2013 William Allan Award: My Multifactorial Journey[1]

Aravinda Chakravarti[2,*]

Photo courtesy of ASHG.

I want to begin by expressing my sincerest appreciation to The American Society of Human Genetics (ASHG) for this wonderful recognition and award; Vivian Cheung, Evan Eichler, and Stanislas Lyonnet for their nominations; and Evan Eichler for his very kind introduction. I want to thank each of you who are here today and particularly my wife and my daughters, who seldom get to hear what I have been up to. I have been a member of ASHG since 1975, have been attending its meetings since 1977, and have, over the years, seen and heard the very best of our scientists and clinicians receive this honor. I am both surprised and delighted to stand beside them and hope to convince you, over the next 30 min, of my laboratory's deep fascination with human genetics.

I do not stand here either alone or unassisted. I am the product of mentorship by many individuals, first while an undifferentiated undergraduate in India (Profs. J. Roy, B.P. Adhikari, T.A. Davis, and R.L. Brahmachari), then while trying to form my own research ideas in graduate school (Drs. Masatoshi Nei and Alfred Knudson), and finally while a beginning assistant professor at Pittsburgh (Drs. C.C. Li and Haig Kazazian). Even when I thought I knew which ideas to pursue, I was immeasurably helped by a talented group of trainees without whom my accomplishments would have been meager. And then there have been my numerous and diverse collaborators who have taught me new corners of genetics.

So, where does one begin describing one's journey into understanding the molecular basis of complex genetic disease? I took my cue from rereading one of my favorite authors, John Steinbeck, this past summer. Steinbeck begins his 1945 *Cannery Row* as follows:

"Cannery Row in Monterey in California is a poem, a stink, a grating noise, a quality of light, a tone, a habit, a nostalgia, a dream…(how can this) be set down alive? When you collect marine animals there are certain flat worms so delicate that they are almost impossible to capture whole, for they break and tatter under the touch. You must let them ooze and crawl of their own will onto a knife blade and then lift them gently into your bottle of seawater."[1]

So, that's what I will do: let the science from my laboratory "ooze and crawl of [its] own will" across time as my ideas were developed, tested, and then refined. It is possible that some might contest my narrative, but it is my eyewitness account. From the beginning, the work from my laboratory had three cardinal features: it involved both statistical and experimental approaches, both family-based and population-based studies, and an evolutionary perspective. But, to understand how my work developed, I have to start at the beginning.

I was born and grew up in Calcutta, a much maligned and blighted city in the Western imagination but one with a long tradition of excellence, independence, and dissent. Founded in 1690, it is an old city and has been continually exposed to foreign cultures, starting with the Mughals and then continuing with a variety of European influences. Calcutta was the nucleus of the Bengal renaissance and was once the center of Indian education, science, culture, and politics. Yes, I said "once": a reminder that current success is ephemeral. It is a city where

[1]This article is based on the address given by the author at the meeting of The American Society of Human Genetics (ASHG) on October 25, 2013, in Boston, MA, USA. The audio of the original address can be found at the ASHG website.

[2]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
*Correspondence: aravinda@jhmi.edu

http://dx.doi.org/10.1016/j.ajhg.2013.11.014.

diversity is its DNA and its residents are interested in the whole spectrum of the arts and humanities, science, and politics. It is this city's crucible that exposed me to more possibilities than my mind could imagine.[2]

Like my two older brothers, I attended the Calcutta Boy's School (CBS), a Methodist missionary school very much like the fictional English Public School in Lindsay Anderson's movie *If*. CBS taught me to read widely, and I had two dedicated biology and mathematics teachers who taught me, through Mendel's pea experiments, that empirical observations are little without an underlying theory to generalize them. After graduation, I was accepted into the 1970 undergraduate class at the Indian Statistical Institute (ISI) into a program designed by P.C. Mahalanobis, a Cambridge physicist turned statistician, and J.B.S. Haldane, the polymath geneticist who renounced his British citizenship to become an Indian in 1961 and came to work at the ISI. Their joint imagination created a program of 4 years of study in mathematics, statistics, and a variety of scientific disciplines (including laboratory work) that represented their broad scientific interests. Haldane had passed away in 1964, but his ghost was still alive at the ISI when I began; naturally, I focused on human genetics.[2] I left the ISI in 1974 to pursue a PhD in human genetics at a new University of Texas Health Science Center at Houston graduate school headed by Al Knudson, the 1978 president of ASHG and the 1991 recipient of the Allan Award. I studied with Masatoshi Nei (the doyen of neutral molecular evolution in the United States), who first made me work and publish on theoretical aspects of genetic diversity (F_{ST}) before he allowed me to graduate with a dissertation on "The Utility of Linked Marker Genes in Genetic Counseling." I was enamored then of how risk counseling would change if disease-associated markers were available, that is, by linkage disequilibrium (LD) rather than mere linkage.[3,4] I graduated in 1979 to my parents' delight, and my father wrote to say, "Today, your mother and I are so happy; two of our children have PhDs, and the third is successful!" Subsequently, thinking I would learn genetic epidemiology, I pursued 9 months of postdoctoral work with Ryk Ward at The University of Washington, Seattle, but my mind was elsewhere.

1979–1980 was the height of the recombinant DNA period, even prior to the Botstein restriction-fragment-length polymorphism (RFLP) paper,[5] and I was afraid I was unprepared. So, I accepted a teaching position at the School of Public Health at the University of Pittsburgh (no startup, no office, and no moving expenses) and learned recombinant DNA technology in the laboratory of another starting assistant professor, Jim Pipas, working on SV40 large T antigen. The benefit of my teaching position was that I had "free" time that allowed me to work with C.C. Li at Pittsburgh and start a very fruitful and long-term collaboration with Haig Kazazian and Stylianos Antonarakis at Johns Hopkins. I prospered in Pittsburgh but was recruited by Hunt Willard to Case Western Reserve University in Cleveland, where my experimental phase continued with even greater emphasis and I started my genomics studies. I was eventually recruited to the Johns Hopkins University School of Medicine in Baltimore to direct the new McKusick-Nathans Institute of Genetic Medicine in 2000 at a time when the genome sequence was in hand and excitement over its applications to understanding human disease was paramount. To highlight my laboratory's scientific research, I will focus today on only two aspects: how I came upon LD mapping and how I used genomic approaches to understand the molecular basis of Hirschsprung disease (HSCR). I apologize to the many whose work I cannot describe here.

My focus on LD, or allelic association, arose from two unrelated events. First, as an undergraduate, I took Haldane's advice that "one can do human genetics in India without a lot of resources."[6] Consequently, and knowing too little to deter me, I designed, collected, and analyzed familial phenotypic transmission of a simple observable trait: hand clasping, routinely used in many anthropomorphic surveys.[7] I demonstrated that it is inherited in a near-recessive manner and was surprised at its estimated high allele frequency. The trait is common around the world, and I was astounded (and still remain so) as to why such a benign (useless?) trait is so variable. I thought perhaps it was in allelic association with some beneficial mutation, but that would require approaches that were not yet available and also unthinkable. Second, November 1978 was a big month for me. Most importantly, I got married to my wife, Shukti. But, the very same month, I came across Kan and Dozy's ground-breaking paper,[8] which showed three successes: identification of the first human DNA polymorphism, its location in noncoding DNA, and its association (LD) with the sickle mutation! This study was a home run for me and opened many doors for me where none existed. As of today, it is *the* paper that set our field on a different trajectory.

The finding of an association between a disease mutation and a specific marker, even outside a gene of interest, was remarkable given that it immediately suggested a way of mapping and identifying specific mutations: in reverse. However, the scale of such mapping in humans was unknown. I knew that in *Drosophila* large regions could be in allelic association, but usually that was from suppressed recombination within common chromosomal inversions; the human genome was not known to harbor these and was 30 times larger. To begin this LD-based reverse genetics, I required a test case. The globins were one of the most attractive genomic regions to tackle because their mutations were known, their genes had been recently cloned and mapped to individual chromosomes, and their physical arrangements had been elucidated. A chance encounter—my pursuit of data to understand molecular patterns of LD and Haig Kazazian and Stylianos Antonarakis's search to identify the molecular basis of β-thalassemias—led to our collaboration on understanding the patterns of diversity and association in the β-globin cluster. These studies, with my graduate student Ken Buetow, were

the first quantitative studies of LD in the human to give us the scale of population recombination, demonstrate the existence of haplotype "frameworks" (now called blocks), and show that associations can dissipate over very short distances. We inferred that meiotic recombination is nonuniform in the human genome and mapped the location of a putative recombination hotspot adjacent to the β-globin-encoding gene.[9] Although these ideas and observations were contested and very controversial when we published them in 1984, they have stood the test of time and are now commonplace, no doubt because of Mark Daly's subsequent studies on an interleukin cluster.[10]

The research I am not recounting, because it's not my own, is Haig and Stylianos's great success in using the β-globin frameworks for identifying a diverse set of molecular thalassemia mutations; Haig received the Allan Award in 2008 partly for these successes.[11] I wanted to do the same but in a situation where the gene for a phenotype was unknown. In the 1980s, positional cloning had come into its own, but all of the successes, except Stuart Orkin and Lou Kunkel's work on chronic granulomatous disease,[12] were based on a chromosomal lesion interrupting the coding sequence of the relevant gene. LD mapping suggested a new approach provided that the mutant allele was common. The perfect candidate was cystic fibrosis (CF) given that its high incidence defied the mutation-selection balance and suggested a single polymorphic (2%) mutant allele maintained through overdominant selection. In 1985, at these meetings, Lap-Chee Tsui presented the first DNA marker linked to CF. I met Lap-Chee at these meetings, and we personally hit it off. This started my laboratory's work, specifically by graduate student Tara Cox Matise, on LD mapping in CF. Importantly, LD mapping played a very critical role in the positional identification of CFTR in 1989, and our elucidation of haplotypes in CF-affected families uncovered a rich haplotype diversity, of which the most frequent one harbored the now-famous common mutant ΔF508.[13] These studies would have been impossible without Lap-Chee's postdoctoral fellow Bat-Sheva Kerem, who painstakingly generated tens of high-frequency RFLP markers mapped relative to one another (an LD map).

It is not surprising then that the CF studies, conducted in an outbred rather than isolated population, imprinted upon me the necessity of a high-frequency-polymorphism map across the human genome for disease mapping; I was not alone in this realization, given that Neil Risch, Eric Lander, and Francis Collins had individually proposed the same.[14–16] The first of these maps was created by Eric Lander's group.[17] However, it took completion of the human genome sequence for a group of us—largely Francis Collins, Eric Lander, David Altshuler, David Bentley, Peter Donnelly and myself—to eventually champion, participate, and create the high-resolution maps of the HapMap Project for common variants.[18,19] Even prior to these studies, it was clear that having a technology for screening human genes directly for sequence variants, particularly candidates for a given disorder, would be immensely helpful. One of my graduate students, Marc Halushka, worked extensively with Jian-Bing Fan at Affymetrix to complete such a study for candidate hypertension genes by using the then novel array technology.[20] These studies, and a near-identical one from Eric Lander's laboratory,[21] were published back to back in 1999 and showed how diverse some human proteins are and that purifying selection on nonsynonymous alterations is widespread, ideas now commonplace. Consequently, by 2007, a near-complete common-variant map including dense coverage of genes was freely available through the HapMap project, a feature that has changed complex disease mapping and its biology forever. Indeed, with the completion of the 1000 Genomes Project,[22,23] which I also championed and participated in, I cannot imagine a nongenomic genetics world.

Let me now turn to another of my laboratory's parallel developments that intellectually merged with the studies I have described so far. Although my statistical training and interests allowed me to be useful to others in genetic studies of Mendelian disorders, I was far more interested in understanding why some disorders are complex. R.A. Fisher's so-called synthesis in 1918 of the Biometrical-Mendelian debate on the nature of genetic inheritance was an artifice: it was a correct statistical description of complex inheritance, but not a molecular one.[24] So, it was quite lucky that I came across an ideal opportunity: I was asked to review the 1986 Master's thesis of a University of Pittsburgh genetic-counseling student who had studied the genetic epidemiology of HSCR.[25] My simple reading of the literature suggested that this disorder might be amenable to genetic dissection using the evolving genetic maps and positional cloning methods—perhaps not in 1986, but soon. What sealed the deal was the conversation between the two senior committee members: the geneticist lamenting that "we might not ever understand the underlying genes" and the pediatric surgeon arguing that "I am no geneticist, but all I know is that I am repairing the children of kids I repaired 25 years ago…that must count for something." In my mind, that counted for a lot!

HSCR (or congenital aganglionosis) is a neurodevelopmental defect of the enteric nervous system and is characterized by a lack of nerve supply in segments of the gut, leading to an absence of peristalsis and a functional intestinal blockage.[26] It is lethal if not surgically resected, and today, rates of postoperative complication can be as high as 30%. It is considered a multifactorial disorder and is fairly frequent at 1 in 5,000 live births in newborns of European ancestry but is twice as common in Asian infants. HSCR also shows a curious sex bias in that four times as many males as females are affected. The available genetic data in 1986 showed some evidence of discrete genetic loci given that numerous syndromic associations, pedigrees in which HSCR segregated in a near-dominant fashion, and recessive mouse models were already known.[26] To settle the issue, one of my MD-PhD students, Judy Badner, performed statistical analysis of both the

extant published data and families we collected through the Children's Hospital of Pittsburgh to show four remarkable features: its very high heritability, high prevalence of associated anomalies, inverse relationship between the extent of aganglionosis and sex ratio, and predicted genetic etiologies in almost all cases![27] Moreover, nearly 15% of our severe long-segment cases were compatible with incomplete dominance and were thus amenable to genetic mapping using the just discovered microsatellite markers.

In the period from 1993 to 1999, my group and others accomplished precisely that. In 1993, Misha Angrist[28] in my group and Stan Lyonnet[29] in Paris simultaneously published the mapping of a "dominant" form of HSCR to proximal chromosome 10q, and within a year, Stan Lyonnet[30] and Giovanni Romeo[31] in Genoa identified deleterious mutations in *RET*, encoding a receptor tyrosine kinase. We followed with a report on diverse loss-of-function mutations throughout the length of *RET* in all forms of HSCR but found fewer coding mutations (at lower penetrance) than others did.[32] Nevertheless, we failed to find any *RET* variants in the vast majority of the families we studied. In parallel, Erik Puffenberger, another graduate student, followed up on an extensive Old Order Mennonite kindred in which HSCR segregated in an apparently "recessive" form by performing a genome-wide linkage study using identity-by-descent sharing to implicate a locus on chromosome 13q.[33] Subsequently, we engaged in an exciting collaboration with Masashi Yanagisawa at University of Texas Southwestern to implicate a common (15%) hypomorphic missense substitution in the G-protein-coupled receptor EDNRB.[34] Interestingly, this variant showed different degrees of penetrance in all three genotypes and also showed interaction with a common *RET* haplotype. In quick succession, on the basis of our own studies and the published biochemistry, we also demonstrated substitutions in these two receptors' ligands (namely GDNF[35] and EDN3[36]) and, in collaboration with Bill Pavan[37] at the National Institutes of Health, in relevant transcription factor SOX10. Mutations in *SOX10* were identified by Michele Goosens' group[38] the previous year as a cause of Shah-Waardenburg syndrome, one of the HSCR syndromes.

However, the mystery of the reduced penetrance of all of the mutations identified still remained. We hypothesized that full penetrance came about from the joint action of both *RET* and *EDNRB* mutations. Given the commonality of the *EDNRB* variant in Mennonites, in 2002, Minerva Carrasquillo, then a graduate student in my laboratory, performed one of the first microsatellite-marker-based genome-wide association studies in the genetically isolated Old Order Mennonites to demonstrate that strong epistasis between *RET* and *EDNRB* mutations leads to higher penetrance.[39] This work would have been less significant were it not buttressed in this[39] and subsequent work (by Andy McCallion, then a postdoctoral fellow in my laboratory) showing that these epistatic interactions could be fully recapitulated and lead to sex-biased expression of aganglionosis in mice.[40]

Honestly, we were quite pleased because in a matter of 10 years, we had gone from a statistical description of HSCR inheritance to knowing some of its molecular details with precision and how the complex phenotype was synthesized. We postulated a crude theory that HSCR resulted from disrupted biochemical interactions either within enteric neuroblasts or between neuroblasts and the growth-factor-secreting cells of the gut mesenchyme. These studies automatically led to a search for multiple rare mutations within individual patients, and these we and others described. A persistent feature of all of these studies was the primacy of *RET* mutations in our patients. Our idea at that time was that HSCR patients had one primary *RET* mutation and at least one other in a *RET*-dependent modifier. Stacey Bolk Gabriel, not the confident person of today but then a quiet graduate student in my laboratory, performed two of these screens to map such modifiers.[41,42] One of these screens, in short-segment HSCR, was performed with Remi Salomon, visiting from Stan Lyonnet's laboratory in Paris, and clearly demonstrated the primacy of the *RET* effect and the existence of at least two polymorphic modifier loci; interestingly, these three loci recapitulated both the HSCR incidence and recurrence risk.[42]

What was to come was even more surprising. Over time, we found fewer and fewer *RET* coding mutations; I reasoned that our early data were biased upward given that we had primarily collected multiplex families for linkage studies. I hypothesized that if *RET* was involved in nearly every HSCR patient and coding mutations were relatively rare, then the remaining mutations must be in a neighboring gene or a noncoding functional element. In 2005, in an elegant series of experiments, my postdoctoral fellow Eileen Emison and the newly independent Andy McCallion, in collaboration with Eric Green at the National Human Genome Research Institute and his postdoctoral fellow Matthew Portnoy, performed old-fashioned sequencing of retrieved bacterial artificial clones containing *RET* from 12 vertebrates and used comparative sequence analysis to identify conserved noncoding elements.[43] Betsy Grice, a graduate student with Andy McCallion, did functional analyses of 18 elements to identify numerous potential *RET* enhancers, including those acting in gut development in the mouse.[44] Our data on sequencing these enhancers in HSCR patients quickly identified a highly polymorphic variant, called RET+3, in the most potent gut enhancer within intron 1 of *RET* as the causal site.[43] To clarify its effect further, we collaborated in 2010 with the International Hirschsprung Disease Consortium (with principal investigators Salud Borrego [Sevilla], Isabella Ceccherini [Genoa], Aravinda Chakravarti [Baltimore], Robert Hofstra [Rotterdam], Stanislas Lyonnet [Paris], and Paul Tam [Hong Kong]) to publish a large study that clearly demonstrated that this enhancer variant has unique properties and explains many features of HSCR's complex inheritance: it had the highest frequency in the commonest and

least severe cases (simplex males with short-segment disease) and the lowest frequency in the rarest and most severe cases (multiplex females with long-segment disease); interestingly, coding mutations had the precise opposite relationship.[45] Indeed, the enhancer genotypes varied in risk by 200-fold. We also demonstrated by functional analysis that these effects arose from the abrogation of SOX10 binding at the enhancer.[45]

Our recent forays into the effects of *RET* enhancers are even more intriguing. In unpublished work (presented at this meeting) by two current postdoctoral fellows, Sumantra Chatterjee and Ashish Kapoor, we show that HSCR risk is affected by two interacting noncoding elements each imparting differential genetic risk.[46] The second element identified by a genome-wide association study is 125 kb upstream of the transcription start site and is the predicted binding site for a retinoic acid receptor, binding of which is attenuated in the risk variant. We need to prove this hypothesis functionally but believe in its likelihood given that independent work by Robert Heuckeroth at Washington University in St. Louis has shown that vitamin A deficiency in a *Ret* heterozygote mouse leads to aganglionosis.[47] Therefore, *RET* genetic risk is affected by at least two "shadow" enhancers. These two *cis*-regulatory elements, which we show in collaboration with Len Pennacchio and Eddy Rubin, are both active in the gut and neurons, act at slightly different developmental time points, and use the two different transcription factors SOX10 and RARA. This is most likely the first description in humans of complex genetic disease arising from two interacting enhancers that, in the wild-type, provide exquisite control over enteric nervous system development.

To uncover all of HSCR's associated genes, in the past year, we have collaborated with Stacey Gabriel of the Broad Institute of Harvard and MIT in Cambridge to complete initial analysis of the exome sequences of a diverse group of 189 HSCR probands. The chief protagonists of this experiment in my laboratory have been two graduate students, K.D. Nguyen and Tychele Turner, who have had considerable help from Courtney Berrios, my genetics counselor, and Albee Ling. Compared to controls, HSCR genomes show a ~1.5% excess of highly deleterious variants. To identify these specifically, we sought all genes with more deleterious sequence variants than dictated by chance and further narrowed this set to those with three or more deleterious variants. We were astonished to net 363 genes! Is this list credible? We strongly believe so given that both *RET* and *EDNRB* were identified, as were mutations in six other known HSCR-associated genes. Importantly, these 363 genes are highly enriched with those expressed in the embryonic gut, as assessed by our RNA-sequencing studies in mice at embryonic days 10.5 and 14.5 (over which time the gut and enteric nervous system develop) and human induced-pluripotent-stem-cell-derived neural crest cells. In total, 225 of these genes are expressed in the gut, and the functional categories enriched in this set are genes affecting synaptogenesis, cytoskeletal architecture, and the extracellular matrix. We also have begun analyzing these exome sequences for copy-number variants (CNVs) with Nik Krumm and Evan Eichler at the University of Washington, Seattle.[48] A further surprise is that HSCR patients have an excess of recurrent CNVs already known to be associated with other neurodevelopmental disorders, such as autism and epilepsy. In other words, these recurrent CNVs might predispose to a broad neuronal dysfunction whose specificity is produced when they reside in a genome with other genes mutated in HSCR.

All of these are recent results that require validation, confirmation, and peer review. I wanted to share them with you to demonstrate that even a rare disease like HSCR is the result of tens, if not hundreds, of genes and has many valuable lessons to teach us. I will argue that we are just beginning to "understand" this complex genetic defect. But, before we do so, are any of the 225 gut-expressed genes with deleterious variants true? We have begun the painstaking work of individually knocking down their function in zebrafish embryos, an excellent model for studying the migration of enteric neurons, to assess their phenotype. The first class we have focused on is a set of genes likely affecting neuronal ubiquitylation by showing that loss of function of *UBR4*, which encodes an E3 ubiquitin ligase, clearly affects normal ganglionosis in zebrafish and demonstrates strong epistasis with *RET*.

So, what have we learned? We now know the identity of the many genes that are mutant in HSCR and are beginning to understand how they bring about their effects. Enteric nervous system development is highly dependent on a group of early transcription factors, such as PAX3 and SOX10, and two key signaling molecules, RET and EDNRB. Normal development requires passage through this rate-limiting developmental step, which is why deleterious coding variants in these genes are associated with severe disease. Weaker mutations, such as the enhancer alleles I spoke of, might allow passage through this step, which is why less severe and common forms of HSCR occur from the joint action of these weaker deleterious alleles with additional mutations in subsequent developmental processes, such as in the formation of neuromuscular junctions or in neuronal activity per se. Quite importantly, HSCR is not exclusively a neuronal disease given that mutations in genes affecting nonneuronal tissue, such as the extracellular matrix, have now been identified. I have no doubt that human genetics will prove its worth in the subsequent few years to further solve these puzzles.

None of the work I described would have been possible without a strong belief in the awesome power of human genetics, the principles of evolution, and my deep, deep respect for basic and fundamental research. Many years ago, the great biochemist Arthur Kornberg laid down his ten commandments, such as the one on evolution: "What's true for *E. coli* is true for elephants; what's not true for *E. coli* is simply not true."[49] He also humorously wrote that Moses, while coming down from the mountain after his meeting with the Almighty, carried down three

tablets, one of which broke but contained the most important 11th commandment: "Thou shalt respect basic research."[49] In these days, when translation of our science is imminent, the need for basic research is even greater, an aspect that I believe is given insufficient support.

Let me end with a few more thanks. First, to my parents, now both deceased. They did much more than give me their genomes; they were giants and have bequeathed to my brothers and me their lives as lessons. Second, to my beautiful wife and our two lovely daughters. I thank them for their daily support and love and all that it takes to make our house a home; our lives are sometimes chaotic but always interesting. My wife, Shukti, is a noted cell and molecular biologist of her own accord, and I have been lucky to have learned "real" biology from her through osmosis over our 35 years together. Also, thanks to this country and the ASHG, which have given me so many opportunities. However, this has not come without some pain. The life of immigrants, with the absence of our language, culture, and family, is largely an unobserved one to most of you and carries with it a permanent sense of dislocation and loss. Still, it has been a great and worthwhile journey, and I am not about to turn back. I will end by quoting the last paragraph from the last chapter of the book *Interpreter of Maladies*, written by one of the best short-story writers in the United States, the Bengali American Jhumpa Lahiri. She writes in the "Third and Final Continent" of a protagonist, modeled after her father and similar to mine, who left Calcutta to study in England and then at the Massachusetts Institute of Technology in this very city. The year is 1969, and the protagonist rents a room in the home of the 103-year-old Mrs. Croft, who simply cannot believe that man has landed on the moon. It is to this simple beginning on which he reflects:

"In my son's eyes I see the ambition that had first hurled me across the world. In a few years he will graduate and pave his way, alone and unprotected. But I remind myself that he has a father who is still living, a mother who is happy and strong. Whenever he is discouraged, I tell him that if I can survive on three continents, then there is no obstacle he cannot conquer. While the astronauts, heroes forever, spent mere hours on the moon, I have remained in this new world for nearly 30 years. I know that my achievement is quite ordinary. I am not the only man to seek his fortune far from home, and certainly I am not the first. Still, there are times when I am bewildered by each mile I have traveled, each meal I have eaten, each person I have known, each room in which I have slept. As ordinary as it all appears, there are times when it is beyond my imagination."[50]

Acknowledgments

I wish to sincerely acknowledge the support, efforts, and hard work of numerous trainees and collaborators whose work I have described, as well as the patients and their families who lent themselves to our investigations. Importantly, I acknowledge the numerous contributions of the members of the International Hirschsprung Disease Consortium, particularly the junior leaders, Jeanne Amiel, Merce Garcia-Barcelo, Raquel Fernandez, Francesca Lantieri, and Yunia Sribudiani. Finally, I remain deeply indebted to the National Institutes of Health, which, despite my frequent complaints, have steadfastly and consistently supported my diverse research ideas since 1984. I am a paid member of the Scientific Advisory Boards of Affymetrix and Biogen Idec. These potential conflicts of interest are managed by the policies of Johns Hopkins University School of Medicine.

References

1. Steinbeck, J. (1945). Cannery Row (New York: The Viking Press).
2. Chakravarti, A. (2008). An Autobiographical Recollection of the Indian Statistical Institute. *In* A Bouquet of Remembrances, Alumni Association, ed. (Kolkata: Indian Statistical Institute), pp. 22–29.
3. Chakravarti, A., and Nei, M. (1982). Utility and efficiency of linked marker genes for genetic counseling. II. Identification of linkage phase by offspring phenotypes. Am. J. Hum. Genet. *34*, 531–551.
4. Chakravarti, A. (1983). Utility and efficiency of linked marker genes for genetic counseling. III. Proportion of informative families under linkage disequilibrium. Am. J. Hum. Genet. *35*, 592–610.
5. Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. *32*, 314–331.
6. Haldane, J.B.S. (1961). An autobiography in brief (The Illustrated Weekly of India).
7. Chakravarti, A. (1974). Certain considerations on the genetics of hand-clasping in human populations. In Human Population Genetics in India, *Volume 1*, L.D. Sanghvi, V. Balakrishnan, H.M. Bhatia, P.K. Sukumaran, and J.V. Undevia, eds. (New Delhi: Orient Longmans), pp. 277–288.
8. Kan, Y.W., and Dozy, A.M. (1978). Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. Proc. Natl. Acad. Sci. USA *75*, 5631–5635.
9. Chakravarti, A., Buetow, K.H., Antonarakis, S.E., Waber, P.G., Boehm, C.D., and Kazazian, H.H. (1984). Nonuniform recombination within the human β-globin gene cluster. Am. J. Hum. Genet. *36*, 1239–1258.
10. Rioux, J.D., Daly, M.J., Silverberg, M.S., Lindblad, K., Steinhart, H., Cohen, Z., Delmonte, T., Kocher, K., Miller, K., Guschwan, S., et al. (2001). Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat. Genet. *29*, 223–228.
11. Kazazian, H.H., Jr., Orkin, S.H., Markham, A.F., Chapman, C.R., Youssoufian, H., and Waber, P.G. (1984). Quantification of the close association between DNA haplotypes and specific beta-thalassaemia mutations in Mediterraneans. Nature *310*, 152–154.
12. Royer-Pokora, B., Kunkel, L.M., Monaco, A.P., Goff, S.C., Newburger, P.E., Baehner, R.L., Cole, F.S., Curnutte, J.T., and Orkin, S.H. (1986). Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location. Nature *322*, 32–38.
13. Kerem, B.-S., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., and Tsui, L.-C.

(1989). Identification of the cystic fibrosis gene: genetic analysis. Science *245*, 1073–1080.
14. Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science *273*, 1516–1517.
15. Lander, E.S. (1996). The new genomics: global views of biology. Science *274*, 536–539.
16. Collins, F.S., Guyer, M.S., and Charkravarti, A. (1997). Variations on a theme: cataloging human DNA sequence variation. Science *278*, 1580–1581.
17. Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., et al. (1998). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science *280*, 1077–1082.
18. International HapMap Consortium (2005). A haplotype map of the human genome. Nature *437*, 1299–1320.
19. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., et al.; International HapMap Consortium (2007). A second generation human haplotype map of over 3.1 million SNPs. Nature *449*, 851–861.
20. Halushka, M.K., Fan, J.-B., Bentley, K., Hsie, L., Shen, N., Weder, A., Cooper, R., Lipshutz, R., and Chakravarti, A. (1999). Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. Nat. Genet. *22*, 239–247.
21. Cargill, M., Altshuler, D., Ireland, J., Sklar, P., Ardlie, K., Patil, N., Shaw, N., Lane, C.R., Lim, E.P., Kalyanaraman, N., et al. (1999). Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat. Genet. *22*, 231–238.
22. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A.; 1000 Genomes Project Consortium (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.
23. Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A.; 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.
24. Fisher, R.A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. Philosophical Transactions of the Royal Society of Edinburgh *52*, 399–433.
25. Inati, M. (1986). Hirschsprung's Disease: Clinical and Familial Findings in Patients Treated at Children's Hospital of Pittsburgh between 1978–1985. MS Thesis (Pittsburgh: Graduate School of Public Health, University of Pittsburgh).
26. Chakravarti, A., and Lyonnet, S. (2001). Hirschsprung Disease. In The Metabolic and Molecular Bases of Inherited Disease, Eighth Edition, C.R. Scriver, A.L. Beaudet, D. Valle, W.S. Sly, B. Childs, K. Kinzler, and B. Vogelstein, eds. (New York, NY: McGraw-Hill), pp. 6231–6255.
27. Badner, J.A., Sieber, W.K., Garver, K.L., and Chakravarti, A. (1990). A genetic study of Hirschsprung disease. Am. J. Hum. Genet. *46*, 568–580.
28. Angrist, M., Kauffman, E., Slaugenhaupt, S.A., Matise, T.C., Puffenberger, E.G., Washington, S.S., Lipson, A., Cass, D.T., Reyna, T., Weeks, D.E., et al. (1993). A gene for Hirschsprung disease (megacolon) in the pericentromeric region of human chromosome 10. Nat. Genet. *4*, 351–356.
29. Lyonnet, S., Bolino, A., Pelet, A., Abel, L., Nihoul-Fékété, C., Briard, M.L., Mok-Siu, V., Kaariainen, H., Martucciello, G., Lerone, M., et al. (1993). A gene for Hirschsprung disease maps to the proximal long arm of chromosome 10. Nat. Genet. *4*, 346–350.
30. Edery, P., Lyonnet, S., Mulligan, L.M., Pelet, A., Dow, E., Abel, L., Holder, S., Nihoul-Fékété, C., Ponder, B.A., and Munnich, A. (1994). Mutations of the RET proto-oncogene in Hirschsprung's disease. Nature *367*, 378–380.
31. Romeo, G., Ronchetto, P., Luo, Y., Barone, V., Seri, M., Ceccherini, I., Pasini, B., Bocciardi, R., Lerone, M., Kääriäinen, H., et al. (1994). Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. Nature *367*, 377–378.
32. Angrist, M., Bolk, S., Thiel, B., Puffenberger, E.G., Hofstra, R.M., Buys, C.H., Cass, D.T., and Chakravarti, A. (1995). Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung disease. Hum. Mol. Genet. *4*, 821–830.
33. Puffenberger, E.G., Kauffman, E.R., Bolk, S., Matise, T.C., Washington, S.S., Angrist, M., Weissenbach, J., Garver, K.L., Mascari, M., Ladda, R., et al. (1994). Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22. Hum. Mol. Genet. *3*, 1217–1225.
34. Puffenberger, E.G., Hosoda, K., Washington, S.S., Nakao, K., deWit, D., Yanagisawa, M., and Chakravart, A. (1994). A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell *79*, 1257–1266.
35. Angrist, M., Bolk, S., Halushka, M., Lapchak, P.A., and Chakravarti, A. (1996). Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient. Nat. Genet. *14*, 341–344.
36. Bolk, S., Angrist, M., Xie, J., Yanagisawa, M., Silvestri, J.M., Weese-Mayer, D.E., and Chakravarti, A. (1996). Endothelin-3 frameshift mutation in congenital central hypoventilation syndrome. Nat. Genet. *13*, 395–396.
37. Southard-Smith, E.M., Angrist, M., Ellison, J.S., Agarwala, R., Baxevanis, A.D., Chakravarti, A., and Pavan, W.J. (1999). The Sox10(Dom) mouse: modeling the genetic variation of Waardenburg-Shah (WS4) syndrome. Genome Res. *9*, 215–225.
38. Pingault, V., Bondurand, N., Kuhlbrodt, K., Goerich, D.E., Préhu, M.O., Puliti, A., Herbarth, B., Hermans-Borgmeyer, I., Legius, E., Matthijs, G., et al. (1998). SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat. Genet. *18*, 171–173.
39. Carrasquillo, M.M., McCallion, A.S., Puffenberger, E.G., Kashuk, C.S., Nouri, N., and Chakravarti, A. (2002). Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. Nat. Genet. *32*, 237–244.
40. McCallion, A.S., Stames, E., Conlon, R.A., and Chakravarti, A. (2003). Phenotype variation in two-locus mouse models of Hirschsprung disease: tissue-specific interaction between *Ret* and *Ednrb*. Proc. Natl. Acad. Sci. USA *100*, 1826–1831.
41. Bolk, S., Pelet, A., Hofstra, R.M., Angrist, M., Salomon, R., Croaker, D., Buys, C.H., Lyonnet, S., and Chakravarti, A. (2000). A human model for multigenic inheritance: phenotypic expression in Hirschsprung disease requires both the RET gene and a new 9q31 locus. Proc. Natl. Acad. Sci. USA *97*, 268–273.
42. Gabriel, S.B., Salomon, R., Pelet, A., Angrist, M., Amiel, J., Fornage, M., Attié-Bitach, T., Olson, J.M., Hofstra, R., Buys, C., et al. (2002). Segregation at three loci explains familial and population risk in Hirschsprung disease. Nat. Genet. *31*, 89–93.

43. Emison, E.S., McCallion, A.S., Kashuk, C.S., Bush, R.T., Grice, E., Lin, S., Portnoy, M.E., Cutler, D.J., Green, E.D., and Chakravarti, A. (2005). A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. Nature *434*, 857–863.
44. Grice, E.A., Rochelle, E.S., Green, E.D., Chakravarti, A., and McCallion, A.S. (2005). Evaluation of the *RET* regulatory landscape reveals the biological relevance of a HSCR-implicated enhancer. Hum. Mol. Genet. *14*, 3837–3845.
45. Emison, E.S., Garcia-Barcelo, M., Grice, E.A., Lantieri, F., Amiel, J., Burzynski, G., Fernandez, R.M., Hao, L., Kashuk, C., West, K., et al.; International Hirschsprung Disease Consortium (2010). Differential contributions of rare and common, coding and noncoding Ret mutations to multifactorial Hirschsprung disease liability. Am. J. Hum. Genet. *87*, 60–74.
46. Chatterjee, S., Kapoor, A., and Chakravarti, A. (2013). Allelic heterogeneity of and interactions between polymorphic RET enhancers affecting Hirschsprung disease risk. Proceedings of the 63rd Annual Meeting of The American Society of Human Genetics.
47. Fu, M., Sato, Y., Lyons-Warren, A., Zhang, B., Kane, M.A., Napoli, J.L., and Heuckeroth, R.O. (2010). Vitamin A facilitates enteric nervous system precursor migration by reducing Pten accumulation. Development *137*, 631–640.
48. Krumm, N., Sudmant, P.H., Ko, A., O'Roak, B.J., Malig, M., Coe, B.P., Quinlan, A.R., Nickerson, D.A., and Eichler, E.E.; NHLBI Exome Sequencing Project (2012). Copy number variation detection and genotyping from exome sequence data. Genome Res. *22*, 1525–1532.
49. Kornberg, A. (2000). Ten commandments: lessons from the enzymology of DNA replication. J. Bacteriol. *182*, 3613–3618.
50. Lahiri, J. (1999). The Third and Final Continent. In Interpreter of Maladies (New York: Houghton Mifflin Company), pp. 173–198.