

Editorial

# Towards Accurate Genotype–Phenotype Correlations in the CYP2D6 Gene

Angel L. Pey 

Departamento de Química Física, Unidad de Excelencia de Química aplicada a Biomedicina y Medioambiente, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain; angelpey@ugr.es

Received: 28 September 2020; Accepted: 1 October 2020; Published: 8 October 2020



**Abstract:** Establishing accurate and large-scale genotype–phenotype correlations and predictions of individual response to pharmacological treatments are two of the *holy grails* of Personalized Medicine. These tasks are challenging and require an integrated knowledge of the complex processes that regulate gene expression and, ultimately, protein functionality *in vivo*, the effects of mutations/polymorphisms and the different sources of interindividual phenotypic variability. A remarkable example of our advances in these challenging tasks is the highly polymorphic CYP2D6 gene, which encodes a cytochrome P450 enzyme involved in the metabolism of many of the most marketed drugs (including SARS-Cov-2 therapies such as hydroxychloroquine). Since the introduction of simple activity scores (AS) over 10 years ago, its ability to establish genotype–phenotype correlations on the drug metabolizing capacity of this enzyme in human population has provided lessons that will help to improve this type of score for this, and likely many other human genes and proteins. Multidisciplinary research emerges as the best approach to incorporate additional concepts to refine and improve such functional/activity scores for the CYP2D6 gene, as well as for many other human genes associated with simple and complex genetic diseases.

**Keywords:** drug metabolism; polymorphism; genotype–phenotype correlations; cytochrome P450

---

The application of next-generation sequencing technologies is revealing the existence of a vast number of genetic variants in human population, although our ability to reliably identify and quantify the phenotypic consequences and potential pathogenicity of such variants is still quite limited [1–7]. In some cases, such as non-sense and frame-shift mutations, large deletions and insertions, or duplications, the phenotypic consequences and pathogenicity of these mutations can be inferred more or less straightforwardly. However, and not surprisingly, due to the lower output of functional analyses compared to that of large-scale new-generation sequencing analyses, about half of the currently identified genetic variants are classified as of uncertain significance [2]. This interpretation is further challenged by the large number (about 50%) of these variants identified in only a single individual [8].

A particularly challenging case is the evaluation of the phenotypic consequences of missense mutations. Good correlations between the effects of missense mutations on molecular function and their pathogenic manifestations are sometimes observed, although these correlations are complex in nature (life is intrinsically complex) [1–3,9,10]. At the simplest level, a missense mutation can simultaneously affect different functional traits in a given protein, including enzyme kinetics, intracellular stability, subcellular targeting, aggregation propensity and interactions with small molecules and other biomacromolecules [3,9,11,12]. Even at this simple level, the characterization of these molecular effects and their correlation with a given molecular or cellular genotype is challenging, even when protein structural information is available [9]. Remarkably, a good correlation between the effect of missense mutations on protein structural stability, the levels of functional protein in the cell and particular loss-of-function phenotypes has been drawn in some cases [3,7,13,14]. The phenotypic manifestation of

mutational effects on molecular functionality also depends on many other factors, such as the genetic and cellular context, environmental factors, stochastic variations, transgenerational transmission of epigenetic information and post-translational modifications (PTMs) [2,3,15–18]. A paradigmatic example of this mechanistic complexity are those isogenic individuals who manifest widely different disease susceptibilities [19,20].

The *CYP2D6* gene is a very well-documented case in which large genetic variability is known to affect patient's phenotype in a complex manner [21]. This gene encodes for the cytochrome P450 2D6 monooxygenase, an enzyme involved in the metabolism of a wide variety of biomolecules (such as lipids) as well as several of the most marketed drugs, including SARS-CoV-2 therapies (e.g., chloroquine and hydroxychloroquine) [22–25]. This gene displays a large genetic diversity among human population, including non-sense, frame-shift and missense mutations and gene duplications [21,26,27]. Despite the inherent complexity in establishing genotype–phenotype correlations in the *CYP2D6* gene, a simple phenotypic score was proposed in 2008 that qualitatively ranked the phenotype (i.e., poor, intermediate, extensive-normal or ultrarapid metabolizers) by assigning a given value for each allele, including activity enhancement due to gene duplications [28]. As these authors recently reviewed, this score has become popular and rather successful, although its simplicity provides an inherent lack of (quantitative) predictive power. This limitation is highlighted, for instance, by the strong interindividual differences in the phenotype within a given genotype or the variable phenotypic manifestation, depending on the drug investigated [21]. This has led to additional factors being proposed that may improve the performance of such a predictive tool for genotype–phenotype correlations in the *CYP2D6* gene such as the presence of alterations (e.g., polymorphisms) in DNA regulatory elements, modulation of *CYP2D6* gene expression by changes in the activity of transcription factors and small, non-coding RNAs, competing effects of substrates and inhibitors of the enzyme and many others [21]. Several additional factors may contribute to improving genotype–phenotype correlations for the *CYP2D6* gene, such as considering the effects of missense mutations on protein structure, energetics and dynamics, or the potential regulatory roles of post-translational modifications (PTMs). First, missense mutations may affect the structure and stability of the protein, leading to changes in the intracellular stability and specific activity. To implement these effects into an activity score for the *CYP2D6* gene, currently available crystal structures for the protein (over 25 are available at the *Protein Data Bank*, including those with bound substrates and inhibitors) [24,29] and suitable structure-based energy force fields could be used for either the prediction of deleterious effects of missense variants detected in human population or for *in silico* saturation mutagenesis to identify regions that are particularly sensitive to these mutational effects (i.e., hot-spots) [3,7,14,30]. These analyses can be cross-correlated with functional information from expression analysis combined with biochemical and biophysical studies of mutant variants [31]. Second, detailed structural analyses on the *CYP2D6* enzyme have supported the existence of a certain level of structural malleability of the protein in adapting its conformation specifically to the presence of substrates and inhibitors [24]. Thus, it is plausible that missense mutations may affect this dynamic behavior of the protein, causing subsequent functional (i.e., enzymatic) alterations. Regarding PTMs, there are currently 10 sites of PTMs identified in the *CYP2D6* enzyme by proteomic analyses, although no information is available for the effects of these site-specific PTMs on the function and stability of the protein [32,33]. Six of these correspond to ubiquitinylation events, with potential effects on the stability or regulation of the protein, whereas three of them correspond to phosphorylation sites and one to an acetylation event. Therefore, it is plausible that the complex dynamics underlying these PTMs could lead to different phenotypic manifestations of a given phenotype in different individuals.

Clearly, an increased knowledge of the mechanisms underlying genotype–phenotype correlations associated with the genetic variability of the *CYP2D6* will improve our predictive power on the metabolizing function of *CYP2D6*. The lessons to be learned from this gene could be more generally useful for the accurate prediction of these genotype–phenotype correlations (particularly for missense mutations) on a genome-wide scale, and the implications for individual response to pharmacological treatments, one of the holy grails of Personalized Medicine.

**Funding:** ALP is funded by the ERDF/Spanish Ministry of Science, Innovation and Universities—State Research Agency (Grant RTI2018-096246-B-I00) and Consejería de Economía, Conocimiento, Empresas y Universidad, Junta de Andalucía (Grants P18-RT-2413).

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Shendure, J.; Akey, J.M. The origins, determinants, and consequences of human mutations. *Science* **2015**, *349*, 1478–1483. [[CrossRef](#)] [[PubMed](#)]
2. Manolio, T.A.; Fowler, D.M.; Starita, L.M.; Haendel, M.A.; MacArthur, D.G.; Biesecker, L.G.; Worthey, E.; Chisholm, R.L.; Green, E.D.; Jacob, H.J.; et al. Bedside Back to Bench: Building Bridges between Basic and Clinical Genomic Research. *Cell* **2017**, *169*, 6–12. [[CrossRef](#)] [[PubMed](#)]
3. Stein, A.; Fowler, D.M.; Hartmann-Petersen, R.; Lindorff-Larsen, K. Biophysical and Mechanistic Models for Disease-Causing Protein Variants. *Trends Biochem. Sci.* **2019**, *44*, 475–488. [[CrossRef](#)] [[PubMed](#)]
4. Daneshjou, R.; Wang, Y.; Bromberg, Y.; Bovo, S.; Martelli, P.L.; Babbi, G.; Lena, P.D.; Casadio, R.; Edwards, M.; Gifford, D.; et al. Working toward precision medicine: Predicting phenotypes from exomes in the Critical Assessment of Genome Interpretation (CAGI) challenges. *Hum. Mutat.* **2017**, *38*, 1182–1192. [[CrossRef](#)] [[PubMed](#)]
5. Hoskins, R.A.; Repo, S.; Barsky, D.; Andreoletti, G.; Moulton, J.; Brenner, S.E. Reports from CAGI: The Critical Assessment of Genome Interpretation. *Hum. Mutat.* **2017**, *38*, 1039–1041. [[CrossRef](#)] [[PubMed](#)]
6. Andreoletti, G.; Pal, L.R.; Moulton, J.; Brenner, S.E. Reports from the fifth edition of CAGI: The Critical Assessment of Genome Interpretation. *Hum. Mutat.* **2019**, *40*, 1197–1201. [[CrossRef](#)] [[PubMed](#)]
7. Pey, A.L.; Stricher, F.; Serrano, L.; Martinez, A. Predicted effects of missense mutations on native-state stability account for phenotypic outcome in phenylketonuria, a paradigm of misfolding diseases. *Am. J. Hum. Genet.* **2007**, *81*, 1006–1024. [[CrossRef](#)] [[PubMed](#)]
8. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.H.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [[CrossRef](#)]
9. Medina-Carmona, E.; Betancor-Fernández, I.; Santos, J.; Mesa-Torres, N.; Grottelli, S.; Batlle, C.; Naganathan, A.N.; Oppici, O.; Cellini, B.; Ventura, S.; et al. Insight into the specificity and severity of pathogenic mechanisms associated with missense mutations through experimental and structural perturbation analyses. *Hum. Mol. Genet.* **2019**, *28*, 1–15. [[CrossRef](#)]
10. Monteiro, C.; Mesgazardeh, J.S.; Anselmo, J.; Fernandes, J.; Novais, M.; Rodrigues, C.; Brighty, G.J.; Powers, D.L.; Powers, E.T.; Coelho, T.; et al. Predictive model of response to tafamidis in hereditary ATTR polyneuropathy. *JCI Insight* **2019**, *4*, e126526. [[CrossRef](#)]
11. Mesa-Torres, N.; Betancor-Fernández, I.; Oppici, E.; Cellini, B.; Salido, E.; Pey, A.L. Evolutionary Divergent Suppressor Mutations in Conformational Diseases. *Genes* **2018**, *9*, 352. [[CrossRef](#)] [[PubMed](#)]
12. Pey, A.L. Biophysical and functional perturbation analyses at cancer-associated P187 and K240 sites of the multifunctional NADP(H):quinone oxidoreductase 1. *Int. J. Biol. Macromol.* **2018**, *118*, 1912–1923. [[CrossRef](#)] [[PubMed](#)]
13. Nielsen, S.V.; Stein, A.; Dinitzen, A.B.; Papaleo, E.; Tatham, M.H.; Poulsen, E.G.; Kasseem, M.M.; Rasmussen, L.J.; Lindorff-Larsen, K.; Hartmann-Petersen, R. Predicting the impact of Lynch syndrome-causing missense mutations from structural calculations. *PLoS Genet.* **2017**, *13*, e1006739. [[CrossRef](#)] [[PubMed](#)]
14. Scheller, R.; Stein, A.; Nielsen, S.V.; Marin, F.I.; Gerdes, A.M.; Marco, M.D.; Papaleo, E.; Lindorff-Larsen, K.; Hartmann-Petersen, R. Towards mechanistic models for genotype-phenotype correlations in phenylketonuria using protein stability calculations. *Hum. Mutat.* **2019**, *40*, 444–457. [[CrossRef](#)] [[PubMed](#)]
15. Lehner, B. Genotype to phenotype: Lessons from model organisms for human genetics. *Nat. Rev. Genet.* **2013**, *14*, 168–178. [[CrossRef](#)]
16. Klosin, A.; Casas, E.; Hidalgo-Carcedo, C.; Vavouri, T.; Lehner, B. Transgenerational transmission of environmental information in *C. elegans*. *Science* **2017**, *356*, 320–323. [[CrossRef](#)]
17. Dhar, R.; Missarova, A.M.; Lehner, B.; Carey, L.B. Single cell functional genomics reveals the importance of mitochondria in cell-to-cell phenotypic variation. *Elife* **2019**, *8*, e38904. [[CrossRef](#)]

18. Fernandez-Higuero, J.A.; Betancor-Fernandez, I.; Mesa-Torres, N.; Muga, A.; Salido, E.; Pey, A.L. Structural and functional insights on the roles of molecular chaperones in the mistargeting and aggregation phenotypes associated with primary hyperoxaluria type I. *Adv. Protein Chem. Struct. Biol.* **2019**, *114*, 119–152.
19. Burga, A.; Lehner, B. Predicting phenotypic variation from genotypes, phenotypes and a combination of the two. *Curr. Opin. Biotechnol.* **2013**, *24*, 803–809. [[CrossRef](#)]
20. Casanueva, M.O.; Burga, A.; Lehner, B. Fitness trade-offs and environmentally induced mutation buffering in isogenic *C. elegans*. *Science* **2012**, *335*, 82–85. [[CrossRef](#)]
21. Gaedigk, A.; Dinh, J.C.; Jeong, H.; Prasad, B.; Leeder, J.S. Ten Years' Experience with the CYP2D6 Activity Score: A Perspective on Future Investigations to Improve Clinical Predictions for Precision Therapeutics. *J. Pers. Med.* **2018**, *8*, 15. [[CrossRef](#)] [[PubMed](#)]
22. Wang, B.Y.L.; Zhang, X.Z.; Huang, S.Q.; Bartlam, M.; Zhou, S.F. New insights into the structural characteristics and functional relevance of the human cytochrome P450 2D6 enzyme. *Drug Metab. Rev.* **2009**, *41*, 573–643. [[CrossRef](#)] [[PubMed](#)]
23. Paniri, A.; Hosseini, M.M.; Rasoulinejad, A.; Akhavan-Niaki, H. Molecular effects and retinopathy induced by hydroxychloroquine during SARS-CoV-2 therapy: Role of CYP450 isoforms and epigenetic modulations. *Eur. J. Pharmacol.* **2020**, *886*, 173454. [[CrossRef](#)] [[PubMed](#)]
24. Wang, A.S.C.; Zhang, Q.; Johnson, E.F. Contributions of ionic interactions and protein dynamics to cytochrome P450 2D6 (CYP2D6) substrate and inhibitor binding. *J. Biol. Chem.* **2015**, *290*, 5092–5104. [[CrossRef](#)] [[PubMed](#)]
25. UniProt. Available online: <https://http://www.uniprot.org/uniprot/P10635> (accessed on 27 September 2020).
26. OMIM<sup>®</sup>—Online Mendelian Inheritance in Man<sup>®</sup>. Available online: <https://http://www.omim.org/entry/124030> (accessed on 26 September 2020).
27. GnomAD Browser. Available online: <https://gnomad.broadinstitute.org/gene/ENSG00000100197> (accessed on 26 September 2020).
28. Gaedigk, A.; Simon, S.D.; Pearce, R.E.; Bradford, L.D.; Kennedy, M.J.; Leeder, J.S. The CYP2D6 activity score: Translating genotype information into a qualitative measure of phenotype. *Clin. Pharmacol. Ther.* **2008**, *83*, 234–242. [[CrossRef](#)] [[PubMed](#)]
29. Protein Data Bank. Available online: <https://http://www.rcsb.org/> (accessed on 27 September 2020).
30. Abildgaard, A.B.; Stein, A.; Nielsen, S.V.; Schultz-Knudsen, K.; Papaleo, E.; Shrikhande, A.; Hoffmann, E.R.; Bernstein, I.; Gerdes, A.M.; Takahashi, M.; et al. Computational and cellular studies reveal structural destabilization and degradation of MLH1 variants in Lynch syndrome. *Elife* **2019**, *8*, e49138. [[CrossRef](#)] [[PubMed](#)]
31. Dalton, R.; Lee, S.B.; Claw, K.G.; Prasad, B.; Phillips, B.R.; Shen, D.D.; Wong, L.H.; Fade, M.; McDonald, M.G.; Dunham, M.J.; et al. Interrogation of CYP2D6 Structural Variant Alleles Improves the Correlation Between CYP2D6 Genotype and CYP2D6-Mediated Metabolic Activity. *Clin. Transl. Sci.* **2020**, *13*, 147–156. [[CrossRef](#)]
32. PhosphositePlus<sup>®</sup>. Available online: <https://http://www.phosphosite.org/> (accessed on 27 September 2020).
33. Hornbeck, P.V.; Kornhauser, J.M.; Latham, V.; Murray, B.; Nandhikonda, V.; Nord, A.; Skrzypek, E.; Wheeler, T.; Zhang, B.; Gnad, F. 15 years of PhosphoSitePlus(R): Integrating post-translationally modified sites, disease variants and isoforms. *Nucleic Acids Res.* **2019**, *47*, D433–D441. [[CrossRef](#)]

