Thaler D.S.
– Biozentrum, University of Basel, Switzerland.
– Program for the Human Environment, Rockefeller University, New York, USA.

Email:
davidsthaler@gmail.com;
david.thaler@unibas.ch

DOI: 10.14673/HE2024121126

KEY WORDS: *Forensics, DNA, Genetics, Intelligence, IQ, Genius.*

Forensics, Genius, and Enthusiasm in the Genetics of Leonardo da Vinci

Why do we want the DNA sequence of Leonardo da Vinci?[1] Three categories, or contexts, suggest themselves: The first is in support of art history, provenance studies, and forensics. The presence of DNA attributable to the master would be evidence of his physical proximity. Depending on details, molecular evidence might lend support to assertions that Leonardo had a hand in creating a particular work. The second category of reasons is to explore the hypothesis that Leonardo's DNA sequence may contain clues to his physical qualities, materials of his work, diet, possible illnesses, and maternal lineage.[2] The possibility that Leonardo's DNA sequence might yield clues to his extraordinary visual sensibilities or even his singular genius deserves to be said explicitly.[3] The third category is not a "reason" in the usual sense but an "apology" in the sense of explanation. It is that this quest is extraordinarily difficult and it engages our enthusiasm. Implications and challenges in each of these three categories are elaborated below.

Art history and Y chromosome evolution

DNA analysis supplies an important means of identification in forensics, tracing family trees, and physical anthropology. It is beginning to be applied to issues of art history[4], provenance, and calibrations of auteurship.

The degree of certainty attributable via DNA analysis depends on several factors including the chain of evidence, the information content yielded by particular methods, the reliability of data on technical grounds, and the databases to which new data are compared.

Historians agree that Leonardo da Vinci fathered no children. However, Leonardo's own father had several other children as did Leonardo's brothers by the same father. The Y chromosome is only passed from father to son. Careful scholarship has identified both deceased and living males who are candidates for the possession of a Y chromosome in unbroken lineage continuity with Leonardo's.[5]

There are three related paths by which Y chromosome analysis may contribute to elucidation of the entire sequence of Leonardo. First, Y chromosome haplotypes could be used as a "ground truth" to support particular relics or archeological findings as likely being Leonardo. Y-corroborated bones or relics would then support further genomic analysis. Secondly, a Y haplotype may indirectly allow reasonable inference of Leonardo's mitochondrial sequence. Thirdly, coalescence analysis, the modeling of how differing alleles among contemporaries descended from a common

ancestor, of multiple Y chromosomes descended from a shared ancestor would allow inference of this portion of Leonardo's genome without danger of information loss associated with the degradation of DNA. Living descendants offer the best- perhaps the only- opportunity to infer complete and accurate T2T (telomere to telomere), i.e. entire and comprehensive coverage of an entire chromosome of Leonardo's genome. Even if exceptionally well-preserved materials were to be obtained, samples donated directly by living relatives are expected to yield the highest fidelity sequence. Complete sequence of the Y from living descendants is needed to obtain as much data and information as possible, rather than limiting the project to forensic identification.

Less is known about Leonardo's mother and no information on the lineage of her mitochondrial genome is currently available.[6] This could change consequent to Y chromosome study. The copy number of mitochondrial genomes in each cell is about a thousand times that of the Y chromosome. In cases of difficult source material mitochondrial sequences are often recoverable when nuclear genes, of which the Y chromosome is one, are not. If even a single case is found where the Y identification is unambiguous and mitochondrial sequences are also unitary, then in other cases where only the mitochondrial sequences are recovered, a reasonable case for identification could be made. Karina Åberg points out that ambiguities of identification based on mixed samples might also be sorted out on the basis of commonalities. There is a prospect, only speculative at present, that mitochondrial sequences might become an additional and quantitatively more sensitive approach to identifying Leonardo's physical intimacy with particular works.

Neither Y chromosomes nor mitochondrial genomes suffer meiotic recombination during gametogenesis and both yield haplotype data. Haplotypes are defined as groups of allelic variants that are inherited together, typically because they are on the same chromosome and meiotic recombination does not separate them.[7] Mutations in the Y and mitochondrial genomes are estimated to occur at several times the rate of somatic chromosomes, an estimate based on sets of parent progeny analysis. The Leonardo project offers an important opportunity to enhance understanding of Y chromosome mutation over multiple generations.

Telomere to telomere, i.e. complete and reliable sequence, was achieved for the human Y chromosome only in 2023.[8] The human Y chromosome is subject to specialized types of sequence variation.[9] Estimates of Y chromosome types and rates of change would be newly empowered via T2T of descendants whose lineage is known over centuries. T2T analysis of multiple descendants of Leonardo's Y would, via coalescence analysis, be best able to accurately predict the primordial sequence. (It is, in many circumstances, as difficult to predict the past as the future.) Coalescence analysis from T2T analysis of multiple contemporary Y's would both reinforce the forensic value of candidate archeological and relics and also give a fresh view of the rates and types of human evolution on the Y. They would be a route to the best and most comprehensive model of the original sequence of Leonardo's Y. A benefit of a

full T2T analysis of contemporary holders of the lineage of Leonardo's Y is that it would offer precise insight, both into Leonardo's genome and more generally, details of human Y chromosome evolution.

The historical development of DNA forensics suffered from underestimates of the importance of characterizing the population to which a new analysis is compared.[10] In the present context, one needs to know the frequency of relevant haplotypes in the population in which Leonardo lived. It may turn out that the mitochondrial haplotype is not as frequent in the local population and therefore more reliable evidence of Leonardo's physical proximity than the Y.

In addition to bones and relics of the conventional sort we note the possibility that single intact cells of Leonardo may one day be found on or in his work, e.g. the notebooks where several fingerprints have been found. Crime scene investigators are developing technologies to isolate individual cells from fingerprints allowing their DNA to be amplified for identification.[11,12] Whole genomic sequence via the single cell amplification is routine in contemporary cell-developmental[13] and cancer biology but has not yet been applied in the contexts of cells recovered from fingerprints or ancient DNA. New methods will need to be developed. In one future-oriented scenario, even a single cell recovered out of a fingerprint of Leonardo's could become a rich source for the master's genomic sequence.

There are works that may have been created by the workshop of Leonardo da Vinci and the master's hand may have been directly involved.[14-16] We can conceive of circumstances in which DNA evidence has the potential to contribute to relevant discussions.

DNA sequence and genius

> *One famous parody describes how a rabbi asked his disciple why the letter peh (p) was needed in the word korah. When the disciple replied that the letter did not appear in the word, the rabbi persisted: "Let us assume for a moment that the letter is placed in this word." "But why should it be needed there?" asked the disciple, to which the rabbi replied: "That is exactly what I asked you."* Steinsaltz, A., *The Essential Talmud*.[17]

At the outset of this section the author should make clear that he is not a genetic determinist on this matter. Genius seems to be recognized in retrospect as a rare coming together of factors including cultural and personal circumstances and luck.[18,19] Having said all that, if one were to consider a type case where DNA sequence might play a detectable role, what better place to consider than Leonardo da Vinci?

Are there clues to Leonardo da Vinci's genius in his DNA sequence? In any non-trivial case guesses are subject to surprise if and when a "black box" can be opened

and its contents examined. It is not necessary to believe in a model to consider the implications "as if" it is true. The following is in the spirit of "AS IF." Suppose, for a moment, there are clues to Leonardo's unique genius in his DNA sequence. Where might they be?

Two distinct hypotheses follow from the "as if" conjecture that Leonardo's DNA sequence will yield clues to his unique genius. The first and less radical possibility is that Leonardo is at the high end of a continuum of human variation in intelligence associated with a particular set of SNPs (Single Nucleotide Polymorphisms). "Generalized intelligence" or "g" is defined as the quantity measured by IQ tests and is reported to overlap with educational attainment. The supposition, challenged in this section, is that by counting and adding up SNPs associated with high IQ, the IQ of a person can be estimated from their genomic sequence.

The second hypothesis, or way of thinking, is more radical: to consider that Leonardo was a qualitative rather than a quantitative variant, a "one off" for which there will be no statistical correlation or continuity with the tail of a normal distribution. These two ways of framing the search for significance in Leonardo's DNA sequence are in analogy to alternative ways to think about the origin of new species in biological evolution.[20] One school of thought considers speciation as continuous with population biology. Another school thinks of speciation events as qualitatively different than everyday processes that alter allelic frequencies within existing populations.

From heritability in monozygotic twins to SNPs in populations: Weak reasoning

Consider the possibility that Leonardo's genius lies at the extreme end of a normal distribution of "g" (General Intelligence, "g", defined as the property measured by IQ tests). There is a literature claiming to map intelligence, as measured by IQ score, to particular SNPs) in the human genome. It has been proposed that the SNP map of Leonardo's genome will give important clues as to his appearance, a subject of interest in its own right. The reasoning here is the same. A critical consideration of SNP mapping of intelligence is best discussed in the context of an important unsolved problem in reconciling the results of monozygotic twin studies with those of SNP studies in unrelated populations.

The reasoning and experimental approach of comparing twins is often associated with Francis Galton in 1875. Galton was the first to use human twin studies in what we now call genetics research but the history of their development in genetic investigations, realizing the critical comparison of monozygotic (identical) twins with dizygotic (fraternal) and the development of "heritability coefficient" as used in contemporary social sciences should not be attributed to Galton.[21]

Monozygotic twins share 100% of the same SNPs, whereas dizygotic twins and other siblings by the same parents share, on average, 50% of the allelic state of each gene ("gene" here is defined as the DNA sequence at a defined locus). To the extent that a population of monozygotic (Mz) twins are more similar for a phenotype than a population of dizygotic (Dz) twins this additional phenotypic similarity is routinely attributed to the additional alleles shared by the Mz twins. Twin studies rate intelligence measured by IQ as a heritable behavioral trait along with many others. An algebraic statement of the reasoning is given by WG Hill: "For quantitative traits, heritability can be estimated as $2(r_{MZ} - r_{DZ})$ where r_{MZ} refers to the co-twin phenotypic correlation for MZ twins and r_{DZ} refers to the correlation for DZ twins."[22]

A diagram of the chain of reasoning that heritability of "g" (20%-80% in different studies) is consequent to SNPs is given in Figure 1. The weaker links in the chain of reasoning are marked in the figure as A), B), C) D), briefly noted in the figure legend and elaborated in the following text.

Let us stipulate for this discussion that many twin studies themselves are well done, interpreted, and complied,[23,24] acknowledging the caveat that experienced environment may be more similar for identical Mz twins than for Dizygotic, Dz twins.[25] In terms of IQ, for young children the hereditary component tends lower and in mature adults it tends higher.[26] IQ heritability in different studies ranges from 20%-80%. Comprehensive review of many studies that collectively cover the lifespan puts the hereditary component of g as measured by IQ at approximately 50%.[23] Creativity is less studied than IQ but literature comparing Mz and Dz twins allows an interpretation less certain but similar to that for other behavioral traits, i.e. significant heritability, plausibly in the range of 50%.[27-29]

Comparison of nature vs nurture or genetic vs environmental influence necessarily depends on the particular environments as well as the genotypes being compared. It is well known that many negative influences of the environment can lead to a decrease in IQ, e.g. heavy metals.[30] There is the less appreciated possibility of unrealized positive environmental influences that go beyond the current idea that the absence of detriment is the best that can be achieved.[31]

The word "heredity" has an ambiguous component, or rather, it is often used with different implicit definitions or easily confused contexts that can lead to substantial misunderstanding and "talking past one another." On the one hand, if environments in a comparison of different genotypes are held constant, then differences in phenotype are often assigned entirely to heredity. On the other hand, in particular environments, some differences in heredity may become either more or less relevant, even irrelevant. Consider for example two genotypes where one of the genotypes requires several times more vitamin C than the other. If a constant environment supplies adequate amounts of vitamin C for both genotypes, then there will be no phenotypic difference between them. If the environment supplies zero vitamin C, both

will be deficient and there also will be no difference between the genotypes. IF AND ONLY IF the environment supplies adequate vitamin C for the low-requiring but not enough for the high-requiring genotype will there be a phenotypic difference between the two genotypes and that difference will be attributable to "hereditary" components.

The "Missing Heritability Problem" refers to a shortfall of SNPs to quantitatively account for results of twin studies.[32] The statistical association of IQ score with each SNP is purported to be in the range of 1/100th of an IQ point. (The range of IQ measurement is ca 50-150 or about 100 points. Put another way each relevant SNP is reported to account for ca 0.005% of the total variance in intelligence.) There are approximately 1,000 IQ-relevant SNPs that are thought to be independent and additive as are most complex traits.[33] In total these 1000 SNPs are believed to account by correlation for ca 10% of intelligence as measured by IQ testing.[34,35] One possibility is that there are many thousands of additional SNPs each of which contributes additional correlation to IQ but that the individual effects are too small to be found even in the sample sizes of hundreds of thousands. Alternatively, the SNP approach has reached a ceiling because non-SNP factors are responsible for the majority of heritability. Two non-SNP possibilities are of special interest: epigenetic marks, and small duplications. The properties of heritability manifested by epigenetic marks and of CNVs differ from SNPs.

Important classes of heritable alterations of genetic material are not expected to show up in SNP studies. These may be candidates for solving the important problem of missing heritability.[36,37] The best studied epigenetic mark in animals is 5-methyl-Cytosine, a covalent modification of DNA that does not change the base sequence. Histone modifications, chromatin remodeling, and non-coding RNAs are additional epigenetic mechanisms that can manifest across one or even several generations.[38,39] CNVs (Copy Number Variants), are expansions and contractions in the genome that can range from a few bases to many thousands. Important consequences follow from the different molecular mechanisms that underlie heritability. Epigenetic marks and CNVs are much less stable over multiple generations than are SNPs.

SNPs are the most stable of genetic variants over many generations; they change only by point mutations at a rate approximated as one or two changes per hundred million bases (1 x 10e-8 > 2 x 10e-8) per generation.[40] CNVs are both more frequent and more dramatic since each can encompass many genes, and occur at rates estimated at 10e-5 to 10e-3 per generation in germ lines[41], a rate one thousand to one hundred thousand times greater than the rate of generation of new SNP variants.[40]

Epigenetic variation is far less stable across generations than either SNPs or CNVs.[42] Twin studies go only over a single generation. The assumption of stability of heritability over multiple generations is hundreds to millions of times more valid for SNPs than for CNV or epigenetic inheritance. The heritability of epigenetic marks is in some systems responsive to modification by parental experience. Epigenetic methylation or small RNA modification of expression across generations and their rapid heritable modifications are not Lamarkian in the sense of specificity, but they violate an absolutist form of the Weissman doctrine that demands complete separation of germ cells from soma and somatic experience. The term "epigenetic" can encompass many different molecular mechanisms and each may have particular characteristics of heritability across generations and propensity to change.[42]

A fascinating study shows that identical twins share epigenetic methylation patterns.[43] This study was well controlled by comparing mono- to di- zygotic twins. Further, the study describes stereotyped epigenomic signatures in sub-telomeric regions that are unique to monozygotic twins. It is now possible to tell if a single person is, in fact, one of a monozygotic twin pair.

Shared patterns of cytosine methylation in monozygotic twins and the instability of cytosine methylation over multiple generations suggest the specific hypothesis that the high phenotypic correlations found in monozygotic twins are attributable to shared patterns of cytosine methylation. Cytosine methylation and other transgenerational epigenetic mechanisms are transmitted over single generations[44] but are not stable over the long term. Cytosine methylation and other epigenetic marks are invisible to SNP studies in either large or small populations.

To conclude: Known SNPs account for approximately 10% correlation with measured IQ scores. One approach is to assume that the remaining "heritability gap" to get to the ~50% or even greater heritability inferred most strongly from Mz twin studies is due to thousands of never-to-be-seen SNPs. There are alternative ideas for how to account for the heritability gap. One approach considers that epigenetics and CNV (Copy Number Variation) are good candidates for the "missing heritability", i.e. the gap between twin-studies-assigned heritability and SNP-assigned heritability. That's because epigenetics and CNV are strong for a single generation but, even when the method allows their detection, too mutable to a give stable signal in population studies.[45]

Epigenetics and CNV are heritable but not with the multi-generational stability of SNPs. Epigenetics and CNVs provide realistic alternatives more amenable to rapid change across generations compared to the SNP-based assignment of unchangeable race-associated "inconvenient" inheritance.[46,47] Molecular mechanisms of IQ heritability may exist but be far less "locked in stone" than supposed by a SNP-limited view of genetics. Stress-induced epigenetic changes are reversible over

a single generation, unlike SNPs.[48,49] It is conceivable but remains to be tested, that one type of epigenetic modification might be inferred from ancient or damaged DNA because 5-methyl cytosine is more prone to deamination than cytosine.[50] Differential U's (and consequent C > T transitions) across the genome might allow inference of local methylation states of different genes or regions.

Leonardo as a "hopeful monster"

Leonardo da Vinci was a singular genius. There is no evidence that any of his relatives or their descendants are of particular talent. Richard Goldschmidt in 1940 proposed a "hopeful monster" theory of speciation through rare and dramatic macromutations.[51] Goldschmidt's theory is in radical contrast to the conventional neo-Darwinian synthesis of population biology.

Let us consider Leonardo a "mutant" or a "hopeful monster" in some unique and positive sense. Thinking in this way has advantages of generating ideas that are more subject to experiment and these are described in more detail in a companion paper on K+ channels and the genetics of vision.

The sequence of Leonardo himself would allow the direct test of hypothesized novel alleles. It is also possible to guess what the allele might be, subsequently to make it by gene synthesis and place it into an appropriate living background. A new type of inspiration from the master: if a unique allele of a gene played a role, then which gene and what allele?

In summary, the hope of getting useful information from SNP analysis for understanding Leonardo's unique genius is slight. At best there might be some sort of correlation and a certain "bragging right" for people who happen to share SNPs with the master. The possibility of special alleles of retinal-expressed potassium channels has the advantage of being an explicit and testable hypothesis. Whether or not Leonardo's DNA is ever analyzed, consideration of the problem has led to ideas that are themselves worthwhile. This point forms a natural transition to the next section concerning another way to think about the Leonardo DNA project.

"Enthusiasm"

Daniel Kahneman and others have shown that human decision making is often of the type I style, i.e. decisions are made at an intuitive level, not with explicit reason or rationality.[52] Reasons often come up later in order to justify decisions already made. When judging work as an appropriate addition to a particular scientific journal or in grant review panels an intangible is often asked of reviewers: What is your level

of "enthusiasm?". I'm told pitch- judging by venture capitalists is similar. "Scientific taste" is a related quality involved in choosing which problems to work on. Communities of practice form around shared enthusiasms and taste, particular ways of doing science encompassing both methods and the type of questions that are of interest.[53] Although narrow communities of aesthetics and practice get routine puzzle-solving work done efficiently, they are necessarily conforming. The rut of communities of scientific conformation can be transcended in three distinct ways. The first is when a routine puzzle turns out to be not merely a puzzle but an insoluble paradox within the reigning paradigm.[54]

The second is when an interdisciplinary group forms to work on problems that require transcending disciplinary silos. A new sort of question invites us to listen and learn from each other. The Leonardo DNA project is this type of creative opportunity. We have the involvement of art historians, curators, restorers, forensic scientists, physical anthropologists, biochemists, and environmental microbiologists of both natural and built environments. In addition to what we bring each other, there is the world around us. COVID-19 dropped into the mix.

The third way a problem inspires is by being hard, apparently insoluble with current knowledge and methods. Exploration into novel and difficult territory is a type of necessity that can be a mother of invention and innovation, perhaps with applications beyond the problem that inspired them. To obtain Leonardo's DNA sequence is a hard, possibly insoluble problem and yet it has the potential to bring out the best in those who undertake to solve it.

One approach the Leonardo DNA project has been pursuing is recovery of DNA from Leonardo's notebooks and drawings. Beyond Leonardo, this concept opens up the exploration of archives and libraries as unintended repositories of DNA, important information not previously envisioned by their curators. Paper archives and libraries had not previously been appreciated in this way as the trend toward digitization and the optionality of hard copy attests. Museums and culture collections have often found uses for their materials that were beyond those originally intended. The use of museum collections for the extraction of DNA has importantly informed obtaining DNA barcodes from type specimens.[55] More recently museum specimens were used to obtain the transcriptome of an extinct Tasmanian tiger proving that under some circumstances RNA and DNA are both stable over the long term.[56]

The COVID-19 global pandemic emerged in 2019 and 2020. How it originated remains unclear and controversial as do the cause-effect relationships of many differing mitigation efforts. COVID-19 is not the first and not likely to be the last pandemic in which ignorance of its origin and the effects of mitigation alternatives limit effective action. The juxtaposition of the Leonardo DNA project and COVID-19 birthed the CoronaCal project, "a biological diary" (CoronaCal) that allows anyone in the community to collect and store serial saliva samples and chart symptoms on ordinary printer paper."[57]

Summary and look to the future

Five forward-looking themes present themselves:

1. The prospect of complete and high-fidelity sequencing of Y chromosomes whose lineage genealogies are of Leonardo. Multiple contemporary Y chromosome data would allow two complementary projects: a) by coalescence analysis, determine the most likely high-detail sequence of Leonardo's Y and b) contribute to knowledge about how Y chromosomes change over generations. Unlike the somatic chromosomes, the Y chromosome does not have a pairing partner during meiosis and its mechanisms of change will be unique.
2. Using the Y as a "ground truth" the potential of locating more of Leonardo's genome both somatic and mitochondrial.
3. The Leonardo DNA project has inspired a specific hypothesis concerning potassium channels expressed in the retina as key to the temporal acuity of vision.
4. The forensic approach to Leonardo's notebooks and related historical correspondence dovetailed with COVID to inspire multilevel biological diaries, the first instantiation being CoronaCal, that together aim to bring historical research methodology into epidemiology.
5. The motivation to recover maximum information from minute amounts of damaged nucleic acids inspires new approaches with wider applications.

Figure 1. The logical progression of Twin studies to SNPs: Assumptions and Weak Points

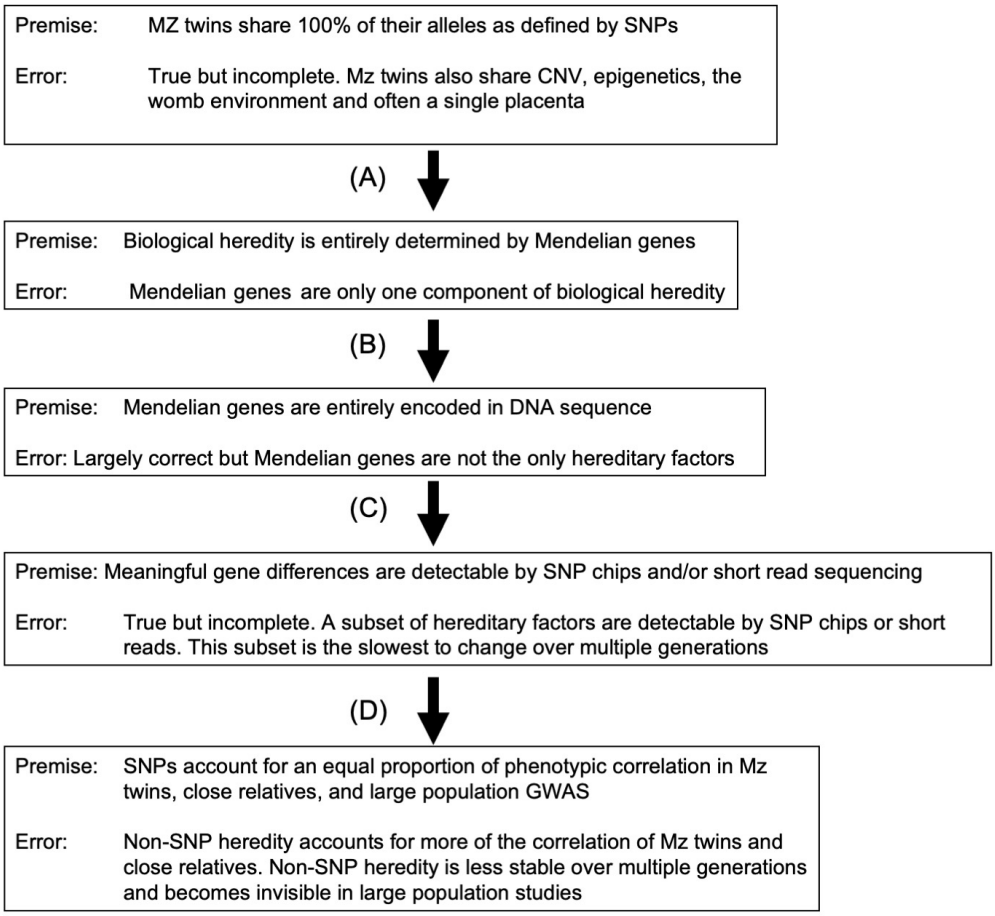

Figure 1. Legend. A) The higher correlation of Mz (Monozygotic) as compared to Dz (Dizygotic) twins is the measure. Measured "heredity" depends on the environments involved. B) Examples of hereditary factors that are not DNA sequence based Mendelian genes include epigenetic modification of DNA (primarily 5-meC), histone modification, and small RNAs. C) The history of human population analysis developed through a SNP chip phase with data from hundreds of thousands of people. This has led to a "light under the lamppost" literature with the assumption that SNPs are all the differences that matter. CNV (copy number variation) and epigenetics are invisible to SNP chips and accompanying analysis. D) Epigenetic heredity

and copy number variations are subject to change over generations at rates many times that of SNP mutation. Perhaps millions of times in excess, arguing that the right conditions may alter relevant hereditary outcomes in a very few generations.

Acknowledgements—Thanks to Karina Aberg, Jesse Ausubel, Guadalupe Pinar, Ylenia Vassallo, Eric Rayman, Thomas Sakmar, Mark Stoeckle, Iddo Wernick, Eric Turkheimer, and all members of the Leonardo DNA project for companionship in related and inspiring work, discussions, and encouragement.

References

1 Ausubel, J. The Search for Leonardo's Genome. *Human Evolution* **37 (3-4)**, 221-228 (2022). https://doi.org/DOI: 10.14673/HE2022341106
2 King, R. Leonardo's Bones: Myth, History,and Evidence. *Human Evolution* **31**, 133-147 (2016). https://doi.org/10.14673/HE201631018
3 Thaler, D. in *Leonard de Vinci, Anatomiste, Pionnier de l' Anatomie, Anatomie comparée, biomécanique, physiognomonie* (eds H de Lumley & Pierre-Marie Lledo) 225-233 (CNRS Éditions, 2021).
4 Begg, T. J. A. *et al.* Genomic analyses of hair from Ludwig van Beethoven. *Curr Biol* **33**, 1431-1447 e1422 (2023). https://doi.org/10.1016/j.cub.2023.02.041
5 Vezzosi, A. & Sabato, A. The New Genealogical Tree of the Da Vinci Family for Leonardo's DNA. Ancestors and descendants in direct male line down to the present XXI generation. *Human Evolution* **36 (1-2)**, 1-90 (2021). https://doi.org/DOI: 10.14673/HE2021121077
6 Kemp, M. *Living with Leonardo: Fifty Years of Sanity and Insanity in the Art World and Beyond*. 288 (Thames & Hudson, 2018).
7 Schulze, T. G. *et al.* Defining haplotype blocks and tag single-nucleotide polymorphisms in the human genome. *Hum Mol Genet* **13**, 335-342 (2004). https://doi.org/10.1093/hmg/ddh035
8 Rhie, A. *et al.* The complete sequence of a human Y chromosome. *Nature* **621**, 344-354 (2023). https://doi.org/10.1038/s41586-023-06457-y
9 Hallast, P. *et al.* Assembly of 43 human Y chromosomes reveals extensive complexity and variation. *Nature* **621**, 355-364 (2023). https://doi.org/10.1038/s41586-023-06425-6
10 Jobling, M. A. Forensic genetics through the lens of Lewontin: population structure, ancestry and race. *Philosophical transactions of the Royal Society of London* **377**, 20200422 (2022). https://doi.org/10.1098/rstb.2020.0422
11 Ostojic, L., O'Connor, C. & Wurmbach, E. Micromanipulation of single cells and fingerprints for forensic identification. *Forensic Sci Int Genet* **51**, 102430 (2021). https://doi.org/10.1016/j.fsigen.2020.102430
12 Huffman, K. & Ballantyne, J. Single cell genomics applications in forensic science: Current state and future directions. *iScience* **26**, 107961 (2023). https://doi.org/10.1016/j.isci.2023.107961

13 Komiyama, T. Diversity of primate brain cells unraveled. *Sci Adv* **9**, eadl0650 (2023). https://doi.org/10.1126/sciadv.adl0650
14 in *Wikipedia: Isleworth Mona Lisa* https://en.wikipedia.org/wiki/Isleworth_Mona_Lisa (2023).
15 Missinne, S. *The Da Vinci Globe*. 291 (Cambridge Scholars Publishing, 2018).
16 Kline, F. *Leonardo's Holy Child: The discovery of a Leonardo da Vinci Masterpiece, A Connoisseurs Search for Lost Art in America*. (Pegasus Books / Simon & Schuster, 2016).
17 Steinsaltz, A. *The Essential Talmud*. (Basic Books A division of Harper Collins, 1976).
18 Robinson, A. *Genius: A Very Short Introduction*. Kindle Edition. edn, 180 (Oxford University Press, 2011).
19 Kalb, C. *Spark: How Genius Ignites, From Child Prodigies to Late Bloomers*. 368 (National Geographic Books, 2021).
20 Gould, S. *The Structure of Evolutionary Theory*. (Belknap Press, 2002).
21 Downes, D. & Turkheimer, E. An Early History of the Heritability Coefficient Applied to Humans (1918–1960). *Biological Theory* **17**, 126–137 (2022).
22 Hill, W. in *'Heritability' in Brenner's Encyclopedia of Genetics (Second Edition)* (Science Direct, 2013).
23 Polderman, T. J. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* **47**, 702-709 (2015). https://doi.org/10.1038/ng.3285
24 Jansen, A. G., Mous, S. E., White, T., Posthuma, D. & Polderman, T. J. What twin studies tell us about the heritability of brain development, morphology, and function: a review. *Neuropsychol Rev* **25**, 27-46 (2015). https://doi.org/10.1007/s11065-015-9278-9
25 Jay, J. *The Trouble with Twin Studies: A Reassessment of Twin Research in the Social and Behavioral Sciences*. (Routledge Taylor and Francis, 2015).
26 Plomin, R. & Deary, I. J. Genetics and intelligence differences: five special findings. *Mol Psychiatry* **20**, 98-108 (2015). https://doi.org/10.1038/mp.2014.105
27 Velazquez, J. A., Segal, N. L. & Horwitz, B. N. Genetic and environmental influences on applied creativity: A reared-apart twin study. *Pers Individ Dif* **75**, 141-146 (2015). https://doi.org/10.1016/j.paid.2014.11.014
28 Segal, N. L. Twin Artists: Unique Sources of Inspiration and Talent/Twin Research Reviews: Ethnic and Racial Factors Affecting Type of Twin Delivery; Outcomes in Twins Conceived Naturally and With Reproductive Assistance; Delivery Time for Growth-Discordant Twins; 3D Study of Twins' Facial Resemblance/In the Media: Baseball Players With Divergent Careers; Basketball Twins in the Elite 8; Twins' Grave and Epitaph; Twins Conceived Three Times in 2 Years; Twins in the Hockey Hall of Fame; Superfetated Twins. *Twin Res Hum Genet* **25**, 206-210 (2022). https://doi.org/10.1017/thg.2022.30
29 Piffer, D. & Hur, Y.-M. Heritability of Creative Achievement. *Creativity Research Journal* **26**, 151-157 (2014). https://doi.org/10.1080/10400419.2014.901068
30 Heidari, S. *et al.* The effect of lead exposure on IQ test scores in children under 12 years: a systematic review and meta-analysis of case-control studies. *Syst Rev* **11**, 106 (2022). https://doi.org/10.1186/s13643-022-01963-y
31 Lederberg, J. Molecular biology, eugenics and euphenics. *Nature* **198**, 428-429 (1963).

32 Matthews, L. J. & Turkheimer, E. Three legs of the missing heritability problem. *Stud Hist Philos Sci* **93**, 183-191 (2022). https://doi.org/10.1016/j.shpsa.2022.04.004
33 Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* **4**, e1000008 (2008). https://doi.org/10.1371/journal.pgen.1000008
34 Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* **50**, 912-919 (2018). https://doi.org/10.1038/s41588-018-0152-6
35 Plomin, R. & von Stumm, S. The new genetics of intelligence. *Nat Rev Genet* **19**, 148-159 (2018). https://doi.org/10.1038/nrg.2017.104
36 Trerotola, M., Relli, V., Simeone, P. & Alberti, S. Epigenetic inheritance and the missing heritability. *Hum Genomics* **9**, 17 (2015). https://doi.org/10.1186/s40246-015-0041-3
37 Slatkin, M. Epigenetic inheritance and the missing heritability problem. *Genetics* **182**, 845-850 (2009). https://doi.org/10.1534/genetics.109.102798
38 Burton, N. O. & Greer, E. L. Multigenerational epigenetic inheritance: Transmitting information across generations. *Semin Cell Dev Biol* **127**, 121-132 (2022). https://doi.org/10.1016/j.semcdb.2021.08.006
39 Zhu, L., Marjani, S. L. & Jiang, Z. The Epigenetics of Gametes and Early Embryos and Potential Long-Range Consequences in Livestock Species-Filling in the Picture With Epigenomic Analyses. *Front Genet* **12**, 557934 (2021). https://doi.org/10.3389/fgene.2021.557934
40 Narasimhan, V. M. *et al.* Estimating the human mutation rate from autozygous segments reveals population differences in human mutational processes. *Nat Commun* **8**, 303 (2017). https://doi.org/10.1038/s41467-017-00323-y
41 Fu, W., Zhang, F., Wang, Y., Gu, X. & Jin, L. Identification of copy number variation hotspots in human populations. *Am J Hum Genet* **87**, 494-504 (2010). https://doi.org/10.1016/j.ajhg.2010.09.006
42 Fallet, M. *et al.* Present and future challenges for the investigation of transgenerational epigenetic inheritance. *Environ Int* **172**, 107776 (2023). https://doi.org/10.1016/j.envint.2023.107776
43 van Dongen, J. *et al.* Identical twins carry a persistent epigenetic signature of early genome programming. *Nat Commun* **12**, 5618 (2021). https://doi.org/10.1038/s41467-021-25583-7
44 Fitz-James, M. H. & Cavalli, G. Molecular mechanisms of transgenerational epigenetic inheritance. *Nat Rev Genet* **23**, 325-341 (2022). https://doi.org/10.1038/s41576-021-00438-5
45 Hu, X. S. *et al.* High mutation rates explain low population genetic divergence at copy-number-variable loci in Homo sapiens. *Sci Rep* **7**, 43178 (2017). https://doi.org/10.1038/srep43178
46 Herrnstein, R. & Murray, C. *The Bell Curve: Intelligence and Class Structure in American Life*. (Free Press, 1996).
47 Wade, N. *A Troublesome Inheritance: Genes, Race and Human History*. (Penguin Random House, 2015).

48 Choi, Y. & Mango, S. E. Hunting for Darwin's gemmules and Lamarck's fluid: transgenerational signaling and histone methylation. *Biochimica et biophysica acta* **1839**, 1440-1453 (2014). https://doi.org/10.1016/j.bbagrm.2014.05.011
49 Wasson, J. A. *et al.* Neuronal control of maternal provisioning in response to social cues. *Sci Adv* **7** (2021). https://doi.org/10.1126/sciadv.abf8782
50 Frederico, L. A., Kunkel, T. A. & Shaw, B. R. A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry* **29**, 2532-2537 (1990). https://doi.org/10.1021/bi00462a015
51 Goldschmidt, R. *The Material Basis of Evolution*. (Yale University Press, 1939; 1982).
52 Morewedge, C. K. & Kahneman, D. Associative processes in intuitive judgment. *Trends Cogn Sci* **14**, 435-440 (2010). https://doi.org/10.1016/j.tics.2010.07.004
53 In *Ludwick Fleck in Stanford Encyclopedia of Philosophy* (First published Mon Mar 19, 2012; substantive revision Fri Oct 8, 2021; Accessed November 27, 2023 https://plato.stanford.edu/entries/fleck/).
54 Kuhn, T. *The Structure of Scientific Revolutions* (University of Chicago Press, 1962).
55 Mulcahy, D. G. *et al.* DNA barcoding of the National Museum of Natural History reptile tissue holdings raises concerns about the use of natural history collections and the responsibilities of scientists in the molecular age. *PLoS One* **17**, e0264930 (2022). https://doi.org/10.1371/journal.pone.0264930
56 Marmol-Sanchez, E. *et al.* Historical RNA expression profiles from the extinct Tasmanian tiger. *Genome research* **33**, 1299-1316 (2023). https://doi.org/10.1101/gr.277663.123
57 Kazmi, M. *et al.* The Coronavirus Calendar (CoronaCal): a simplified SARS-CoV-2 test system for sampling and retrospective analysis. *Front. Epidemiol.* **3** (2023).