PERSPECTIVE

From Phenotype to Genotype: Enter Genomics and Transformation of Primary Health Care around the World

H.C. Slavkin

Abstract: *The progress in phenotype descriptions, measurements, and analyses has been remarkable in the last 50 years. Biomarkers (proteins, carbohydrates, lipids, hormones, various RNAs and cDNAs, microarrays) have been discovered and correlated with diseases and disorders, as well as physiological responses to disease, injury, stress, within blood, urine, and saliva. Three-dimensional digital imaging advanced how we "see" and utilize phenotypes toward diagnosis, treatment, and prognosis. In each example, scientific discovery led to inform clinical health care. In tandem, genetics evolved from Mendelian inheritance (single gene mutations) to include Complex Human Diseases (multiple gene-gene and gene-environment interactions). In addition, epigenetics blossomed with new insights about gene modifiers (*e.g., *histone and non-histone chromosomal protein methylation, acetylation, sulfation, phosphorylation). We are now at the beginning of a new era using human and microbial whole-genome sequencing to make significant healthcare decisions as to risk, stratification of patients, diagnosis, treatments, and outcomes. Are we as clinicians, scientists, and educators prepared to expand our scope of practice, knowledge base, integration into primary health care (medicine, pharmacy, nursing, and allied health science professions), and clinical approaches to craniofacial-oral-dental health care? The time is now.*

Key Words: personalized medicine and dentistry, whole-genome sequencing, bioinformatics, epigenetics, craniofacial-oral-dental diseases and disorders, phenomics.

Background

Sixty-one years ago, Watson and Crick published their one-page paper on the structure of DNA (deoxyribonucleic acid) with possible biological implications (Watson and Crick, 1953). The DNA preparation that was used by Rosalind Franklin for her x-ray crystallography studies was prepared by a young dentist named Norman Simmons. Thirty-nine years ago, I attended the Recombinant DNA Workshop (1975) held at Asilomar, California, when practical guidelines were established, recommended, and adopted by the international community. Twenty-six years ago, James Watson was appointed to direct the Federally sponsored Human Genome Project (HGP). He was replaced a year later by Francis Collins, who led the HGP to completion. In 2000, President Bill Clinton signaled the near-completion of the HGP (Lander *et al.*, 2001; Venter *et al.*, 2001). In 2004, the HGP was completed under budget and under time; it took 13 years and $2.7 billion (Collins, 2010; Feero *et al.*, 2010).

During the early years of the HGP (1988-2004), it took many months to years to complete a single human genome at a cost of millions of dollars. Meanwhile, Federal investments were being made in bioinformatics, instrumentation, and equipment for high-throughput nucleic acid sequencing to accelerate the speed and lower the cost of whole-genome sequencing. In 2007, individual genome sequences with annotations were completed for J. Craig Venter and James Watson (Venter, 2007; Watson, 2009). This heralded the beginning of literally personalized health care. In 2010, Kevin Davies projected the $1,000 individual genome and introduced the revolution in DNA sequencing and the new era of personalized medicine (and dentistry) (Davies, 2010). One year ago, the FDA approved the first methodology and equipment with the capacity to complete a human genome within 24 hours at a cost of less than $5,000 (Collins and Hamburg, 2013).

We now know that the amount of genetic variation between any

DOI: 10.1177/0022034514533569. Center for Craniofacial Molecular Biology, Ostrow School of Dentistry, University of Southern California, 2250 Alcazar Street CSA-103, Los Angeles, CA 90033, USA; Slavkin@usc.edu

two unrelated people is one base or nucleotide (A, adenosine; T, thymidine; C, cytosine; and G, guanosine) *per* thousand bases, or 0.1%. Each person possesses variations in about 3 million bases out of the total 3 billion bases that comprise the human genome. Information from each of our genome sequences, such as single-nucleotide polymorphisms (SNPs), is now used to understand how DNA variations within our genome can affect our health (Kornman and Duff, 2012). Another key consideration to explain genomic variation is called copy number variants (CNVs). There are at least 1,500 CNVs—segments of chromosomes that are duplicated or lost in different people—and these are scattered around the human genome. We are now at the beginning of a new era of using genomic information to make critical healthcare decisions for personalized dentistry and medicine (Collins, 2010; Feero *et al.,* 2010).

Enter Genomics

Historically, genetics is the study of heredity, the process in which a parent passes certain genes on to their children. A person's appearance or phenotype—height, hair color, skin color, and eye color—is determined by genes (genotypes). Additional characteristics or phenotypes affected by heredity or spontaneous genetic mutations include metabolism, mental abilities, natural talents, and susceptibility or resistance to certain diseases or disorders.

Once the HGP was completed in 2004, a much clearer understanding of the human genome emerged. Adult humans have ten trillion somatic cells, each containing 46 chromosomes [2 sex chromosomes and 22 pairs of autosomal chromosomes (named according to their size, 1 being largest and 22 the smallest)]. There are 21,000 functional genes and 19,000 pseudogenes in the human genome within the nucleus of every somatic cell in the body. Slight variations occur in less than 1% of the DNA sequence, resulting in variants of a particular gene called alleles (Feero *et al.,* 2010). Genes are sections of DNA, and the location of a gene is called the gene locus. Most functional genes encode information within exons, separated by introns, to make proteins. Each gene has an upstream enhancer and promoter sequence as well as a terminal stop sequence. Each gene can produce more than one protein through a process called alternative splicing. In addition, the intracellular organelle mitochondria also contain DNA (mitDNA) inherited exclusively from our mothers.

Amelogenesis imperfecta, dentinogenesis imperfecta, familial tooth agenesis, Papillon-Lefèvre syndrome, and sickle cell disease are a few examples of the 10,000 diseases and disorders that are inherited from single gene mutations *via* Mendelian inheritance. More than 60 hereditary disorders are known to result from changes (mutations) in mitDNA associated with a number of phenotypes such as blindness, hearing loss, short stature, and metabolic disturbances. The vast majority of human diseases and disorders are polygenic and reflect many gene-gene and gene-environment interactions coupled with epigenetic gene modifiers.

Curiously, we know that monozygotic twins share a common genotype. Do they share a common phenotype? Recently, several studies have determined that monozygotic twins present phenotypic discordance, such as differences in susceptibilities to disease as well as a wide range of anthropomorphic features. Whereas monozygotic twins are epigenetically indistinguishable during the early years of life, older twins exhibit significant differences in the distribution of 5-methylcytosine DNA and histone acetylation, thereby affecting their gene expression portrait. Succinctly, many newer studies now provide evidence for an appreciation of epigenetics toward understanding different phenotypes that can originate from the same genotype.

Yet another dimension of genomics is that of inherited as well as acquired mutations such as found in various cancers. Multiple endocrine neoplasia type 2 (MEN 2) is an example of an autosomal-dominant hereditary cancer syndrome caused by missense gain-of-function mutations of the RET proto-oncogene and presents strong genotype-phenotype correlations (Frank-Raue and Frank-Raue, 2010). Environmental insults from carcinogens or mutagens, such as found in tobacco products and benzene, are examples of acquired mutations during the lifespan that cause neoplasia as presented in oral and pharyngeal cancers. Several different gene mutations associated with the regulation of the cell cycle and intracellular signaling networks cause various cancers. Genome-wide sequencing of DNA samples from patient lesion biopsy can rapidly inform oncology diagnosis and treatment strategies for chemotherapy (McDermott *et al.*, 2011).

Exemplars: Sickle Cell Anemia, Diabetes, and Periodontal Diseases

Three examples are selected to highlight the introduction of personalized medicine and dentistry. More than half a century after the discovery of the molecular basis of sickle cell disease, the causes or explanations of the phenotypic heterogeneity of the disease are just becoming clear. Sickle cell disease is a genetic disorder in which the beta-chain of the human hemoglobin (Hb) gene is mutated, leading to an abnormal Hb. This mutation causes red blood cells (RBCs) to acquire a sickle shape under conditions of hypoxia, resulting in an array of phenotypes such as anemia, cell adhesion, vaso-occlusion, severe pain, stroke, and organ failure. The genetic mutation is caused by a single amino acid substitution of glutamic acid replaced by valine at the sixth position of the beta-globin chain. This is due to a single nucleotide substitution, GAG → GTG, in codon 6 of the beta-globin gene located on chromosome 11p15.5. Recent studies using SNP genotyping in patients with various phenotypes have discovered significant involvements of SNPs of different genes other than the beta-globin chain (Driss *et al.,* 2009). SNPs in genes implicated in the transforming growth factor-beta/bone morphogenetic protein (TGF-beta/BMP) pathways are associated with a number of phenotypic features of

subsets in patients with sickle cell disease (Driss *et al.*, 2009).

Diabetes is a chronic disease defined by hyperglycemia. Various degrees of insulin resistance and/or dysfunction of the insulin-producing beta cells of the pancreas cause diabetes. The disease phenotype is heterogeneous and can be divided into subtypes depending on the underlying genetic cause(s). The common forms of diabetes—type 1 diabetes (TIDM) and type 2 diabetes (T2DM)—present significant genetic underpinnings but without a clear pattern of inheritance (Hornstein and Shuldiner, 2004). The disease phenotypes are produced from interactions among multiple gene variants with environmental factors. Therefore, performing genetic studies in multiple human populations can identify disease risk alleles that are common in one population but extremely rare in others. This approach has the potential to illuminate pathophysiology, health disparities, and the population genetic origins of disease alleles. A recent study discovered that sequence variants in the SLC16A11 allele are common risk factors for type 2 diabetes in Mexico but extremely rare in European and African populations (Williams *et al.*, 2013). SLC16A11 gene products alter lipid metabolism and produce intracellular increases in triacylglycerol metabolism (Williams *et al.,* 2013). Curiously, poorly controlled diabetes patients (*i.e.,* those people at risk for retinopathy, neuropathy, and macrovascular diseases) would be susceptible to or at risk for periodontitis with progressive connective tissue and alveolar bone loss (Pihlstrom *et al.,* 2005).

Periodontal disease(s) are highly prevalent and affect more than 70% of the global population. This disease results in a significant loss of connective tissue and bone loss associated with tooth loss in adults. In addition to the various micro-organisms that constitute the biofilm, genetic and environmental factors, especially tobacco use, contribute to this disease. Genetic, dermatological, hematological, granulomatous, immunosuppressive, and neoplastic disorders are each associated with periodontal manifestations (Pihlstrom *et al.,* 2005). Additional insights are obtained from the discovery that four different mutations within exons of the cathepsin C gene found on chromosome 11q14 are responsible for Papillon-Lefèvre syndrome (Hart *et al.,* 1999). Other genetic disorders with periodontal disease manifestations include familial and cyclic neutropenia, Langerhans cell disease, Chediak-Higashi, Ehlers-Danlos, Marfan's, Down's, and Kindler's syndromes.

Genome-wide screening can enable craniofacial-oral-dental health professionals to identify and stratify patients at high risk for periodontal disease(s) (Giannobile *et al.,* 2013a,b) and many other craniofacial-oral-dental diseases and disorders (*e.g.*, craniofacial syndromes, head-and-neck trauma, dental caries, head-and-neck cancers, autoimmune diseases such as Sjögren's syndrome, osteoporosis, temporomandibular dysfunctions, chronic facial pain, osteoporosis, etc.). There is a significant opportunity to utilize international collaborations around the theme of genotype-phenotype interactions. Such an effort can advance the discovery of specific genes and complex gene networks that identify subpopulation genotypes at risk, advance precise diagnosis, and design therapeutics that target specific patient genotypes (Slavkin, 2014; Slavkin and Santa Fe Group, 2014).

Revisiting Scope of Practice

The near future for the scope of oral health care practice within the United States and countries around the world is uncertain (DePaola and Slavkin, 2004; Glick, 2009; Frenk *et al.*, 2010). We have become increasingly aware that health reforms with various permutations have emerged and are emerging in different states within the United States and beyond (Glick, 2009; Frenk *et al.,* 2010). Several studies have recommended revisions of the traditional primary healthcare workforce to better address the diseases and disorders within society (Frenk *et al.*, 2010). Several prominent groups have argued for genomics, pharmacogenomics, and immunogenomics as well as the microbiome associated with the human condition to become part of the education and competencies of oral health professionals (Collins and Tabak, 2004; Johnson *et al.,* 2008; Glick, 2009; Slavkin, 2012c).

Education, Research, and Interprofessional Health Care

Since the late 1930s, numerous oral health professionals have embraced human genetics and recommended that it be included as a required competency in dental education, and that multidisciplinary teams, such as craniofacial teams, be formed and maintained to address the special needs of head-and-neck birth defects, trauma, and cancers (Cooper, 1942, 1953; Slavkin, 2012a,b; Fox and Stone, 2013; Slavkin *et al.*, in press). In contemporary terminology, such Interprofessional Education and health teams (IPE) can significantly improve the depth, breadth, and quality of comprehensive and coordinated health care across the lifespan.

Clinician and scientist members within the International Association for Dental Research (IADR), the American Association for Dental Research (AADR), the National Institutes of Health (NIH), and many other biomedical research organizations around the world engage in the discovery of fundamental knowledge about the nature and behavior of living individuals and systems, and apply that knowledge to enhance the human condition. Science and scientific discovery make a difference in each of our lives. Scientific discovery informs clinical health care! We must invest in genomics to enhance clinical oral health care in the 21st century for all people (Slavkin, 2012a,b,c)!

Acknowledgments

The author wishes to acknowledge the pioneering and valiant efforts of Drs. Herbert Cooper, Sy Kreshover, Sam Pruzansky, Carl Witkop, Bob Gorlin, Ray Stewart, Michael Cohen Jr., Tom Hart,

Bob Genco, Chuck Shuler, and Larry Tabak to promote genetics within the education and clinical practices of oral health professionals. The author received no financial support and declares no potential conflicts of interest with respect to the authorship and/or publication of this article.

References

Collins FS (2010). Language of life: DNA and the revolution in personalized medicine. New York, NY: Harper Publisher.

Collins FS, Hamburg MA (2013). First FDA authorization for next-generation sequencer. *N Engl J Med* 369:2369-2371.

Collins FS, Tabak LA (2004). A call for increased education in genetics for dental health professionals. *J Dent Educ* 68:807-808.

Cooper HK (1942). Crippled children? *Am J Orthod Oral Surg* 28:35-40.

Cooper HK (1953). Integration of services in the treatment of cleft lip and palate. *J Am Dent Assoc*. 47:27-35.

Davies K (2010). The $1,000 genome: the revolution in DNA sequencing and the new era of personalized medicine. New York, NY: Free Press.

DePaola D, Slavkin HC (2004). Reforming dental health professions education: a white paper. *J Dent Educ* 48:1139-1350.

Driss A, Asare KO, Hibbert JM, Gee BE, Adamkiewicz TV, Stiles JK (2009). Sickle cell disease in the post genomic era: a monogenic disease with a polygenic phenotype. *Genomics Insights* 2:23-48.

Feero WG, Guttmacher AE, Collins FS (2010) Genomic medicine—an updated primer. *New Engl J Med*. 362:2001-2011.

Fox LM, Stone PA (2013). Examining the team process: developing and sustaining effective craniofacial team care. In: Cleft lip and palate: diagnosis and management. Berkowitz S, editor. Berlin, Germany: Springer-Verlag.

Frank-Raue K, Frank-Raue RS (2010). Molecular genetics and phenomics of RET mutations: impact on prognosis of MTC. *Mol Cell Endocrinol* 322:2-7.

Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, *et al.* (2010). Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. *Lancet* 376:1923-1958.

Giannobile WV, Braun TM, Caplis AK, Doucette-Stamm L, Duff GW, Kornman KS (2013a). Patient stratification for preventive care in dentistry. *J Dent Res* 92:694-701.

Giannobile WV, Kornman KS, Williams RC (2013b). Personalized medicine enters dentistry: what might this mean for clinical practice? *J Am Dent Assoc* 144:874-876.

Glick M (2009). Expanding the dentist's role in health care delivery: is it time to discard the Procrustean bed? *J Am Dent Assoc* 140:1340-1342.

Hart TC, Hart PS, Bowden DW, Michalec MD, Callison SA, Walker SJ, *et al.* (1999). Mutations of the cathepsin C gene are responsible for Papillon-Lefèvre syndrome. *J Med Genet* 36:881-887.

Hornstein RB, Shuldiner AR (2004). Genetics of diabetes. *Rev Endocr Metab Disord* 5:25-36.

Johnson L, Genco RJ, Damsky C, Haden NK, Hart S, Hart TC, *et al.* (2008). Genetics and its implications for clinical dental practice and education: report of panel 3 of the Macy Study. *J Dent Educ* 72(2 Suppl):86-94.

Kornman KS, Duff GW (2012). Personalized medicine: will dentistry ride the wave or watch from the beach? *J Dent Res* 91(7 Suppl):8S-11S.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, International Human Genome Sequencing Consortium, *et al.* (2001). Initial sequencing of the human genome. *Nature* 409:860-921.

McDermott U, Downing JR, Stratton MR (2011). Genomics and the continuum of cancer care. *N Engl J Med* 364:340-350.

Pihlstrom BL, Michalowicz BS, Johnson NW (2005). Periodontal diseases. *Lancet* 366:1809-1820.

Slavkin HC (2012a). Birth of a discipline: craniofacial biology. Newtown, PA: Aegis Communications.

Slavkin HC (2012b). The answer lies in the genome. *Global Health Nexus* 14:6-15.

Slavkin HC (2012c). The evolution of the scientific basis for dentistry: 1936 to now and its impact on dental education. *J Dent Educ* 76:28-35.

Slavkin HC (2014). The future of research in craniofacial biology and what this will mean for oral health professional education and clinical practice. *Aust Dent J* 59(1 Suppl): 1S-5S.

Slavkin HC, Santa Fe Group (2014). Revising the scope of practice for oral health professionals. *J Am Dent Assoc* 145:228-230.

Slavkin HC, Sanchez-Lara P, Chai Y, Urata M (2014). A model for interprofessional health care: lessons learned from craniofacial teams. *J Calif Dent Assoc* (in press).

Venter JC (2007). A life decoded: my genome. New York, NY: Penguin Books.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, *et al.* (2001). The sequence of the human genome. *Science* 291:1304-1351.

Watson J (2009). Living with my personal genome. *Personalized Medicine* 6:607-610.

Watson JD, Crick FH (1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171:737-738.

Williams AL, Jacobs SB, Moreno-Macías H, Huerta-Chagoya A, Churchhouse C, Márquez-Luna C, The SIGMA Type 2 Diabetes Consortium, *et al.* (2013). Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 506: 97-101.