AYURGENOMICS: A NEW APPROACH IN PERSONALIZED AND PREVENTIVE MEDICINE

MITALI MUKERJI[1] AND BHAVANA PRASHER[2]

Genomics has ushered in an era of predictive, preventive and personalized medicine wherein it is hoped that not too far in the future there would be a paradigm shift in the practice of medicine from a generalized symptomatic approach to an individualized approach based on his or her genetic makeup. Several approaches are being attempted to identify genetic variations that are responsible for susceptibility to diseases and differential response to drugs, however, have met with only a limited success. Ayurveda, an ancient Indian system of medicine documented and practiced in India since 1500 B.C has personalized approach towards management of health and disease. According to this system, every individual is born with his or her own basic constitution, termed Prakriti which to a great extent determines inter individual variability in susceptibility to diseases and response to external environment, diet and drugs. This system is in contrast to contemporary medicine, where a preventive and curative regime can be adopted only after an individual suffers or shows signs of an impending illness and there are no methods to identify healthy individuals who would be differently susceptible to disease. We thought an integration of Ayurveda and genomics if attempted in a systematic manner which we call as Ayurgenomics could help fill the gap. In an exploratory study we have provided evidence that healthy individuals of contrasting Prakriti types i.e. Vata, Pitta and Kapha identified on the basis of Ayurveda exhibit striking differences at the biochemical and genome-wide gene expression level. Subsequently, we could also demonstrate that these differences are meaningful since we could follow one of the cues from gene expression differences and identify a genetic marker that is associated high altitude adaptation and a high altitude illness. Our studies have provided a novel molecular framework for integration of these two disciplines for predictive and personalized medicine.

There has been a steady rise in prevalence of common diseases like diabetes, asthma, cardiovascular disorders, epilepsy, stroke, schizophrenia and bipolar disorder in the world. These have been mainly attributed to changes in life styles and dietary habits as well as inability to adapt to new habitats and environments. Nearly 1% of the world population suffers from these diseases and in some cases like diabetes and obesity, the numbers are considerably higher and also dangerously on the rise. With the steady increase in average life expectancy of humans, this has become a major concern since most of these common diseases extend throughout life and require long-term medication often associated with added complications and/or expensive interventions. Identifying factors that predispose individuals to these diseases and predict their progression as well as designing customized drug regime for each individual to minimize side-effects is a major challenge.

It is now well acknowledged that a large number of diseases are inherited in the families. With the discovery of the double helical structure of DNA and subsequent advancements in the field of molecular biology and genetics, the cause of pathogenesis in many of these

[1] Genomics and Molecular Medicine, Institute of Genomics and Integrative Biology, (IGIB) CSIR, Mall Road, New Delhi, 110007, India.

[2] Planning and Performance Division, Council of Scientific and Industrial Research, Anusandhan Bhawan, 2, Rafi Marg, New Delhi, 110001, India

Correspondence to Mitali Mukerji (mitali@igib.res.in) Phone 91-11-27666156, Fax 91-11-27667471 and Bhavana Prasher (bhavana@csir.res.in) 91-11-23710158, Fax 91-11-23736842

diseases has been traced to changes in the DNA from one individual to another. The human genome, that is the entire genetic content, is composed of 23 chromosome pairs (diploid) where each set (haploid) has 3 billion base pairs of DNA inherited from either of the parents. There are large numbers of variations in the human genome sequence which are called Single Nucleotide Polymorphisms (SNP). Some of these variations are present in large number of individuals and are called as common variations and some are rare. If the variations are present in less than 1% of the population they are mostly classified as mutations. Many rare diseases like hemophilia, beta-thallasemia etc. are monogenic, caused due to mutations in single genes. Most of the common diseases such as diabetes, asthma, cardiovascular disease and so on are multigenic complex disorders involving many genes. It is generally observed that common diseases are a consequence of cumulative effect of a large number of variations in the genome which independently have small effects that are not sufficient to cause the disease. However, it is now being increasingly realized that even those diseases that were considered to be monogenic sometimes exhibit differences in manifestation of disease in different individuals in spite of carrying the same mutations. This is thought to be due to presence of variations in other genes that could modify the effect of the primary mutation. Further there is a complex interplay of gene and environment involved in most of the diseases. For example, in cardiovascular disease (CVD), various parameters like blood pressure, levels of lipoproteins (HDL and LDL), triglycerides and total cholesterol in the blood along with life style habits such as diet, smoking and lack of exercise, stress etc have been identified as risk factors in these diseases. Each of these parameters can be modulated by a large number of genes. Thus an astronomical number of possibilities of combination of variants from different genes and environment could contribute not only to differences in clinical manifestation of disease but also to the variability in age of onset, severity and symptoms of the diseases. Another aspect of the disease is the drug dosage management. Most of these diseases require long term drug administration and there is a high variability in individual response to drug dosage and adverse effects due mainly to variations in the genes responsible for drug transport and drug metabolism within the individual's system. Therefore design of optimum dosage with least side-effects is difficult to establish. Thus an important starting point in understanding the factors responsible for these diseases and how the treatment regime can differ from individual to individual, is to study the prototype sequence of a human genome that could be used as a reference for comparison between healthy and affected individuals. With the availability of the complete sequence of the human genome, it is now possible to entertain the thought that not too far in the future, each individual would have a personalized health regime based on his/her genetic make-up.

The first draft of the Human Genome sequence made available in 2001, came with its own share of surprises, the most striking amongst them was that we did not have "The Human Genome Sequence" as mentioned above[1,2]. Two unrelated humans differ on an average one nucleotide every 1000 base pairs. If we extend that to the entire genome (haploid size $3x10^9$ base pairs) these would translate to over 6 million differences between any two individuals. This makes it nearly impossible to reconstruct a unique prototype sequence of 3 billion nucleotides which we could refer as "The Human Genome Sequence". Re-sequencing of entire genomes in the recent times of a couple of prominent humans has further confirmed that there are extensive differences between individuals not only with respect to single nucleotides (i.e. SNP) but also large scale gain and loss of chromosomal segments due to genome rearrangements. Discovering meaningful variations (between the healthy and affected) from this vast pool of common variations (between two healthy) is now an insurmountable challenge.

The observed variability in the human genome sequence led to a natural progression to an International HapMap Consortium project[3,4]. This project was undertaken to catalog all human genome variations and to infer patterns of variations across global populations. The project was primarily aimed at speeding the discovery of genes related to common illnesses through capturing extent of genome-wide diversity across 269 individuals from Caucasian, Chinese, Japanese and Yoruban populations. These efforts have uncovered nearly 11 million SNPs in the human genome wherein nearly 4.2 million SNPs have frequency higher than 1% in all the major populations. Though all the variations may not be functional, these variations could act as landmarks and a combination of these landmarks (defined as haplotypes) can define uniqueness of genomes of different individuals/populations. It is generally seen that specific mutations mostly arise once in the population and spread with the subsequent growth of the population and are introduced in different populations due to migrations and subsequent admixture. If a mutation arises in a specific haplotype background (that can be defined by a combination of SNP landmarks) the frequency of this haplotype would differ between the cases (diseased individuals) and controls (healthy

individuals). Once these haplotype backgrounds are identified the causal mutation can be traced relatively easily through sequencing the regions demarcated by the landmarks (haplotypes). This is similar to reaching a destination following a combination of landmarks wherein each landmark per se might not be so informative. The haplotype information generated in the HapMap data for different populations can be used for identification of mutations through comparison of frequencies between controls and cases. Based on the differences in frequencies of these haplotypes between affected and unaffected individuals an ODDs ratio can be calculated for that haplotype to be associated with the disease. Therefore it was anticipated that this basal variation data would be useful for not only improved understanding of disease etiology, mechanism and discovery of new drug targets but would provide predictive marker for disease risk assessment, prevention, progression as well as differential drug response and side effects.

Since most of the genes in complex disorders are undiscovered, it has been presumed that an unbiased and exploratory approach which could compare genome-wide variations between cases and controls through comparison of variations at genome wide scale could reveal causal haplotypes. With advancement in genomic technologies it has become possible to interrogate the entire genome-wide variations of an individual in one single experiment. This led to studies called as "Genome Wide Association Studies (GWAS) " where variations in its entirety for many diseased and healthy subjects could be compared to identify regions that had sufficiently different frequency and could be associated with the disease[5]. Once these regions are identified one could look at the genes and start to make sense of the underlying biology of the disease. These experiments have provided valuable insights into genetic underpinnings of diseases and also seemingly unrelated diseases are getting linked through common set of genes and pathways identified through these studies. Although in these experiments many disease gene associations were observed but with very less ODDs ratio that was predicted to confer only slightly higher risk to the diseased population as compared to controls. As these studies were carried out in large number of individuals, the power of the study was adequate however, the inferences were inadequate[6]. In most of these studies the cases and controls are chosen from the same ethnicity, gender and age to remove the effect of any of these factors on the disease condition. These studies have been highly dependent on strong contrasts in disease susceptibility between cases and controls. Controls do not have obvious clinical disease and in general comprise of a heterogeneous set of individuals who are differently predisposed or protected from diseases. It is being increasingly realized that identification of sub-groups within normal controls corresponding to contrasting disease susceptibility is likely to lead to more effective discovery of predictive markers for diseases. There are however no modern methods available to look at inter-individual differences within ethnically matched healthy populations. We, at the Institute of Genomics and Integrative Biology, have been exploring the concept whether Ayurveda, an ancient Indian system of predictive and personalized medicine, can fill this gap and help in identification of predictive markers for some of these complex diseases. This integrated field of research which has been initiated since 2001.

Ayurveda, an ancient Indian system of Medicine that has been documented and practiced since 1500 B.C. is a living tradition of healthcare even today. This can be judged by the fact that nearly 65% of India uses traditional medicine with its growing acceptance as an alternative medicine in many parts of the world. Ayurveda has personalized approach in predictive, preventive and curative aspects of medicine. It deals with inter-individual variability in assessing susceptibility, establishing diagnosis, and prognosis mainly on the basis of constitution type of the individual ("*Prakriti*"). selection of a suitable dietary, therapeutic and life style regime is made on the basis of clinical assessment of the individual keeping one's *Prakriti* in mind[7-10]. This is in contrast with modern medicine wherein assessing susceptibility might be based on genetic markers, diagnosis based on objective parameters, dietary and life style recommendation are disease based and treatment is mostly symptomatic with the dosage management mostly empirical. Though many sophisticated and state-of-the-art methods are available, there is minimal cross-talk among any of these steps mentioned above. We felt it important to explore whether the ancient and documented system of constitution types described in Ayurveda could be correlated with modern biology (Figure 1).

Prakriti is a consequence of the relative proportion of three entities (Tri-Doshas), Vata (V), Pitta (P) and Kapha (K), which are not only genetically determined (Shukra Shonita), but also influenced by environment (Mahabhuta Vikara), especially maternal diet and lifestyle (Matur Ahara Vihara), and age of the transmitting parents (Kala - Garbhashaya). The ethnicity, familial characteristics as well as place of origin of an individual are also described to influence development of *Prakriti* besides the afore-mentioned individual specific factors. The *Prakriti* of an individual is fixed at the time of birth and remains invariant

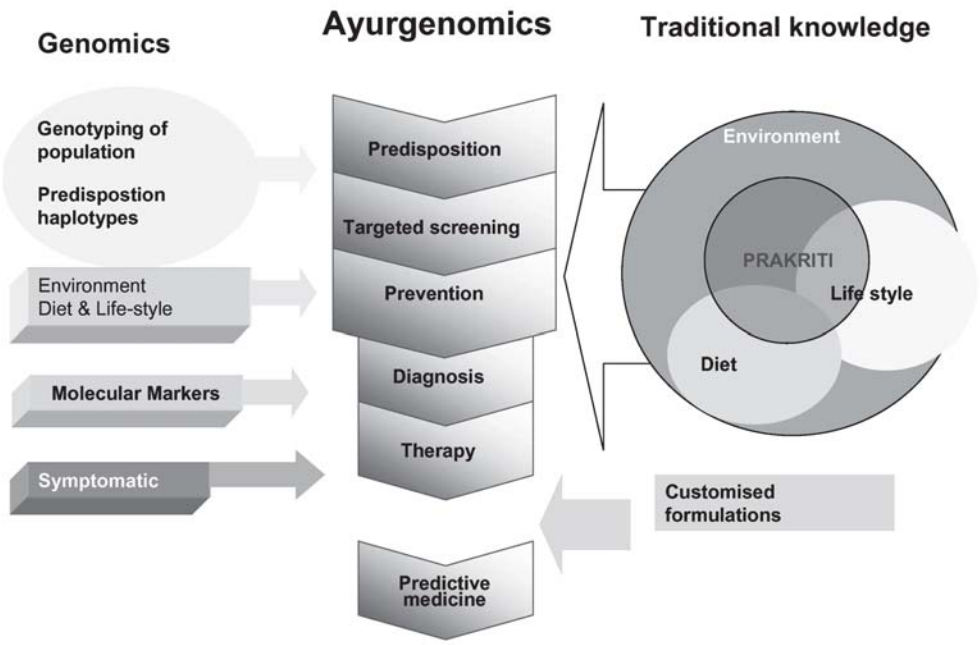

Figure 1. Proof of concept of the study

throughout the lifetime. In an individual, the tri-doshas work in conjunction and maintain homeostasis throughout the lifetime starting from fertilization. Distinct properties and functions have been ascribed to each Dosha. The kinetic components of a system have been ascribed to Vata, the metabolic components to Pitta and the structural and stability components to Kapha. For instance, Vata contributes to manifestation of shape, cell division, signaling, movement, excretion of wastes, cognition and also regulates the activities of Kapha and Pitta. Kapha is responsible for growth and maintenance of structure, storage and stability. Pitta is primarily responsible for metabolism, thermo-regulation, energy homeostasis, pigmentation, vision, and host surveillance. Hence the differences in Tridoshic proportions right from the time of fertilization are manifested as different phenotypes that can be with respect to external appearances, body physiology, and response to external environment etc. Thus a continuum of relative proportions of Doshas results in seven possible constitutional types namely Vata, Pitta, Kapha, Vata-Pitta, Pitta-Kapha, Vata-Kapha and Vata-Pitta-Kapha. Amongst these, the first three are considered as extremes, exhibiting readily recognizable phenotypes that are evident not only at the anatomical and physiological level but also at the level of mental aptitude. However, the individuals of such predominant Vata, Pitta, Kapha *Prakriti* are relatively infrequent in the population. At the anatomical level these constitution types differ with respect to body frame and build, skin, eye and hair colour texture and composition; at the physiological level the differences are exhibited with respect to food and bowel habits, tendency to gain weight, disease resistance and healing capacity, tolerance for specific weather, metabolism of toxic compounds etc. Besides these constitution types have specific likes, dislikes and suitability of tastes and there are differences in memory retention as well as aptitude differences (see Table).

The constitution type and *Prakriti* levels of dosha are considered as normal for that individual. Any perturbation from an individual's homeostatic state of Doshas leads to diseases. Elevation of Vata, Pitta Kapha beyond an individual's threshold leads to manifestation of specific doshic disorders. Amongst the Vatic disorders developmental, neurological, dementia, movement and speech disorders, arrhythmias etc are described. In Pitta elevation, ulcer, bleeding disorders, skin diseases etc. and in Kapha obesity, diabetes, atherosclerotic conditions etc. are described. The perturbation of specific doshas in an individual is assessed through the symptoms and the aim of the Ayurveda physicians is to measure the amount of perturbation and bring back the doshas to his or her homeostatic state by appropriate dietary and therapeutic regime. Each of the food or medicine including lifestyle related things have been described to enhance or reduce a particular doshic state and therefore an individual-specific customized treatment is provided. Thus the beauty of Ayurveda lies in the fact that an individual, a disease condition, drug, diet as well as environment is described in terms of doshic components and appropriate customizations can be provided to balance these states.

From the foregoing description it is evident that tridoshas and *Prakriti* are the important basis of personalized tenets of Ayurveda for application in predictive medicine, it is imperative to establish their molecular basis. We felt that since Doshas are also explained to be elicited differently in different constitution types, it might be worthwhile to use modern genomic approaches to get a molecular explanation of the Tridosha concept by studying individuals who have predominance of the Doshas. Our basic working hypothesis was that if

TABLE: Distinguishing features of three contrasting *Prakriti* types *Vata, Pitta* and *Kapha* and their disease predisposition as described in the original text.

S.No	Features	*Vata*	*Pitta*	*Kapha*
1	Body frame	Thin	Medium	Broad
2	Body build and musculature	Weakly developed	Moderate	Well developed
3	Skin	Dry and cracked	Soft, thin, with tendency for moles, acne and freckles	Smooth and firm, clear complexion
4	Hair	Dry, thin, prone to breaks	Thin, oily, early graying	Thick, smooth and firm
5	Weight gain	Recalcitrant	Fluctuating	Tendency to obesity
6	Food and bowel habits	Frequent, variable and irregular	higher capacity for food and water consumption	Low digestive capacity and stable food habits
7	Movements and physical activities	Excessive and brisk		Less mobile
8	Tolerance for seasonal weather	Cold intolerant	Heat intolerant	Endurance for both
9	Disease resistance and healing capacity	Poor	Good	Excellent
10	Metabolism of toxic substances	Moderate	Quick	Poor
11	Communication	Talkative	Sharp, incisive communication with analytical abilities	Less vocal with good communication skills
12	Initiation capabilities	Quick, responsive and enthusiastic	Moderate, upon conviction and understanding	Slow to initiate new things
13	Memory	Quick at grasping and poor retention	Moderate grasping and retention	Slow grasping and Good at retention
14	Ageing	Fast	Moderate	Slow
15	Disease Predisposition/ Poor prognosis	Developmental, Neurological, dementia, movement and speech disorders, Arrhythmias	Ulcer, bleeding disorders, Skin diseases	Obesity, diabetes, atherosclerotic conditions

Ayurveda describes Vata to be the kinetic component, Kapha as a structural component and Pitta as a metabolic component, do individuals of different constitution types who are explained to be Vata, Pitta or Kapha types express different set of genes that governs these processes? (Figure 2) If this were to happen then one could identify such genes and take it further to see whether these genes confer differences in susceptibility to diseases or differences in drug responsiveness. Differences in HLA gene polymorphism between Prakriti group has been described earlier[11]. No studies at the genome-wide scale have however been attempted.

We adopted the Ayurvedic method of *Prakriti* analysis and identified normal healthy individuals belonging to the three extreme and contrasting *Prakriti* groups - Vata, Pitta, Kapha. In order to rule out effect of ethnicity related genetic variation confounding our study we carried out our Ayurveda study on individuals primarily of Indo-European origin (Figure 3). Our studies have shown that normal healthy individuals from the three most contrasting constitution types exhibit striking differences at the genome wide expression levels and biochemical and hematocrit parameters measured using peripheral blood[12]. For instance lipid profiles, uric acid, haemoglobin, blood clotting time serum zinc levels etc varied from one group to the other. At the expression level we also observed enrichment in core biological processes like transport, immune response, blood coagulation etc. There were correlations among biochemical profiles, functional categories of differentially expressed genes and the Ayurvedic descriptions as well. We observed higher levels of markers of metabolic syndrome and chronic inflammation (TG, total cholesterol, LDL, VLDL, High LDL/HDL, low HDL, uric acid, SGPT) in Kapha males compared to Vata and this was also consistent with over-expression of genes involved in inflammatory response in these individuals. Prothrombin time, indicative of blood coagulation process, was observed to be low in

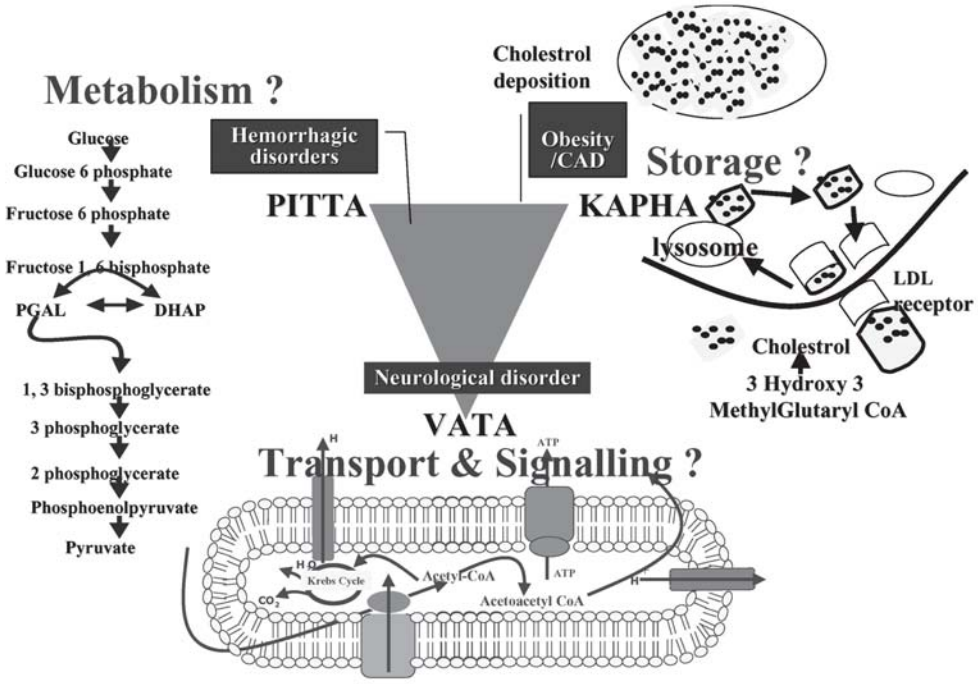

Figure 2. Working hypothesis of the study.

perturbation from the threshold. 30% of the entire data set of the genes that were differentially expressed among *Prakriti* groups were reported to be associated with complex and monogenic diseases. There were a significant number of hub genes (these genes interact with more than ten genes) which link complex diseases related to diabetes, immunological, infectious, cardiovascular, neuro-psychiatric disorders and cancer. Ayurveda based method of *Prakriti* classification thus allowed us for the first time to identify biochemical and expression differences amongst normal healthy individuals. We reasoned that identification of genetic variations which can be correlated to *Prakriti* phenotypes would facilitate predictive marker discovery.In order to test this hypothesis we studied genetic differences between a set of genes that were observed to be differentially expressed. To our surprise we observed significant differences in frequency of Kapha males. Further, higher levels of expression of haemoglobin genes in Pitta compared to Vata and Kapha also corroborates with the differences in haemoglobin levels between the *Prakriti*s and correlates with the redness of skin as a phenotype in Pitta individuals. Ayurveda proposes that the proportions of Doshas are restrained within allowable limits and disease is a consequence of

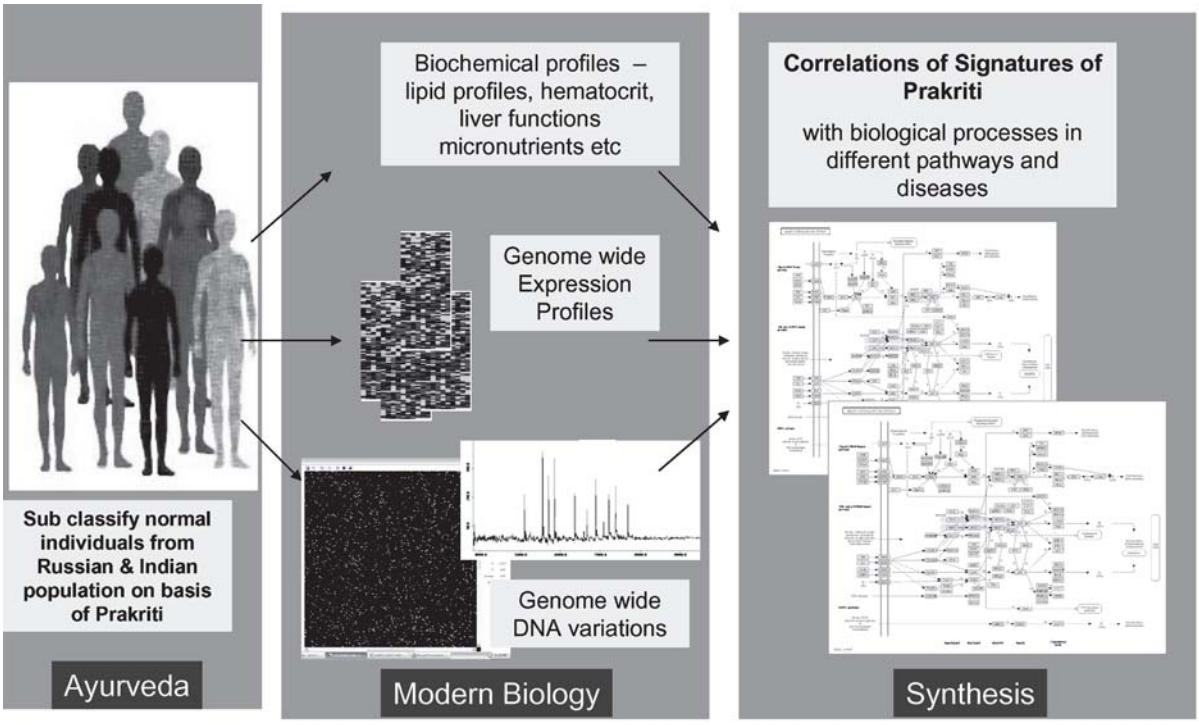

Figure 3. Genotype-Phenotype Correlation Based on the Principles of Ayurveda With Special Focus on Prakriti.

variations of a few of these genes among the three contrasting constitution types. Most importantly if all the individuals of contrasting constitution types were pooled together, the frequency of these genetic variations assumed an average value which was similar to that of the population of the same ethnic background. This highlights the importance of this method which allowed us to partition the common variations in some of the genes even within genetically homogeneous population. Since these constitution types are described to be differently predisposed to diseases, we ask the question, could these markers that were partitioned according to constitution types also serve as markers for differential predisposition? In order to test this further we followed a gene EGLN1 which is a key oxygen sensor that can switch on a subset of genes when required that allows a body to adapt to low oxygen conditions. In our dataset this gene differed both with respect to its expression level as well as at genetic level between *Pitta* and *Kapha* constitution types, and the expression differences were coretable to genetic variations. We therefore reasoned if these variations could be involved in the response of an individual to high oxygen or low oxygen conditions. One of the physiological conditions where oxygen levels are low is at high altitudes to which natives get acclimatized and often un-acclimated sojourners suffer from High Altitude Pulmonary Edema (HAPE). High altitude region, according to Ayurveda is considered as Kapha-Vata predominant region where disorders of a Kapha-Vata are more prevalent and *Pitta* was anticipated to have higher adaptive capacity. To our amazement we actually observed the *Pitta* genotype to be highly represented in natives of high altitude and that of the *Kapha* genotype in those individuals who develop HAPE. In order to substantiate our observation we further studied patterns of distribution of these variations across 24 Indian populations residing at different geographical locations. We observed that populations that reside in lower altitude but are genetically similar to the natives have a significantly lower frequency of the *Pitta* genotypes. In addition, we also observed the marker linked to high altitude adaptation in India to be conserved across different world populations residing at high altitude[13]. Thus we anticipate that individuals who are of the *Pitta Prakriti* or who have the marker linked to the high altitude phenotype may be able to perform better in high altitude conditions. Thus using the Ayurgenomics approach we could identify pathways and genes that differ at the expression level as well as genetic level between contrasting constitution types that are differently predisposed. Led by this observation we subsequently discovered a genetic marker in an oxygen sensing gene that confers differences in responsiveness to low oxygen (hypoxia) and was found to be responsible for high altitude adaptation[13]. A conventional approach wherein inherent molecular differences amongst healthy individuals.

Our studies have provided a molecular framework for Ayurgenomics and further studies need to be undertaken to understand the holistic concepts of this system of medicine. These studies have highlighted that an Ayurgenomics approach can accelerate/assist predictive marker discovery. World-over, different approaches are now being taken to find solutions for addressing the issue of "missing heritability" – the dark hole in the genetic studies. This has brought in a new era of system biology for holistic understanding of diseases and integrative approaches involving more omics (proteomics, transcriptomics, metabolomics) and intensive economics' are being adapted. In this context, validation of the tenets of Ayurveda and subsequent integration of Ayurveda phenotyping in complex diseases has the potential to be the least invasive and most affordable form of assessing an individual's susceptibility and prognosis. Further it could guide therapeutic and dietary recommendations. Presently we are trying to develop objective measures of the phenotyping concepts that would allow global standards to be set for different constitution types. So far Ayurveda has been principally been explored to identify active principles from herbs described in the treatment of different disease although a rigorous testing of the tenet on which these medicinal plants are described to be effective has not been attempted. Our efforts are towards establishing the molecular basis of the principles of Ayurveda in order to expand the scope of applicability and global acceptability of this branch of science.

Acknowledgements

We thank Prof. Samir K. Brahmachari for conception of Ayurgenomics and also the Indian Genome Variation project. Financial support from the Department of Science and Technology (DST) Grant B6.25 (to M.M.) to help initiate the project, and Council of Scientific and Industrial Research Grants CMM0016 and MLP3601 (to M.M.) is also acknowledged. ❐

References

1. JC Venter et al The sequence of the human genome. Science; **291**(5507):1304-51 (2001).
2. E Lander *et al* Initial sequencing and analysis of the human genome.Nature; **409** (6822):860-921 (2001).
3. The International HapMap Consortium. The International HapMap Project. *Nature* **18**; 426 (6968):789-96 (2003).

4. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs *Nature* **449**, 851-861 (2007).
5. LA Hindorff, P Sethupathy, HA Junkins, E M Ramos, J P Mehta, FS Collins and TA Manolio . Potential etiologic and functional implications of genome-wide association loci for human diseases and traits *Proc. Natl. Acad. Sci. USA* **106**: 9362-9367 (2009).
6. KA Frazer, SS Murray, NJ Schork, EJ Topol. Human genetic variation and its contribution to complex traits Nat. Rev. Genet. **10**: 241-251 (2009).
7. Caraka Samhita (Text with English translation). Chaukhamba Orientalia; (2000).
8. Susruta Samhita (Text with English translation). Chaukhamba Visvabharati; (2000).
9. MS Valiathan: The Legacy of Caraka. Orient Longman; (2003).
10. PV. Sharma *Caraka Samhita (Text with English translation)* (Chaukhamba Orientalia, Varanasi, India) (2000).
11. P Bhushan, J Kalpana and C Arvind Classification of human population based on HLA gene polymorphism and the concept of Prakriti in Ayurveda. *J. Altern. Complement Med*, **11**: 349-353 (2005).
12. B Prasher, S Negi, S Aggarwal, AK Mandal, TP Sethi, SR Deshmukh, SG Purohit, S Sengupta, S Khanna, F Mohammad, G Garg, SK Brahmachari, Indian Genome Variation Consortium, Mukerji M Whole genome expression and biochemical correlates of extreme constitutional types defined in Ayurveda. *J. Transl. Med.* **6:48**. :48 (2008)
13. S Aggarwal, S Negi, P Jha, PK Singh, T Stobdan, MA Pasha, S Ghosh, A Agrawal, Indian Genome Variation Consortium, Prasher B, Mukerji M *EGLN1* involvement in high altitude adaptation revealed through genetic analysis of extreme constitution types defined in Ayurveda *Proc. Natl. Acad. Sci. U S A.* 2010 Oct 18. (2010) [Epub ahead of print]

Website references

- The Human Genome Project (http://www.ornl.gov/sci/techresources/Human_Genome)
- The International HapMap Project http://hapmap.ncbi.nlm.nih.gov/
- A Catalog of Published Genome-Wide Association Studies. http:// www.genome.gov/gwastudies