Luigi Luca Cavalli-Sforza (1922–2018)

Luigi Luca Cavalli-Sforza passed away on 31 August 2018. Born in Genoa, Italy, in 1922, at age 16 he began his undergraduate studies in medicine in Torino and continued them the next year in Pavia. At age 20 he began publishing research papers on quantitative measurements of bacterial virulence. In this field he worked with his classmate Giovanni Magni, performing experiments using mice inoculated with virulent bacteria. They discovered a linear relationship between mean death time and the logarithm of dose (i.e., the number of bacteria used for inoculations) and proposed an interpretation that isolated the two factors of virulence: the reproduction time of the bacteria and their toxicity (which are related, respectively, to the slope and the intercept of the linear relationship) (Cavalli and Magni 1947).

In 1944, Cavalli-Sforza finished his undergraduate studies and had already published eleven research papers. These were followed by six papers the following year, including his first one coauthored with Adriano Buzzati-Traverso (Buzzati-Traverso and Cavalli 1945), who was his teacher in a genetics course in 1942. Cavalli-Sforza had met Buzzati-Traverso after three years of looking for a mentor who could teach him how to become a researcher (Cavalli-Sforza and Cavalli-Sforza 2005). Buzzati-Traverso’s complete dedication to science deeply impressed Cavalli-Sforza, who considered him one of the people who had a major influence on his life. Their first joint research papers dealt with population genetics of Drosophila (Buzzati-Traverso and Cavalli 1945) and planktonic organisms in lakes (Baldi et al. 1945).

At the end of World War II no jobs were offered at Italian universities, and Cavalli-Sforza worked as a doctor in a hospital during 1944–1945. Discouraged by the lack of drugs to help patients, he found a job at the Istituto Sieroterapico Milanese, a pharmaceutical institute in Milan (Cavalli-Sforza and Cavalli-Sforza 2005). There, during the mornings he extracted blood from patients and did other works, and in the afternoons he managed to perform research experiments that led to publications on quantitative analyses of bacterial resistance to X-rays and other mutagens (Buzzati-Traverso et al. 1945).
He also found time to learn mathematics and statistics on his own.

Without prospects of improving his professional situation, he felt a strong desire to find a full-time research position in England or the United States. Only in 1948 he was awarded a grant to spend some months in England, in the lab of Kenneth Mather, a former student of Ronald A. Fisher. The latter was one of the most important evolutionary geneticists at the time and had also done major foundational work in modern statistics. Fisher had been trying to test his mathematical theory of DNA crossing over using experiments with mice, but he wanted to do it with bacteria because Joshua Lederberg had shown in 1946 that bacteria recombine, and they have a much shorter generation time than mice. Thus, Fisher was looking for a researcher with experience in bacteriology. Mather presumably recommended Cavalli-Sforza to Fisher, who offered Cavalli-Sforza a position in Cambridge during a conference in 1948 (Cavalli-Sforza 1990; Edwards 2009).

Cavalli-Sforza accepted Fisher’s offer, and that same year he moved to Cambridge and built a bacteriology laboratory from scratch in less than four months (Cavalli-Sforza and Cavalli-Sforza 2005). Although unexpected complications made it impossible to test the Fisher mathematical theory of DNA crossing over using bacteria (Cavalli-Sforza 1990), Cavalli-Sforza obtained new results during his stay in Cambridge (his publications during this period included two papers in *Nature*). Lederberg sent him the *E. coli* strain that he had used for his recombination experiments, and Cavalli-Sforza succeeded in isolating a mutant that exhibited an enormously higher frequency of recombination (Cavalli-Sforza 1950). He attained this success by repeating his previous experiments on *E. coli* mutants resistant to agents, but now using Lederberg’s strain (Cavalli-Sforza and Cavalli-Sforza 2005).

Cavalli-Sforza was happy with his research work at Cambridge, but the university did not fulfill the promise of a stable position. More worrying for him was his difficulty in establishing collaborations with researchers outside the Fisher group (Cavalli-Sforza and Cavalli-Sforza 2005). The new director of the Istituto Sieroterapico Milanese visited Cavalli-Sforza to offer him what then seemed a good position in Italy (Cavalli-Sforza 1990), and he accepted it, leaving Cambridge in 1950. During the next years in Italy, Cavalli-Sforza produced what he considered his first important research (Cavalli-Sforza 2007). It dealt with bacterial recombination and was performed in collaboration with Lederberg (Nobel Prize in 1958) and Lederberg’s wife. They discovered that some crosses of *E. coli* strains are fertile and others are not. Thus, if F stands for fertility, the cross $F_+ \times F_+$ is fertile and the cross $F_- \times F_-$ is sterile. The cross $F_+ \times F_-$ is also fertile, because an $F_+$ cell can transfer what they called an F factor (later shown to be DNA) to an $F_-$ cell and convert it to an $F_+$ cell (Lederberg et al. 1952). Cavalli-Sforza and Lederberg did not meet personally until 1954. They then introduced a new, more quantitative method to isolate resistant mutants indirectly (i.e., in the absence of the toxic agent or drug) and applied it to show that all *E. coli* mutants resistant to the antibiotic streptomycin arise spontaneously (Cavalli-Sforza and Lederberg 1955).

Cavalli-Sforza considered Fisher and Lederberg the two most intelligent people he had ever met (Manni 2010) and praised their huge interdisciplinary knowledge and important contributions to several fields (Cavalli-Sforza 2007). James D. Watson, Nobel Prize winner in 1962, wrote in his book *The Double Helix* that Lederberg had performed such an enormous number of successful experiments, and his models were so complicated, that nobody except Cavalli-Sforza dared to work in his field. Clearly Cavalli-Sforza was at the top of world research on bacterial genetics. However, he disliked the pressure of having to publish new results very rapidly, before competing researchers. This was especially difficult in Italy, where he was working on his own, with few means and far from the best research groups in the world. Thus, Cavalli-Sforza gradually shifted to another research field, human genetics.

At the Istituto Sieroterapico Milanese during 1950–1957 Cavalli-Sforza was now director of research. Unexpectedly for him, after a few years this work became depressing because it dealt with pharmaceutical products of very little interest (because the director of the institute had changed again) (Cavalli-Sforza and Cavalli-Sforza 2005). However, Cavalli-Sforza had free time for research and to lecture in genetics and statistics part-time at the Universities of Parma and Pavia (1951–1960). In 1951 one of his students in Parma, Antonio Moroni,
showed him parish books where the Catholic Church had recorded births, marriages, deaths, and dispensations for consanguineous marriages since the sixteenth century. Cavalli-Sforza immediately grasped the great potential of such long-term data for genetic studies, specifically for the analysis of drift.

Genetic drift can be defined as the fact that statistical fluctuations in gene frequencies tend to be larger in smaller populations, due to random sampling. Cavalli-Sforza used parish books to reconstruct the complete genealogy of a village of 150 people (it occupied a whole wall of his studio at his Milano home) (Cavalli-Sforza and Cavalli-Sforza 2005). With such a genealogy, it is possible to detect consanguineous marriages (i.e., those between relatives) in the village. Such marriages are more frequent in the mountains (where villages are smaller) than in the plains. He also began collecting blood samples from the villages in Parma, and after three years, he had enough data to see that genetic variation of blood groups between villages was also greater in the mountains, that is, for lower population densities (Cavalli-Sforza 2003b). This pattern of genetic variation indicated the possible importance of drift.

To check this possibility, beginning in 1963 Cavalli-Sforza, Moroni, and Gianna Zei computed the expected genetic variation using computer simulations. The computer program considered 22 villages, with a total of about 5,000 virtual individuals. Each initial individual was given a blood group for each of three systems (ABO, MN, and Rh) using probabilities equal to the observed blood group frequencies in the total population (thus, there were no initial genetic variations between villages). Simulation runs used a time step of 10 years and parameter values obtained from parish books (e.g., the probability that a person of a given village and age would mate another person, as a function of the village and age of the latter, and the age-dependent probabilities of marriage, reproduction, and death). The simulations correctly reproduced the observed genetic variations between villages (Cavalli-Sforza 2003b). This pattern of genetic variation indicated the possible importance of drift.

In 1960, Cavalli-Sforza left the field of bacterial genetics to devote his efforts exclusively to human population genetics. He became professor of genetics at the University of Parma (1960–1963) and then at Pavia (1962–1970), where he founded the Institute of Genetics. During his years at Cambridge (1948–1950), Cavalli-Sforza had begun to think about the possibility of reconstructing the past of humanity by using genetic data of modern populations (Cavalli-Sforza 1991). He had imagined using present gene frequencies to produce an evolutionary tree that started from a single population (Cavalli-Sforza and Cavalli-Sforza 2005). His reasoning was that, when part of a group migrates, the gene frequencies of both subgroups will gradually diverge. A problem could be that, if natural selection had been the only important evolutionary mechanism, such a tree would mainly describe the history of the local environments encountered by the populations. However, Cavalli-Sforza had established (thanks to the study of Parma populations) the importance of drift. This suggested that pairs of populations displaying more differences (in gene frequencies) should correspond to those that had separated earlier (rather than to those that had lived in more different environments). Thus, such a tree could correctly describe the history of populations. This belief was later supported by the work by Motoo Kimura, who showed that most mutations are selectively neutral (Kimura 1968).

Since 1961, statistical geneticist Anthony Edwards (a former student of Fisher’s) worked with Cavalli-Sforza in Pavia developing several methods to reconstruct phylogenetic trees (Edwards 2009). They were also the first to apply principal component analysis to genetics (Cavalli-Sforza and Edwards 1965). This statistical method makes it possible to assign to each population a point on a graph, by dealing with a few dimensions rather than with all gene frequencies. They found that this approach and a totally independent one (a phylogenetic tree) yield very similar graphs, in which populations on the same continent tend to cluster together. This was one of Cavalli-Sforza’s published works that he liked most, although he regretted that it was practically forgotten because it was published in the proceedings of a conference rather than in a journal (Manni 2010). Another interesting result of this research has been that, when a phylogenetic tree is projected on a world map, the
result agrees with archaeological reconstructions of the out-of-Africa expansion of modern humans (Cavalli-Sforza 1991, 2003b). And in 1988, a strong correlation was found between the world genetic and linguistic trees (Cavalli-Sforza et al. 1988), as expected from the fact that the fission of a human group leads to subgroups that develop their own genetic and linguistic features (Cavalli-Sforza 1991).

In 1971, Cavalli-Sforza moved from Pavia to Stanford, as professor of genetics, finally accepting the repeated invitations by Joshua Lederberg, founder and chairman of Stanford’s genetics department (Cavalli-Sforza and Cavalli-Sforza 2005). This allowed Cavalli-Sforza to develop high-quality collaborations with Stanford professors, including geneticist Walter F. Bodmer (another former student of Fisher’s) (Cavalli-Sforza and Bodmer 1971), biomathematician Marcus W. Feldman (Cavalli-Sforza and Feldman 1981), linguists Joseph Greenberg (Cavalli-Sforza 2004) and Merritt Ruhlen (Cavalli-Sforza 2003b), and many others.

Finally the tiresome bureaucratic duties as director of the Institute of Genetics in Pavia were over. Another difference with his previous position was that Cavalli-Sforza considered that in Italy there was little money for research, and it was badly distributed. In fact, he had received economic support in Italy from the Rockefeller Foundation, both for bacterial and for human genetics research (Cavalli-Sforza and Cavalli-Sforza 2005). Cavalli-Sforza viewed the university systems in Italy, France, and Germany as largely based on tradition, bureaucracy, and pseudo-democracy, and he praised the meritocratic system of private universities in the United States (Cavalli-Sforza and Cavalli-Sforza 2005).

Cavalli-Sforza had interest in the origins of agriculture because he had reasoned that it must have caused a high population growth rate, which should have greatly affected the genetics of populations (Cavalli-Sforza 2004). He met archaeologist Albert Ammerman in a conference on quantitative archaeology in 1970 and began to collaborate with him (Ammerman 2003). They measured the rate of the spread of the Neolithic (i.e., agriculture) from the Near East across Europe (Ammerman and Cavalli-Sforza 1971). Ammerman and Cavalli-Sforza introduced the term “demic” diffusion (from the Greek démos, which means people) to name the diffusion of people (farmers in this case) and distinguish it from cultural diffusion, which refers to the diffusion of culture only (i.e., the conversion of hunter-gatherers into farmers). Cavalli-Sforza invited Ammerman to Stanford, where they continued their collaboration (1972–1977).

In 1973, Ammerman and Cavalli-Sforza applied a demic mathematical theory due to Fisher (called the wave of advance) to model the spread of the Neolithic. They showed that this demic theory (with ethnographically realistic parameter values for the reproduction rate, dispersal distance, and generation time) yields a spread rate close to the observed rate (about one kilometer per year) (Ammerman and Cavalli-Sforza 1973). This suggested that cultural diffusion played a minor role (contrary to the opinion of most archaeologists at the time), but Ammerman and Cavalli-Sforza stressed that most likely a small amount of cultural diffusion would also have been present and that this should have led to genetic clines, that is, gradual variations of gene frequencies as a function of distance from the Near East, where the Neolithic originated (Ammerman and Cavalli-Sforza 1971). This prediction was confirmed later, when such clines were detected by analyzing modern genetic data using maps obtained with the statistical technique of principal component analysis mentioned above (Menozzi et al. 1978). An important, intuitive reason for Cavalli-Sforza to believe in the primacy of demic over cultural diffusion in the spread of the Neolithic was that, in the expeditions to Africa that he had organized since 1965 to study pygmies (Cavalli-Sforza 1986), he had observed the difficulty and rareness of the shift of those hunter-gatherers to a farming way of life (Cavalli-Sforza et al. 1994; Cavalli-Sforza and Cavalli-Sforza 1995). Another reason was that the spread rate of the Neolithic in Europe (about one kilometer per year) was much slower than for mainly cultural prehistoric expansions, such as that of pottery (Cavalli-Sforza 2003a).

About four decades after the proposal by Ammerman and Cavalli-Sforza that the spread of the Neolithic was mainly demic (Ammerman and Cavalli-Sforza 1971), ancient DNA finally showed conclusively that demic diffusion was indeed more important than cultural diffusion in the spread of the Neolithic in Europe (Mathieson et al. 2015), thereby dismissing previous criticisms to their proposal (Cavalli-Sforza 2002). The application of the principal component method to modern European
populations (Menozzi et al. 1978) was extended to the whole world in a book of 900 pages (Cavalli-Sforza et al. 1994), which includes a tremendous amount of archaeological and linguistic information of major processes in prehistory and history. This book appeared after 16 years of work gathering published data and analyzing them (Cavalli-Sforza 2003b). In 2010, Cavalli-Sforza addressed some criticisms to his applications of principal component analysis in an interview published in Human Biology (Manni 2010).

In his many expeditions to study African pygmies (Cavalli-Sforza 1986), Cavalli-Sforza was impressed by the differences between their culture and ours, and this motivated him to devote a large effort to the quantitative study of cultural evolution (Cavalli-Sforza 2004). He published many papers on this topic with Marcus Feldman, including a monograph (Cavalli-Sforza and Feldman 1981) and work comparing their theories to observed data (e.g., Cavalli-Sforza et al. 1982). Cavalli-Sforza considered that anthropology would have progressed much more if it had given cultural transmission the attention it deserves (Cavalli-Sforza 2004). In fact, he had a chance to exert a stronger influence over American anthropologists when he was offered the chair of the anthropology department at Harvard, but he refused the offer because he judged that the Stanford environment was better for his research (Cavalli-Sforza 2000). Among many other applications, the mathematical treatment of cultural transmission (Cavalli-Sforza and Feldman 1981) has been used in the recent development of wave-of-advance models that combine cultural and demic diffusion in a unified framework (Fort 2012).

Since 1992, when Cavalli-Sforza became an emeritus active professor at Stanford, he began to spend half of his time in Italy (Manni 2010). This enabled him to analyze further the genetic data of the Parma valley. In 2004 he published a monograph on these topics with Moroni and Zei (Cavalli-Sforza et al. 2004). Among other results, it reports the remarkable agreement between observed and expected frequencies of consanguineous marriages, between expected genetic drift and that observed using blood group data, and between drift estimated from genes and from surnames. The latter were obtained from telephone directories and had been previously shown to be useful in estimating migration rates (Piazza et al. 1987).

In 2005, Cavalli-Sforza and colleagues found a linear relationship between genetic and geographic distance of two populations in a worldwide sample, with a correlation coefficient of about \( r = 0.9 \) (Ramachandran et al. 2005). According to Cavalli-Sforza, this is essentially the highest correlation ever found in biology (Manni 2010). Moreover, genetic variation within populations decreases linearly with geographic distance from Africa, the origin of modern humans (again, with a correlation coefficient of about \( r = 0.9 \)). They suggested that both correlations were due to repeated drift during the out-of-Africa expansion of modern humans (a so-called serial founder effect). Spatial numerical simulations yield a decrease of genetic variation that agrees with the observed one. This suggests that drift has shaped human evolution and that natural selection may be responsible only for about \( (1 - r^2) \times 100 = 20\% \) of present difference among populations (Ramachandran et al. 2005).

In a popular book (Cavalli-Sforza 2007), published when Cavalli-Sforza was 85, he wrote that he did not feel tired despite his age and that he had discovered new pleasures, such as writing science textbooks for children and popular science books, as well as directing a history of Italian culture in 12 volumes (Cavalli-Sforza 2009). Cavalli-Sforza moved from Stanford to Milan in August 2008.

Cavalli-Sforza often liked to recall that Fisher had a crucial influence on his career. However, they never published a paper together. In what sense was Fisher so influential? Mainly in providing knowledge. Some examples are (a) a theory to analyze data on the selection for heterozygotes in artificial populations of Drosophila, (b) a transformation that Cavalli-Sforza adapted to find a genetic distance useful to build evolutionary trees, and (c) the logarithmic distribution (originally developed by Fisher to analyze the abundances of species), which Cavalli-Sforza and colleagues used to fit surname distributions (Cavalli-Sforza 1990). But for Cavalli-Sforza, the most rewarding use of Fisher’s work was the wave-of-advance model of the spread of the Neolithic in Europe (Ammerman and Cavalli-Sforza 1973). This was an application of Fisher’s wave-of-advance model of advantageous genes (Fisher 1937), which Cavalli-Sforza had learned more than two decades earlier in Cambridge (Cavalli-Sforza 1990, 2003a).

Cavalli-Sforza published about 600 research
papers and numerous research and popular books. His academic honors include memberships of the American Academy of Arts and Sciences (1973), the US National Academy of Sciences (1978), and the UK Royal Society (1992), as well as honorary doctorates by numerous universities, the Golden Medal of the Italian National Research Council (1990), the Catalonia International Award (1992), the Fyssen International Award (1994), the Balzan Prize (1999), and many others.

He has passed away, at age 96, leaving behind very important contributions to knowledge. His intelligence and intuition led to ideas that revolutionized several fields. For example, he correctly predicted major roles for drift in genetics and for demic diffusion in archaeology. These two ideas were not accepted by most experts when he proposed them, but are now well established. He was a genuine researcher, happy to have devoted a substantial part of his life to scientific work. His loss has caused a strong sense of emptiness and sadness, not only because of his discoveries but also because he has had a crucial influence on the scientific life of numerous researchers (as one of many examples, I consider him my scientific father because my career and present research would be totally different had I never known about his work). It will be impossible to forget his huge intelligence, his immense knowledge, his modesty, his kind personality, and his willingness to help. Scientifically, there is no doubt that Cavalli-Sforza has been one of the highest-quality interdisciplinary researchers. Indeed, it seems difficult (perhaps impossible) to find someone who surpasses his proficiency and research results in such diverse fields of the sciences and humanities. This difficulty may perfectly persist during several (perhaps many) generations.

Received 27 September 2018; accepted for publication 1 October 2018.

LITERATURE CITED
