity.

STALKING GENETIC TREASURE

We're currently experiencing a somewhat similar pursuit in genetics. It is the pursuit of treasures that, like those pursued by the Spanish and buccaneers, many are convinced lie hidden in our genomes (Fig. 3). There is aggressive hunting for this genetic wealth, especially as it relates to disease. The media and journals alike are filled with announcements of the discovery of the genes causally responsible for almost any trait you could name. The most sought-after prizes are genes "for" common complex diseases rather than those that, like many pediatric diseases, are due to the effects of a single gene but generally are of low frequency in the population. Common diseases present potentially lucrative therapeutic markets for the pharmaceutical or gene-therapy industries.

The basic approach, which has become pandemic in the research community, is called genome-wide association studies, or GWAS (pronounced GEE-waz, which is enticingly close to GEE-wiz). GWAS are a complementary approach to classical family studies, which have gone out of fashion for various logistical and technical reasons. But like family studies, GWAS start with a set of genetic markers; that is, highly vari-

Ken Weiss is Evan Pugh Professor of Anthropology and Genetics at Penn State University. Email: kenweiss@psu.edu

DOI 10.1002/evan.20218
Published online in Wiley InterScience (www.interscience.wiley.com).

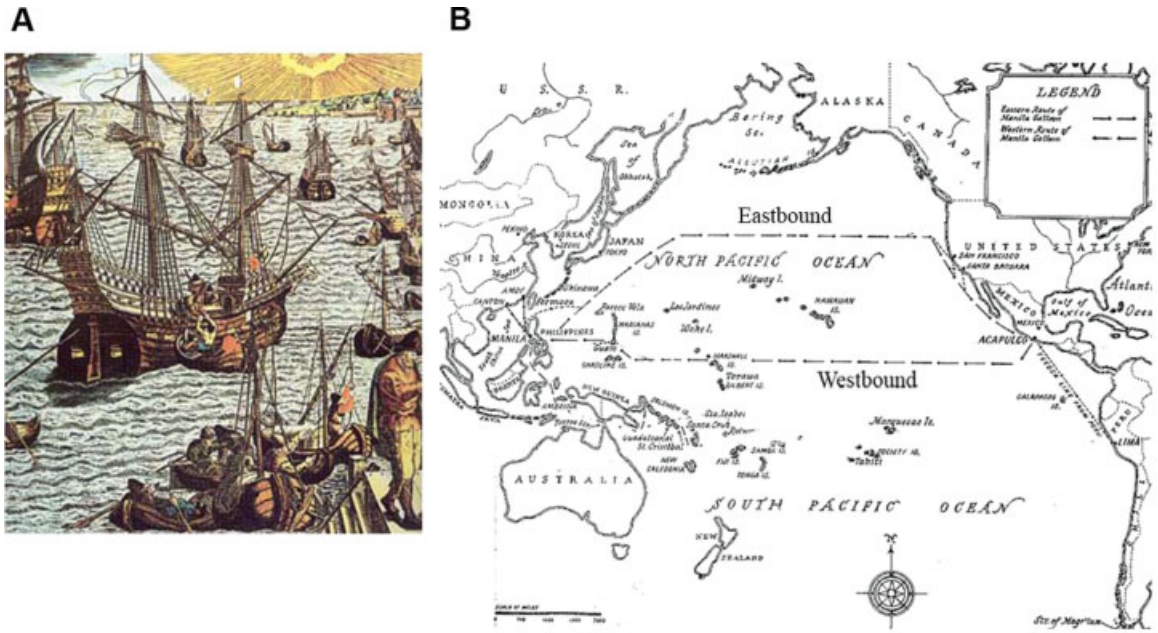

Figure 1. Spanish treasure fleet. A. Loading up for the long voyage (public domain). B. Typical Pacific routes; eastbound: top, westbound bottom. Modified from[1] Shurz. (Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.)

able spots spaced more or less evenly across our genomes (all our chromosomes). The markers are chosen for their variability and genomic location, not their function. (Usually, they have no known function.) With hundreds of thousands of markers that can easily be genotyped using current technology, every place along every chromosome in our genome is likely to be rather close to some of these marker sites.

Samples of individuals, such as cases and controls, are collected and each individual is genotyped for the same set of variable markers. The results are searched for in areas of the genome in which the marker genotypes correlate with the individual's phenotypes (like disease or other traits that can be measured). The idea is that if individuals with similar phenotypes share similar genotypes in some chromosomal region, they may share *untyped* functional DNA sequence variation in that region as well. Those parts of the genome become "candidates" that are searched in detail to find the causal variation.

Figure 2. Mexican pieces of 8 (*pesos*), 1806 and other years. Courtesy Nancy Buchanan. (Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.)

About 100 articles appear every *week* in the major journals to review, and often to hype, the GWAS mania and its findings. "Hype" is a fair description because, regardless of how good the findings actually are, the articles are often written by authors who have strong vested interests in the furtherance of the research programs. (For different views, see Altshuler, Daly, and Lander[4]; Weiss[5]; Manolio, Brooks, and Collins[6]; Bodmer and Bonilla[7]; and The Wellcome Trust Case).[8] Conflicts of interest do not prove incorrectness (though the track record is not good), but there is a bit of *caveat emptor* in that one must at least read these promotions circumspectly.

At the same time that we see an ocean of dazzling reports of gene discovery, many have begun openly acknowledging what's been quietly known for quite a while by those who have been paying attention: Despite some clear successes, GWAS as a whole are not yielding their promised bounty[9] and the prospects may not have their expected glitter. So what's going on here? Are we succeeding or not? An evolutionary perspective can help answer these questions.

In the main, for complex traits that have been studied with any in-

Figure 3. The lure of hidden treasure, twenty-first-century style. Many companies, especially pharmacological ones, promise ways to find hidden genetic treasures. ``Missing is not an option'' says this ad, which ran in science journals in 2000. Copyright and reprinted with permission from *Compugen*. (Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.)

tensity, many contributing genes, often tens of them, have been identified. For some traits, like obesity and diabetes-related susceptibility, it seems likely that there are *hundreds* of such genes. But most of the individual genetic variants that provide detectable evidence have small effects. Alternatively, if their effect is large, it's rare in the population. Also, "large" usually means raising the relative risk of a tested disease by 10–30%, which translates to modest or small absolute risk effects (Fig. 4). The genes we find are usually called quantitative trait loci (QTL), but often turn out to be what I have elsewhere said may be more aptly referred to as quixotic trait loci.[5] In other words, these effects may be real but, like the treasures in the Spanish galleons, they are elusive and difficult to capture.

The meaning of "complex" is vague, and perhaps is intentionally kept so in the competitive arena of research, promises, and advocacy. However, in general it implies traits that aggregate but do not segregate in families. This means that close family members have similar trait values (like blood pressure or stature), or that risk is higher if one has a close relative who has the trait (like schizophrenia or a type of cancer). Twin concordance, the identification of more than one gene having variation that is clearly associated with a trait or proteins that are expressed in relevant tissues constitute evidence that the trait has a substantial genetic component. Based on such criteria, for most traits the *heritability*, or the fraction of trait variation attributable to genetic contributions, seems to be substantial, usually 30%-60% and sometimes more. But this means that while GWAS are certainly identifying genes, they are generally *not* accounting for most of the total genetic contribution to the tested traits. Those *c*ontributions remain unmapped, in what a story in *Nature* called a hunt for "hidden heritability".[9] That may be a cute way for a gossipy magazine to characterize the story, but it's actually quite misleading. The heritability is not hidden: it's there for all to see. What is unknown are the identities of the individual genes having effects that are hidden within their *aggregate* contribution.

As a case in point, let's look at stature, a simple, relevant, easily measured trait. Stature is one of the most heritable traits, with heritability sometimes reaching around 80% or, in some studies, even greater than 90%. It has a nice, normal distribution in populations, as most

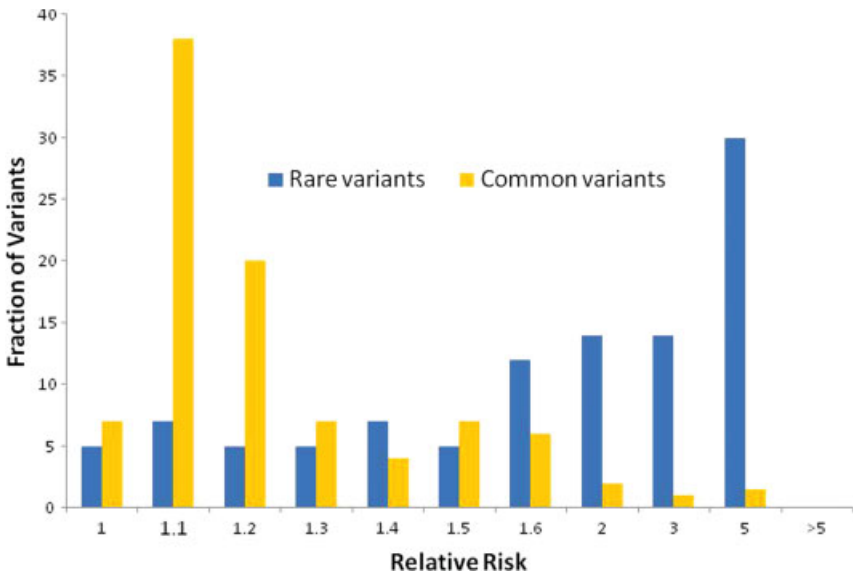

Figure 4. Most common variants found by very large GWAS have small relative risk. Alleles may have greater effect, but are mostly rare. Redrawn from[7] Bodmer and Bonilla.[7] (Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.)

quantitative traits do. Figure 5A shows recruits mustering for military service in World War I in 1914. Tall and short soldiers are unusual; most of them (and us) are in the middle somewhere. If you account for sex and cohort differences, variation in stature is highly correlated among relatives. We know that secular trends in factors such as health and nutrition make large differences, so that shifts in the stature distribution between cohorts, like those in Switzerland over a century (Fig. 5B) are essentially *all* due to changes in life style. However, *within* any given cohort, genetic variation accounts for the great preponderance of stature differences. This suggests that we ought to be able to dissect stature into its individual contributing genes, and many have tried. There are, of course, serious growth dysgenesis syndromes, and these can be monogenic (due to mutations in single, easily identified genes).[10] However, these clear but abnormal extremes generally have reduced fitness, keeping them at very low frequency. They do not contribute to the "normal" variation as usually conceived (Fig. 4). For the latter, results have been striking and less than encouraging for the pursuit of genetic pieces of eight.[10–12]

The bottom line is quite simple. Genetic studies show that stature, as expected, reflects what has theoretically been thought for nearly a century about quantitative and other complex "polygenic" traits. Pooled results from recent, multiple, very large GWAS have identified at least 54 chromosomal regions that have statistically significant effects on stature. Candidate genes that are at least suggestive lie in some of these regions. These genes are ones that are at least involved in relevant physiological processes, such as bone growth.[12] But there is little overlap among studies, and if the top *twenty* such genes are examined together they account for only *3%* of the overall variation in stature, not counting cohort and sex effects.[12] If this is accurate, and the effects of other still-unidentified genes are correspondingly less, as the data suggest, it would take the entire genome to account for all the heritability of stature.

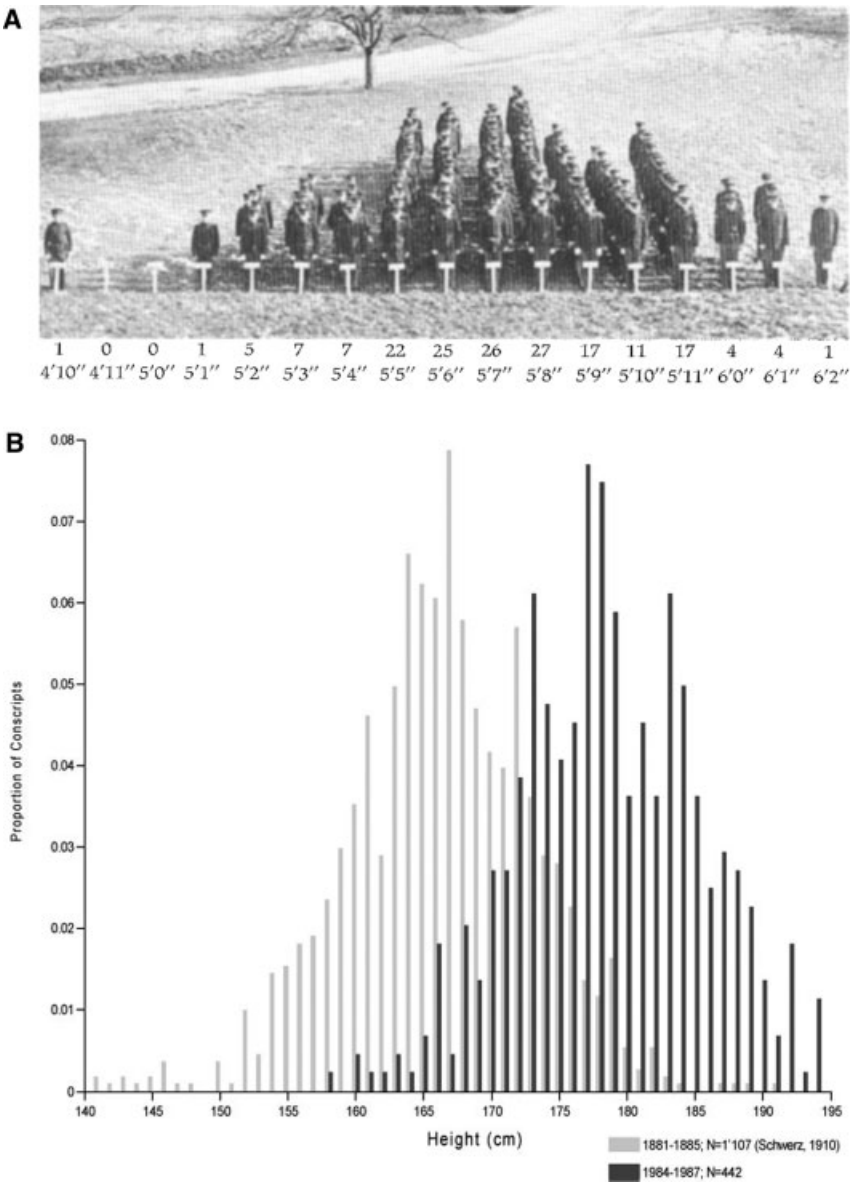

Figure 5. Homoplasy and stature. A. Army draftees for World War I arranged by stature (count and stature bin shown at the bottom).[13] B. Change in stature over a century in Switzerland.[14] Reprinted with permission from J. Wiley & Sons.

HOW IDIOGENIC ARE BIOLOGICAL TRAITS?

A disease is called *idiopathic* if the specific cause is not known. We can use a similar word, *"idiogenic"*, for an instance of a biological trait having a specific genetic cause that isn't known. The degree to which traits are idiogenic varies from clearly *non*idiogenic traits like those of Mendel's peas, where the genetic basis of every yellow or wrinkled pea is known, to traits for which, based on current knowledge, this is hardly ever true, such as schizophrenia. A relevant question for anthropology and, indeed, most areas of biology is the degree to which traits we are interested in are idiogenic. Stature is an example of this within populations, but the same applies to the host of traits among living or fossil primate species, such as limb length, hip shape, and face breadth. This turns out to be highly relevant to the concept of *homoplasy,* similar traits in different species. There can be intense discussion about whether a given instance of homoplosy is due to com-

mon ancestry or parallel evolution. Homoplasy is, of course, important to systematics and taxonomy.

Returning to stature as an example, no two people (other than identical twins) have the same stature-related genotype. And people with the *same* stature have different genotypes. Extremes in the population distribution, such as very tall or short stature (Fig. 5A), are quite rare and might, like genetic stature disorders, be due to single genes, unlike the more common trait values closer to the average. However, GWAS results for many disease as well as normal traits suggest that even most of these are due to unusual combinations of many small effects.

As with other complex traits, most of the individual alleles (genetic variants) contributing to stature have only small effects. Whether they are rare or common in a given population changes from generation to generation, largely by genetic drift (chance aspects of survival and reproduction). But even when a trait like stature is being molded by natural selection, the favored individuals (say, the taller) will be genetically heterogeneous, since many genotypes can confer similar stature and hence similar fitness. This means that the selective effect relative to any one of the contributing genetic variants will be small. That in turn, means, that even when selection is operating, their frequencies change mainly by genetic drift.

As a result, over time within as well as between geographic regions, the spectrum or "landscape" of genotypes differs because the shared alleles vary in frequency and locally unique alleles will have arisen by mutation. Most stature within and between populations is idiogenic. This implies that similar stature among individuals represents what one might call *intraspecific homoplasy*.

Now all one need do is to extend the time horizon enough for species to form, and it is easy to see that intraspecific homoplasy morphs into the classical or interspecific homoplasy that so often is the subject of debate in the systematics community. Whether by drift or natural selection, similar traits in different species can be expected almost invariably to represent different genotypic routes to similar outcomes, like facelength among baboon species. Because developmental mechanisms such as those responsible for traits like stature or facelength are deeply shared phylogenetically, the same genes or genetic pathway networks will at least partly, and usually greatly, be involved and conserved. That, in turn, means that trait similarity usually involves similar genetic systems. To that extent, homoplasy is rare in nature. Yet, at the same time, in most cases the specific allelic variants in those genes, or the members of a gene network making the largest contribution, will differ. To that extent, homoplasy is universal.

Biological systems are built up over eons of evolutionary time. Genes arise by occasional duplication events and new genes can take on related, but diverging functions. Interactions among genes are established bit by bit; we afterwards refer to them as networks or systems. Such interactions and their evolutionary changes occur as their genes' regulatory DNA accumulates mutations, to co-express the genes combinatorially in appropriate cells and tissues at appropriate times and intensity levels.[15] Individual networks that affect development, like signaling systems that induce cell differentiation in developing limbs, teeth, stomachs, or skulls, involve tens of genes. Each trait, even a simple tooth, involves combinations of such systems, whose usage changes during embryological development in a sea of cooperation involving countless factors coded in our genome.[15] Each of these factors is subject to mutation. Although very harmful mutations are weeded out by selection, empirically most mutations have little effect, and the intensity of adaptive selection is usually low.

What this means is that variation is tolerated in this fleet of factors. The gradual buildup of such systems and their variation over countless generations has enabled our traits to evolve in the first place. This also is what makes the traits polygenic, and their individual instances idiogenic. These statements apply to many, if not most traits of interest to anthropology, quantitative as well as qualitative.

All of this makes complete evolutionary sense. It was understood nearly a century ago, and should be entirely be surprising. But its implication is that most complex traits will not be entirely dissectible genetically. If many or most of the variants in the contributing genes are rare in the population or species, each of us bears a different, essentially unique combination of them. We share the pathways and developmental systems, but not the specific set of genotypes in each idiogenic instance of a trait, whether it be the skull, blood pressure, language ability, or stature. Unless this understanding, derived from an evolutionary perspective and totally consistent with the sea of GWAS and other data, is very wrong, it does not bode well for gene-hunting techniques. If there are treasure galleons out there somewhere, we may not be able to find more than a small fraction of them.

In the end, we have a consistent story. Children resemble their parents. A fertilized human egg develops into a human, with its many parts in order. Complex traits are indubitably "genetic" in this sense. Other factors contribute, but genetic mechanisms are ubiquitous. The genetic factors that contribute to a trait can be identified, at least in model systems like the mouse. We can assume that in general these factors are so well conserved that even if we differ in some detail, we basically use homologous systems in our own development, the components and interactions of which can be generally characterized. But the extent to which, even in principle, we can identify the *variation* in these genes that accounts for variation in the trait, or whether we can develop specific *predictions* of the trait in individuals based on their genotype is a more serious and less clear question. The idiogenic variation from person to person means that the individual genotypic contributions are quixotic, if they are even identifiable. Ubiquitous idiogenesis in detail as well as homoplasy in basic genetic systems may also bedevil efforts to

attribute variation in fossils to specific genotypes. Even if not utterly unique, the specific combinations of individual contributors to a given phenotype may simply not occur often enough to be detectable with achievable statistical samples. Even extracting DNA from the fossils themselves may not help, for the same reasons of causal complexity in the living.

"THE SEA CLAIMED DOZENS OF SHIPS AND THOUSANDS OF MEN AND MANY MILLIONS IN TREASURE"[1:21]

The silver lodes that the Spanish found in Bolivia and Mexico did not slake their search for easy wealth. They continued to chase dreams ranging from El Dorado in the New World to islands called Rich in Gold and Rich in Silver (Rica de Oro and Rica de Plata) rumored to be in the Pacific somewhere northeast of the Philippines.[1] Is there likely to be a comparable genetic bonanza that we've not yet hunted down? The answer probably depends on what one wants to find. Whether or not current approaches are the best we could take, we'll certainly accumulate a knowledge bonanza. How or if the GWAS approach will lead to a *financial* bonanza is anyone's guess, but there are many ways to make money, and entrepreneurial cleverness will probably turn that knowledge into some sorts of wealth. But that does not mean we're on the brink of understanding idiogenic causation at a deeper level.

Finding genetic riches is at least as hard and expensive as hunting for elusive galleons. The gleaming genetic sparkles that we can imagine in the phenotypic sea may be an irresistible lure. But the nature of genetic complexity presents a picture in which the wealth of our patrimony is distributed more uniformly, rather than being concentrated in a few lonely galleons sailing somewhere in the vast ocean of possibilities.

As with the Spanish *pesos*, the question is not whether the genetic *pesos* underlying the heritability of complex traits exist, but what we need to know about them, how many we can find, and whether the cost of tracking them down will be rewarded. The question is a serious and fundamental one in biology, human or otherwise.[16,17] There is no obvious answer, because we don't yet have an adequate way to handle the *aggregate* effects of genes on *individuals*, yet that is what we want to know for biomedical genetic prediction or interpretation of fossils. So far, the quest for the genomic pieces of eight has done to us what the sea did to the Spanish. It has engaged dozens of technologies and thousands of investigators and many millions in research funds. Much has been found, but much remains undiscovered.

In the end, the Spanish treasure fleets sailed into history. The last Manila galleon departed in 1815. An international finance system was established based on paper promises such as money, checks, and banking documents, that obviated the need to trade directly in coinage. The billions of pieces of eight that had survived the voyages were distributed, like individual genetic variants, around the world, but were no longer needed in international trade. The modern economy was born of this transition.

Perhaps improved genetics will invent a similar transition, and we won't have to continue the risky, often unproductive enterprise of searching for the genetic coinage beneath the phenotypes we care about. As it did for Spanish treasure, the search for genomic treasure is providing a bonanza for the industries that supply it, including makers of molecular technology, journals, and professors, who are hungry for grants. But that's not the same as finding the lasting wealth that is being promised.

Wealth evolves gradually; each *peso* counts, even if it's not individually identified. In the same way, the causal complexity of biological traits has evolved gradually. Every gene counts, too, but organic wealth may be understood better as a whole, without having to count every genetic penny. That's because evolution itself is pound wise rather than penny foolish.

NOTES

I welcome comments on this column: kenweiss@psu.edu and related blog at http://ecodevoevo.blogspot.com. I have a feedback and supplemental material page at http://www.anthro.psu.edu/weiss_lab/index. shtml. I thank Anne Buchanan and John Fleagle for critically reading this manuscript. This column is written with financial assistance from funds provided to Penn State Evan Pugh professors.

REFERENCES

1 Shurz WL. 1939. The Manila galleon. Manila: RP Garcia Publishing.

2 Walton TR. 1994. The Spanish treasure fleets. Sarasota, FL: Pineapple Press.

3 Norris G, editor. 1994. Buccaneer explorer: William Dampier's voyages. Woodbridge, UK: Boydell Press.

4 Altshuler D, Daly MJ, Lander ES. 2008. Genetic mapping in human disease. Science 322:881–888.

5 Weiss KM. 2008. Tilting at quixotic trait loci (QTL): an evolutionary perspective on genetic causation. Genetics 179:1741–1756.

6 Manolio TA, Brooks LD, Collins FS. 2008. A HapMap harvest of insights into the genetics of common disease. J Clin Invest 118:1590–1605.

7 Bodmer W, Bonilla C. 2008. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 40:695–701.

8 The Wellcome Trust Case Control Consortium. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678.

9 Maher B. 2008. Personal genomes: the case of the missing heritability. Nature 456:18–21.

10 Weedon MN, Frayling TM. 2008. Reaching new heights: insights into the genetics of human stature. Trends Genet 24:595–603.

11 Visscher PM. 2008. Sizing up human height variation. Nat Genet 40:489–490.

12 Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JR, Stevens S, Hall AS, et al. 2008. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 40:575–583.

13 Blakeslee AF. 1914. Corn and men. J. Hered 5:511–518.

14 Ruhli F, Henneberg M, Woitek U. 2008. Variability of height, weight, and body mass index in a Swiss armed forces 2005 census. Am J Phys Anthropol 137:457–468.

15 Weiss KM, Buchanan AV. 2009. The mermaid's tale: four billion years of cooperation in the making of living things. Cambridge: Harvard University Press.

16 Buchanan AV, Sholtis S, Richtsmeier J, Weiss KM. 2009. What are genes "for" or where are traits "from"? What is the question? Bioessays 31:198–208.

17 Weiss KM. n.d. Having a jolly good time—*together* or, evolution by cooperative interaction. Curr Anthropol. In press.