

# Phenotype First or Genotype First: The Conundrum

Divya Agarwal<sup>1</sup>, Meenakshi Lallar<sup>2</sup>, Chaitanya Datar<sup>3</sup>, Koumudi Godbole<sup>4</sup>

<sup>1</sup> Apollo Indraprastha Hospitals, New Delhi, India

<sup>2</sup> Prime Diagnostics and Prenatal Imaging, Chandigarh, India & MedGenome Labs, Bengaluru, India

<sup>3</sup> Bharati Hospital, KEM Hospital, Surya Hospital and Aditya Birla Memorial Hospital, Pune, India

<sup>4</sup> Deenanath Mangeshkar Hospital and Research Center, Pune, India

Correspondence to: Dr Divya Agarwal Email: dr.divya2512@gmail.com

## Abstract

Evaluation of patients with suspected genetic disorders as well as gene disease research has predominantly relied on studying the signs and symptoms (phenotyping) of patients or research cohort, followed by doing appropriate genetic tests (genotyping) to diagnose the genetic condition or characterize the causative gene. As the costs of gene sequencing are going down, clinicians and researchers are evaluating the diagnostic yield, cost-benefit ratio, advantages and limitations of this approach versus genotyping all individuals without any phenotypic biases and correlating the genetic results with clinical features as a secondary step. In this article we put forward the debate on the traditional, tried and tested 'Phenotype first' approach versus the contrasting new hypothesis of 'Genotype first' approach for patient care and genomic research.

**Keywords:** Phenotype, genotype, sequencing, genetic disorders, genomic research.

## The Phenotype First Approach

*"Genotype without phenotype leads to missense or nonsense"* - Prof David Rimoin.

The **phenotype** (*phainein*: 'to show', and *typos*, meaning 'type') of an organism is the composite of the organism's observable characteristics or traits which would include the physical form and structure, developmental processes, biochemical and physiological properties, and behaviour, including the products of behaviour. Phenotype results from two basic factors: the expression of an organism's genetic code or its genotype, along with the influence of the environment. The

**phenome** refers to the set of all phenotypes expressed by the cell, tissue, organism, and species. **Phenomics** is the systematic study of phenotypes.

In simple terms, phenotyping is documentation of all our clinical examination findings and having some differential diagnosis before ordering any genetic test. In fact, the choice of test whether a karyotype or microarray or exome/ genome sequencing or any other molecular test would be guided by the phenotype. For example, if spinal muscular atrophy (SMA) is the clinical suspicion, then multiplex ligation-dependent probe amplification (MLPA) of the *SMN1* gene would be considered first, while if the phenotype is epilepsy, generally exome sequencing would be the first line genetic test.

Phenotyping is done in various ways that include:

- External evaluation by clinical dysmorphology review, fetal autopsy, etc.
- Internal examination including evaluation for organomegaly, cardiac signs, neuromuscular examination, fundus examination, etc.
- Investigation findings such as tissue histopathology, skeletal surveys/ radiographs, magnetic resonance imaging (MRI)/ computed tomography (CT) scans, electromyography/ nerve conduction studies (EMG/NCS), routine and special biochemical investigations, ultrasound findings etc.

All the above would help in reaching a clinical differential diagnosis based on which the laboratory will search for relevant genes of significance. In fact, a professional genetic

laboratory would ask for detailed phenotypic data before initiating a genomic test.

With the advent and rapid progress of next-generation sequencing (NGS) it had been suggested that NGS will decrease the need for phenotyping in general. However, it is now clear that a phenotype-driven approach is necessary to decipher various genotypes. In fact, a gene is listed in the Online Mendelian Inheritance in Man portal (OMIM) as significant only if it has a defined phenotype. It may be remembered that a correct phenotype will more likely provide you with a correct genotype.

The utility of phenotyping is dual in the current NGS era: a pre-NGS differential diagnostic mode (forward phenotyping) and post-NGS diagnostic assessment mode (reverse phenotyping). Both these are very critical for drawing any meaningful conclusions out of the genetic results.

The points in favour of a 'phenotype first' approach are as follows:

- It helps in more accurate search for genes (as relevant to the clinical examination findings) in the primary investigation step. Further data reanalysis can be guided by pointers based on the phenotype evolution as per the age and natural history of the disorder.
- One cannot rely on the laboratories entirely to give a genetic diagnosis. Providing proper phenotype handles would minimize errors from genome analysts or bioinformaticians who are generally non-clinical personnel. In fact, the Human Phenotype Ontology (HPO) has been developed for the reason that computerised NGS analysis should include the accurate phenotype.
- Phenotype blends and causes for variability in presentation and severity (due to reduced penetrance, variable expressivity, environmental or epigenetic interferences) can only be dissected by phenotyping.
- Planning immediate management for inborn errors of metabolism or prognostication regarding the severity of the condition is enabled by various examination and investigation findings and do not always depend on the genotype results.
- Phenotyping enables assigning significance to variants of uncertain significance (VUS) and sometimes downgrading pathogenic

variants (deemed pathogenic by only the in-silico predictions tools or as per the available literature from other populations).

- A proper phenotyping based on examination and investigations will help in deciding on specific gene panels which can save costs and time. Some phenotypes may not need costly NGS-based testing and confirmation of diagnosis may be possible by a simple and less costly targeted test.
- A phenotype-driven approach (with pretest and post-test counselling) will help to reduce the psychosocial anxiety associated with the condition for the patient and reduce medicolegal liabilities for the clinician and the laboratory.

As rightly pointed out by Hennekam et al. (2012), "there will be a critical need for phenotyping and clinical analysis and Medical Geneticists are uniquely positioned to address the need".

## The Genotype First Approach

Clinicians have learnt and practiced medicine in the order of history, examination, basic investigations, and advanced investigations.

The '**genotype first**' approach is the process in which the patient/ individual undergoes genomic testing/sequencing with subsequent determination of the associated phenotypes of interest. Radical scientists and researchers believe that the 'genotype first' approach has the potential to take genomic medicine beyond ascertainment biases and can truly take medicine and health towards the prevention of all disease (Wilczewski et al., 2023). At present there are no formal recommendations, but we are gradually moving towards this change in clinical, laboratory as well as research settings.

### In the clinical setting

A 'genotype first' approach for patients with suspicion of a genetic disorder, will optimize the health care system capacity. This is particularly true in our part of the world, where there is insufficient medical genetics expertise and no well-defined referral system. Majority of genetic tests are ordered by non-specialist physicians with information about only the basic clinical symptoms of the patient. As the cost of genotyping is going down, this practice will further increase.

It is only when the patient reaches the specialist/ geneticist with the genetic report, that reverse deep phenotyping, segregation studies and reanalysis of the genotyping data are done. Inadvertently, the 'genotype first' approach is already being followed for clinical diagnosis. Also, in our setting, 'genotype first' is more relevant as it reduces the time to diagnosis and is cost effective. In-depth phenotyping like cerebrospinal fluid (CSF) studies, magnetic resonance spectroscopy (MRS), repeated MRIs to look for myelination abnormalities, biopsy and special staining, etc. are cost-prohibitive. If we add up the cost of patient visits back and forth for clinical assessment, then for the reports, and then for advanced tests vs patient getting the genotyping in the first visit, the latter approach has a more favourable cost-benefit ratio.

Even if we leave costly investigations out of the equation, phenotyping by only clinical examination has inherent flaws. First, it is subjective and dynamic. The presence or absence of neurologic signs, dysmorphic features, skin findings etc. could be subjective and transient which may bias the analysis of genotyping based on the phenotype. With expanding genetic knowledge, we know that classic phenotypes of genetic syndromes can be seen only in a subset of patients. There is random combination of symptoms, subtle symptoms, and new symptoms being described for genetic syndromes. Genetic analysis based on the clinical differentials will be inaccurate in all patients who do not present with classic phenotypes. This is especially true in the fetal / prenatal setting where clinical details gathered from fetal ultrasound are operator-driven or may be easily missed due to their transient nature.

### **In the laboratory setting**

In the 'genotype first' approach, genes are fully sequenced, and all of the thousands of variants are carefully evaluated for properties that make them more likely to be disease-causing. The variants are not filtered out of the analysis at the first step based on the patient's phenotype. During reporting, variants are compared to the patient's phenotype to see if they explain all, or part, of the phenotype. This allows for identification of variants in patients with atypical or rarely reported presentations and has the potential to diagnose more than one condition. It also allows for the identification of suspicious variants in genes where a disease association is not yet established or only newly described.

In the approach where variants are annotated, classified, and reported only on the basis of symptoms, it is difficult to uncouple the variant classification and reporting in spite of them having evidence of being disease-causing. This will indeed lead to more variants of uncertain significance.

### **In the research setting**

Genotyping of a cohort of individuals followed by reverse phenotyping can help in identification of new causative genotypes. Much of research in autism and neurodevelopmental delay has taken this approach with a good yield. Also, public datasets of NGS have been analyzed for different phenotypes including actionable germline cancer variants, and cardiovascular phenotypes like connective tissue disorders and RASopathies.

Advantages of the 'genotype first' approach in research settings are new gene discovery, new gene-phenotype correlation, better genotype-phenotype correlation, and characterization of background modifiers causing variable expression and penetrance. A novel genotype disease association can never be established by phenotype ascertainment bias.

Genotype-phenotype associations are limited in known genes when researchers select participants strictly based on the phenotype. For example, severe metabolic derangements like insulin resistance, severe diabetes, end organ complications due to *LMNA* gene mutation without the typical signs and symptoms like lipoatrophy could be established only by the 'genotype first' approach (Decaudain et al., 2007).

### **In population research**

The 'genotype first' approach has the potential to shift from reactive medicine to preventive medicine. Multiple studies have demonstrated that population genotyping can identify pathogenic/ likely pathogenic variants in genes related to adult-onset conditions particularly unmasking the risk of malignancies like *BRCA* genes, *NF1*, etc. (Safonov et al., 2023).

## **Conclusion**

The 'genotype first' approach is gaining momentum due to several factors like wider and more accessible NGS, evolving artificial intelligence (AI) tools, public genomic databases, and availability of electronic medical records (EMRs). With technological evolution 'genotype first' is a rapidly advancing approach for genomic research avoiding the phenotypic ascertainment

(Wenger et al., 2021). However, in clinical settings, it comes with various challenges like absence of a single genomic test for detection of all type of genetic variations and all genetic disorders, along with management and counselling of patients with genetic findings and no phenotypic expression of these findings. Overall if we move towards a time where universal genome sequencing is offered as part of routine health care, we should be well versed with the complexities of interpretation of genotyping and of genotypic expression (phenotype) (Bodian et al., 2016).

The debate will continue but genotyping and phenotyping are like two legs taking an individual forward; sometimes one puts the right foot first and sometimes left! Judicious use of the right diagnostic technique and at the right time requires wisdom. Astute clinicians will understand the power of both and are more likely to make the correct diagnosis in less time and at a lower cost. In addition to phenotyping skills and the knowledge of genetic disorders, clinicians of the genomic era need to be empowered with the skill to negotiate the maze of databases to solve the diagnostic conundrum.

**Conflict of interests:** None

## References

1. Bodian DL, et al. Utility of whole-genome sequencing for detection of newborn screening disorders in a population cohort of 1, 696 neonates. *Genet Med.* 2016; 18: 221–230.
2. Decaudo A, et al. New metabolic phenotypes in laminopathies: LMNA mutations in patients with severe metabolic syndrome. *J Clin Endocrinol Metab.* 2007; 92: 4835–4844.
3. Hennekam RCM, et al. Next generation sequencing demands next generation phenotyping. *Hum Mutat.* 2012; 33: 884–886.
4. Safonov A, et al. A genotype-first approach identifies high incidence of NF1 pathogenic variants with distinct disease associations. *medRxiv [Preprint].* 2023 Aug 10:2023.08.08.23293676. doi: 10.1101/2023.08.08.23293676.
5. Wenger BM, et al. A genotype-first approach to exploring Mendelian cardiovascular traits with clear external manifestations. *Genet Med.* 2021; 23:94–102.
6. Wilczewski CM, et al. Genotype first: Clinical genomics research through a reverse phenotyping approach. *Am J Hum Genet.* 2023; 110: 3–12.

