Monogenic traits are not simple

lessons from phenylketonuria

The classification of genetic disease into chromosomal, monogenic and multifactorial categories is an oversimplification. Phenylketonuria (PKU) is a classic 'monogenic' autosomal recessive disease in which mutation at the human *PAH* locus was deemed sufficient to explain the impaired function of the enzyme phenylalanine hydroxylase (enzymic phenotype), the attendant hyperphenylalaninemia (metabolic phenotype) and the resultant mental retardation (cognitive phenotype). In the era of molecular genetics, expectations for a consistently close correlation between the mutant genotype and variant phenotype have been somewhat disappointed, and PKU is used here to illustrate how and why this might be the case. So-called monogenic traits do, indeed, conform to long-accepted ideas about the expression of 'major' loci and their importance in determining parameters of phenotype, but the associated features are as complex, in their own ways, as those in so-called complex traits.

Science can be considered an assault on ignorance and its legacies are arrays of concepts, databases and technologies[1]. The genome projects are some of its major contemporary manifestations, and from them flows information that can be used to understand the variant phenotypes and diseases as they are being catalogued, for example, in a knowledgebase for the human genome[2] (http://www.ncbi.nih.gov/omim). The McKusick Catalog, which can be electronically linked with other genomic databases[3], classifies phenotype by inheritance, implying that a major locus harbouring mutation is a cause of the particular variant phenotype. It was a hope that delineation of genotypes with new methods for the detection of mutations[4,5] would enable the prediction of variant phenotypes; in the case of human genetic disease, this would have added value for prognosis and treatment. This was the point of view held, for example, in an editorial[6] accompanying a celebrated article[7] relating the long-recognized mendelian phenotype of hyperphenylalaninemia (HPA; see Box 1), in the disease phenylketonuria (PKU), to the newly identified mutations at the human *PAH* locus. HPA is one of the most widely ascertained inherited metabolic traits in human society[8].

The traditional classification of genetic disease into chromosomal, monogenic (mendelian) and multifactorial categories is an oversimplification. Here, we use the classic autosomal recessive disease PKU (MIM 261600) to show that even the category of monogenic or 'simple' traits is an artificial division. The evolution towards seeing single-gene traits as versions of complex traits has been under way for some time, as illustrated by: (1) widely different phenotypes that can be accounted for by allelic variation in a single gene[9]; (2) the blurring of predicted relationships between genotype and phenotype in several monogenic disorders[10]; (3) modifier genes and non-genetic factors that contribute to the phenotypes of monogenic disorders (Fig. 1). The complexity of the relationships between genotype and phenotype is illustrated by three of the monogenic diseases mentioned in Fig. 1. (1) Cystic fibrosis (MIM 219700), where both allelic variation at the *CFTR* locus and the expression of an important modifier locus should be taken into account to explain the difference between the pancreatic-sufficient and insufficient forms of the disease, and patency and occlusion of the vas deferens[11]. (2) The Hartnup phenotype (MIM 234500), where, in addition to mutation at the major locus, which accounts for the disorder of amino acid transport in kidney and intestine, there is also evidence for a multifactorial threshold phenomenon that affects the overall homeostasis of plasma amino acid levels. The latter accounts primarily for the difference between a patient having only the variant Hartnup amino acid transport phenotype (a biochemical trait) and having Hartnup disease, a pellagra-like clinical entity[12]. (3) The

BOX 1. Glossary

HPA
Hyperphenylalaninemia

PAH
Phenylalanine hydroxylase enzyme (phenylalanine 4-monooxygenase, EC 1.14.16.1)

PAH
Human phenylalanine hydroxylase gene (cDNA sequence, GenBank U49897)

PKU
Phenylketonuria (MIM 261600)

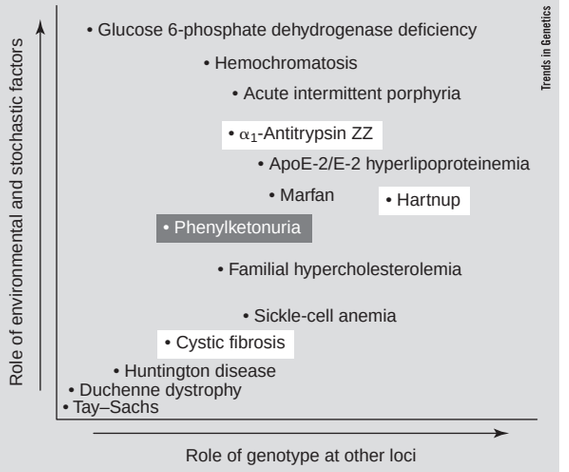
Charles R. Scriver
mc77@musica.mcgill.ca

Paula J. Waters
cywt@musica.mcgill.ca

DeBelle Laboratory for Biochemical Genetics, Montreal Children's Hospital, 2300 Tupper Street, Montreal, Quebec, Canada H3H 1P3.

0168-9525/99/$ – see front matter PII: S0168-9525(99)01761-8

FIGURE 1. Genotype–phenotype relationship



The potential influence of genotype and non-genetic factors on the phenotype of various monogenic disorders. The equation $V_P = V_G + V_E$ (Ref. 61) implies that variation in genotype and environment contribute to variation in phenotype. The diagram indicates the estimated relative importance of background genotype and environment as contributors to the phenotype of several monogenic diseases. The figure does not include the effect of allelic variation at the major locus. Boxed entries are mentioned in the text. Adapted from Ref. 62.

PiZZ form of α-1 antitrypsin deficiency, where some of the variation in the clinical phenotype involving the lungs is explained by environmental factors (e.g. smoking), whereas interindividual differences involving the liver are related to the handling of the mutant protein by chaperones and proteases[13,14]. Accordingly, it should be no surprise that PKU can be seen to behave as a 'complex' trait when considered at its cognitive and metabolic (hyperphenylalaninemia) levels of phenotype, which are beyond the control of the *PAH* locus by itself; nor that various cellular entities can interact directly with the phenylalanine hydroxylase protein (PAH; phenylalanine 4-monooxygenase, EC 1.14.16.1) and, thus, modulate the enzymic phenotype.

Understanding PKU

This has been a long journey of discovery[8,15]. Folling reported his discovery of the disease, now known as PKU, in the 1930s (Ref. 16); he recognized that its effect on cognitive development (mental retardation) and the occurrence of a metabolic marker (HPA) were explained by autosomal recessive inheritance. In the 1950s, a deficiency of hepatic cytosolic PAH enzyme activity was demonstrated in PKU patients. The homotetrameric PAH enzyme catalyses the conversion of phenylalanine to tyrosine, this hydroxylating reaction being the major determinant of the catabolic flux leading to the complete oxidation of phenylalanine to carbon dioxide and water; normally, this pathway accounts for at least 75% of the disposal of dietary phenylalanine. Also in the 1950s, an artificial diet that restricts the intake of the essential nutrient phenylalanine was shown to prevent HPA and allow normal cognitive development. In the 1960s, a simple screening test to detect HPA in the newborn infant was applied to populations worldwide, providing early diagnosis and access to treatment of affected persons. In the 1970s, it was recognized that not all mendelian HPA was PKU; some forms were caused by disorders of synthesis or the recycling of a cofactor (tetrahydrobiopterin) that is essential for the catalytic function of PAH; thus, locus heterogeneity for HPA was identified. In the 1980s the human *PAH* gene (cDNA sequence, GenBank U49897) was mapped to human chromosome 12q24.1, cloned and some mutations identified. By 1999, over 400 mutant alleles have been detected, largely by members of the *PAH* Mutation Analysis Consortium[17], and documented in a locus-specific mutation database (http://www.mcgill.ca/pahdb)[18]. In the 1990s, *in vitro* expression analysis of many *PAH* alleles flourished[19]; several domains, including the catalytic core of the human enzyme[20,21], were crystallized and visualized at 2 Å resolution. There are also crystal structure data for the rat enzyme[63].

Thus, it had become obvious (Fig. 2) that 'mendelian' HPA (a core trait in PKU) is also 'multifactorial', being the result of causes both ultimate (a mutant genotype) and proximate (dietary phenylalanine). It reflects locus heterogeneity (in about 2% of propositi)[8] and it reflects allelic heterogeneity in the *PAH* gene. The allelic spectrum fits an emerging paradigm[22] that only a few of the many disease-causing alleles at human loci are prevalent, whereas the vast majority are rare. Meanwhile, two important challenges to the hypothesis that a *PAH* genotype would consistently predict a 'monogenic' phenotype had emerged: (1) patients, even sibs, sharing identical mutant *PAH* genotypes could have greatly different cognitive and metabolic phenotypes[23–25]; (2) there are many instances of discordance between the mutant *PAH* genotype, its predicted effect on enzyme function, and the associated metabolic phenotype[26,27]. We examine these challenges at the three levels of phenotype: cognitive, metabolic and enzymic.

Cognitive phenotype

Mental retardation, the major disease phenotype, is attributed to a toxic effect of excess phenylalanine on brain development and function, but not every proband with untreated PKU has impaired cognitive development[8]. IQ is a complex trait, many factors account for it and, as one would expect, IQ scores in untreated PKU sibs or patients do not always correlate closely with the predicted effect of their mutant *PAH* genotype upon PAH enzyme function[23,24,28]. Moreover, individuals can have very different cognitive phenotypes, even when their metabolic phenotype (HPA) is similar[29,30]. Such discordance between metabolic and cognitive phenotypes appears to be explained by differences in the function of the blood–brain barrier and the modulation of free phenylalanine content of the brain[29,30]. This implies that genotypes at loci that control the mediated transport of phenylalanine in the brain are among the factors that contribute to the cognitive phenotype in the HPA state.

Metabolic phenotype

Two metanalyses[26,27] have documented genotype–phenotype correlations in >1000 HPA patients. The untreated ambient blood phenylalanine level and the dietary tolerance for phenylalanine, each a measure of the metabolic phenotype, reveal discordant correlations in a significant number of individuals when they are related to the predicted effect of the mutant *PAH* genotype on enzyme function (see below). Metabolic homeostasis of the size of the phenylalanine pool and control of its metabolic fluxes are

complex processes[8,31], involving the intestinal absorption and hepatic uptake of dietary phenylalanine, its incorporation into proteins and disposal by hydroxylation, transamination and decarboxylation (Fig. 3). The control of metabolic homeostasis is distributed as a systemic property of the overall network of reactions[32]. Differences in the disposal of excess phenylalanine by transamination when the hydroxylation reaction is blocked explain the differences between two PKU sibs, with identical mutant *PAH* genotypes, in their tolerance to dietary phenylalanine[25]. Another factor, as yet unmeasured in human subjects but analogous to the transport of phenylalanine into the brain, could be variation in the carrier-mediated uptake of the amino acid by liver; in the rat, this step has a large sensitivity coefficient in the overall homeostasis of extracellular phenylalanine[33].

Enzymic phenotype

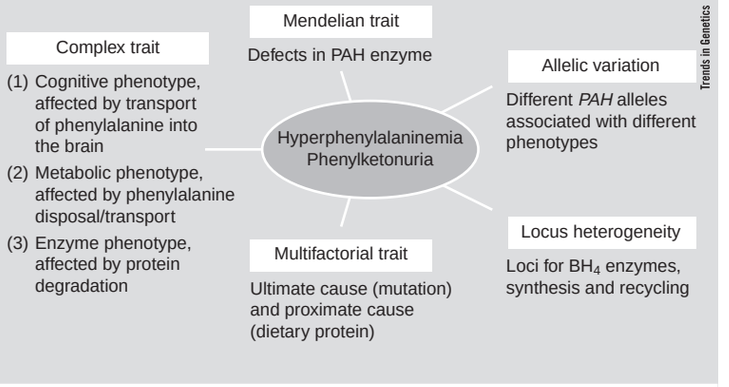
Is the PKU phenotype also complex at the enzymic level?

A compelling case for this would first require evidence that patients with identical *PAH* genotypes possess different phenylalanine hydroxylation activities *in vivo*. Significant differences are seen in *in vivo* phenylalanine oxidation between heterozygotes carrying the same mutant *PAH* allele[34]. In this case, the parameter assayed was flux through the whole pathway from phenylalanine to carbon dioxide, a parameter that shows a gene-dosage effect[34]. While the data are suggestive of differences in PAH enzymic activities, because hydroxylation is the key regulating step with a high sensitivity coefficient[32] in this pathway[8], we cannot discount the possibility that flux modulation is due to variation at other steps, or competing routes for phenylalanine disposal. The meaningful *in vivo* assay of the hydroxylation reaction itself has proven fraught with practical difficulties and, while recent insights might have resolved earlier technical problems[35,36], there is little reliable data correlating *in vivo* hydroxylation rates with the *PAH* genotype. Therefore, attempts to predict the severity of the genotype effect upon enzyme function have relied on more indirect means.

Most of the 400 mutant *PAH* alleles are presumed to be disease-causing through their primary effect on the PAH protein (enzyme integrity and function). Some alleles (nonsense, frameshift, deletion, splicing), as predicted from sequence data, will be 'nulls' in expression, abolishing enzyme function completely. But over 60% of *PAH* alleles are missense mutations[18], and their effect can only be assessed experimentally by *in vitro* expression[19]. Moreover, three-quarters of HPA probands have heteroallelic *PAH* genotypes[26,27], thus complicating the predicted relationships between genotype and phenotype.

Genotype–phenotype relationships should be clearest when the mutant genotypes are homoallelic or functionally hemizygous. Even in these cases, inconsistencies exist[26,27] between the observed metabolic (HPA) phenotype and the 'predicted residual activity' (PRA), as calculated[7] from the mean of the monoallelic *in vitro* PAH enzyme activities for each mutation comprising the genotype. Some discrepancies can be partly explained by differences owing to methods of enzyme assay[19]; or by chronic overestimation of *in vitro* activities (mutant PAH as a proportion of the wild type) inherent in the expression systems used[19]. This is evidenced by comparisons with PAH enzyme activities in liver biopsies or with *in vivo* phenylalanine oxidation data[34]. Yet, even after accounting for these effects and for errors in metabolic phenotype classification[27], inconsistencies remain. Complexity at the metabolic level might be an explanation, when certain mutations, each associated with moderate levels of residual enzyme activity *in vitro*, are associated, in different individuals, with different metabolic phenotypes (Table 1). On the other hand, these particular mutant proteins could possess properties especially predisposing the enzyme to direct modulation. *In vitro* expression analysis of *PAH* mutations begins to reveal how they exert their effects; and these findings in turn point to the opportunities for modulation of the mutant protein phenotype.

FIGURE 2. Factors influencing phenotype in phenylketonuria



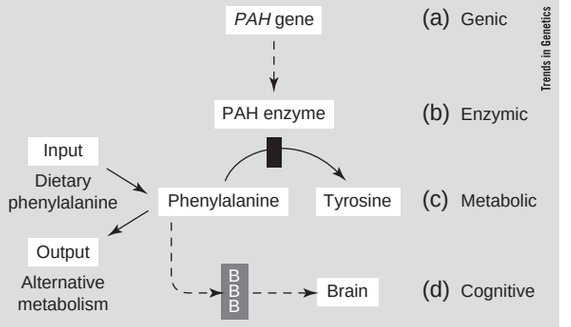
The autosomal recessive trait (hyperphenylalaninemia; HPA in text) and associated disease (phenylketonuria; PKU) have explanations for phenotype beyond a monogenic (mendelian) cause. PKU is MIM 261600 in the McKusick catalog[2]. Other monogenic causes of HPA are the disorders of tetrahydrobiopterin (BH_4) homeostasis, (locus heterogeneity for HPA): they are entered in MIM under 233910, 261630, 261640 and 264070). Symbols: *PAH* for the gene on chromosome 12q24.1; PAH for the homotetrameric enzyme product.

How do *PAH* missense mutations alter enzyme integrity and function?

Most mutant PAH proteins are expressed in mammalian cells at levels that are well below those found in the wild type[19]. Those missense *PAH* alleles, such as the examples in Table 1, which decrease intracellular enzyme activity and PAH protein levels to the same extent, have no effect on specific enzyme activity, while others decrease both the abundance and the specific activity of PAH, and only a few impair function (by catalytic or kinetic effect) without reducing the level of protein. Thus, most *PAH* missense mutations probably reduce the intracellular stability of the protein, by which we mean that the process accounting for reduced expression levels of the protein within cells is not simply the equivalent of decreased inherent (thermodynamic) stability of the protein itself[37,38].

The analysis of six different missense mutant PAH proteins by pulse-chase time courses in rabbit reticulocyte lysates[39] reveals their accelerated disappearance, owing to proteolytic degradation. This process provides an explanation for the observed decrease in levels of mutant protein in transiently transfected mammalian cells. At the same time, these findings emphasise one reason why the PRA derived from mammalian cell studies does not necessarily show a simple 1:1 correlation with the corresponding *in vivo* activity. The activity determined *in vitro*, using cell lysates harvested at a single arbitrary time point after transfection, can reflect only a 'snapshot' of the PAH

FIGURE 3. *PAH* mutations and PKU



Scheme for factors acting at different levels to modify the effect on phenotype of disease-causing human *PAH* alleles. (a) Most *PAH* alleles are missense with a range of effects on enzyme function, but potential interactions between alleles and the effects of regulatory alleles within the gene or elsewhere have not been dismissed. (b) The effect of a mutant allele on enzyme function *in vivo* cannot be predicted precisely from *in vitro* expression (see text). (c) The black bar indicates impaired flux in conversion of phenylalanine to tyrosine in phenylketonuria (PKU); the resulting degree of hyperphenylalaninemia is a function of distributed controls in a complex system of homeostasis, involving inputs and outputs of phenylalanine by factors other than the PAH enzyme. These will modulate the hyperphenylalaninemia phenotype in PKU. (d) Excess phenylalanine in blood is toxic to cognitive development and neurophysiological function. Its transports across the blood–brain barrier (BBB) and into brain cells are mediated; constitutional variation in these transport functions modulates the effect of hyperphenylalaninemia on cognitive phenotype in PKU.

protein level, which will be continually changing as existing protein is degraded and the transient synthesis of new protein tails off. By contrast, the *in vivo* steady-state level of PAH protein will reflect the balance between fluxes through synthetic and degradative pathways that, under physiological controls, will differ quantitatively and qualitatively from those in cultured cells.

The degradation of mutant PAH proteins appears to be triggered by structural (conformational) abnormalities. Detailed study, first of the G46S allele[40], and now of 13 other missense mutations[41,42] (and P.J. Waters *et al.*, unpublished), has revealed altered oligomerization and increased aggregation of PAH protein in each case. Figure 4 shows a schematic model of the competing pathways[38,43] that newly synthesized PAH subunits can enter. Missense mutations that promote PAH aggregation probably do so by affecting the subunit folding and/or the assembly of oligomers, thus altering the kinetic partitioning of protein between the pathways, and reducing the proportion of protein that will form a functional native state.

Some missense *PAH* mutations reveal no predictable effect when examined against the three-dimensional crystal structure of wild-type native PAH (Ref. 20), a finding that is not surprising, because amino acid substitutions can affect protein folding critically, without affecting the stability or activity of the protein's native state[38]. Indeed, the architecture of the PAH protein, unlike many other enzymes[38], seems exquisitely sensitive to misfolding and/or mis-assembly caused by substitutions dispersed almost anywhere throughout the enzyme (http://www.mcgill.ca/pahdb). Because most mutant *PAH* genotypes are heteroallelic, two-hybrid expression systems could prove informative about effects on oligomer assembly where two different mutant subunits are involved[44,45]. In such systems, a semiquantitative estimate of the strength of interaction between two polypeptides is obtained. This is achieved by expressing each polypeptide as a fusion with one domain of a transcriptional activator protein, and assaying the transcription of reporter genes that occurs when the two complementary activator domains are brought into proximity, causing functional reconstitution of the activator[46].

How might the mutant enzyme phenotype be modulated?

Protein stability within cells reflects a fine balance between proteolytic systems and the actions of 'molecular chaperones'[47]. Chaperones protect partially folded proteins from entering non-productive pathways (Fig. 4); while allowing them to fold correctly; inter-individual variation in chaperone function could modify the disease phenotypes associated with missense mutations, when those mutations do not destroy the activity of the native enzyme. Initial data from the co-expression of chaperones with the S349L mutant PAH protein suggest that chaperones can modulate the soluble yield of the mutant PAH protein *in vitro*[45].

Proteolytic pathways that handle aberrant protein are independent factors that can modulate the protein phenotype.

TABLE 1. *PAH* mutations associated with variable metabolic phenotypes

	In vitro[b]		*In vivo*[d]				
Mutation[a]	**Monoallelic data: proportion of wild type (%)**		**Genotype[a]**	**Metabolic phenotype (number of patients)**			
	Enzyme activity	**PAH protein**		**Predicted proportion of normal activity[c] (%)**	**PKU**	**Variant PKU**	**Non-PKU HPA**
I65T	26	25	I65T / I65T	26	1	1	1
			I65T / null	13	4	3	1
R261Q	30	30	R261Q / R261Q	30	3	8	–
			R261Q / null	15	6	3	–
Y414C	50	50	Y414C / Y414C	50	–	4	4
			Y414C / null	25	6	44	6

Abbreviations: HPA, hyperphenylalaninemia; PKU, phenylketonuria.
[a]*PAH* alleles and corresponding genotypes (homozygous or functionally hemizygous) known to be associated with a range of HPA phenotypes. Data from Ref. 26.
[b]Data for *in vitro* expression from Refs 7, 19.
[c]The predicted activity of PAH enzyme is the mean of the *in vitro* expression activities for each allele[7]. It assumes that *in vivo* PAH activity is an exact reflection of the *in vitro* measurement.
[d]The metabolic phenotypes are classified according to criteria described elsewhere[8,26]. PKU is a severe (harmful) phenotype, non-PKU HPA is relatively benign, and variant PKU is intermediate in effect.

Certain mendelian disorders impair proteolytic pathway functions[37], implying that allelic variation at the corresponding loci might modify the variant enzymic phenotype in PKU. Whereas degradation of mutant PAH is proposed to involve the ubiquitin-dependent proteasomal pathway[40,48], other pathways have also been implicated[39], suggesting that modulation in multiple proteolytic pathways owing to allelic variation in genes encoding pathway components could affect the mutant PAH enzymic phenotype.

Thus, PKU has joined a growing repertoire of monogenic disorders in which altered protein folding and/or degradation are implicated in the pathogenesis of the variant phenotype[43,49–51], in ways that escape a direct correlation between *in vivo* catalytic activity and the predicted effect of the allele on it. The challenge then for PKU, as for many other disorders, is to identify what is happening to the mutant protein *in vivo*. Does it, in fact, interact with any chaperones *in vivo*[52]? If so, with which ones; and with which components of proteolytic pathways? If these interactions are important for stability *in vivo*, and thus for the disease phenotype, they could hold promise for combinatorial drug design in the treatment of 'disorders of misfolded proteins'[38].

Further considerations: interactions between alleles?

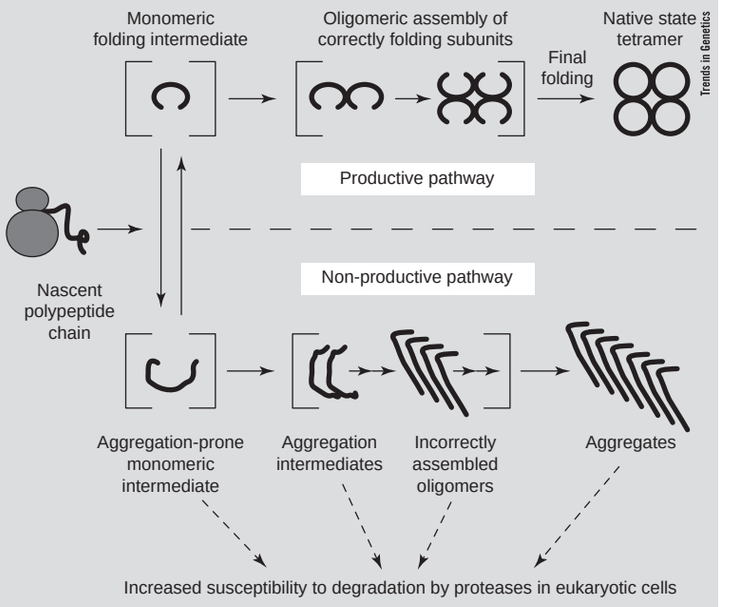
There is further complexity to consider. With regard to pathogenic alleles, there might be more than two mutations on the homologous alleles, implying that complete scanning of the gene to identify all nucleotide changes could be necessary to interpret the effect on phenotype of the mutant genotype at a particular locus. For example, one E203K allele *in cis* with one of a pair of activity-reducing N314D ('Duarte') alleles normalizes galactose-1-phosphate uridyl transferase activity[50]; introduction of the second amino acid substitution alters net charge and neutralizes the destabilizing effect of the aspartate substitution. At least six examples of two mutations *in cis* on one *PAH* allele, the majority in adjacent codons, have been identified in association with HPA (C. Aulehla Scholtz, pers. commun. and the PAHdb website http://www.mcgill.ca/pahdb), but their effects on phenotype have yet to be analyzed in detail.

It should also be kept in mind that the *PAH* locus (100 kb), like many others, harbours an array of polymorphisms; in the case of *PAH*, at least 18 single-nucleotide polymorphisms, eight biallelic restriction fragment length polymorphism sites and two multiallelic sites (STR and VNTR). It is conceivable that these individual polymorphisms and their corresponding extended haplotype structures, along with potential allelic variation in the 5′ untranslated region of the *PAH* gene[53,54], might influence *PAH* gene transcription and expression in ways yet to be discovered. Such an enquiry would be in keeping with the emerging commitment to 'functional genomics'[55].

Genomes speak biochemistry

The genome does not issue instructions to make this or that feature of the organismal phenotype, for example, a brain. Its only orders are to make a polypeptide with a particular function. It speaks 'biochemistry' not 'phenotypes'[56]. In earlier times, it was appropriate to think of PKU and other inborn errors like it as mendelian disorders that affect controlled metabolic pathways, in cycles and networks[57]. One learned to understand them in biochemical terms, to perceive the molecular basis of their disordered phenotype, and to identify the critical enzymic component with the highest sensitivity coefficient that acts as a control element in the functional pathway contributing to metabolic homeostasis[32]. Here, our analysis of PKU does not alter the validity of such thinking. It only shows that the overall PKU phenotype at its enzymic, metabolic and cognitive levels is more complex than it was presumed. It is monogenic, multifactorial and complex (Table 2); or, to borrow a point of view, one might say it is both Mendelian–Garrodian and Galtonian–Fisherian (for an explanation, see Ref. 2) in nature. The hope that the mutant genotype at the major locus would predict the variant phenotype originated in the new ability to study

FIGURE 4. Model of competing folding and aggregation pathways of PAH



Species in brackets [ ] are putative intermediates. The monomeric folding intermediate is in equilibrium with an aggregation-prone intermediate. Mutations that affect folding shift this equilibrium towards the aggregation-prone intermediate, resulting in a greater proportion of PAH protein entering the non-productive pathway, and suffering degradation by proteases. Adapted from Ref. 38.

TABLE 2. Genotype–phenotype relationships in phenylketonuria: factors at the levels of the genome and the phenotype

Level	Entity	Modifiers	Descriptor
Genotype			
	PAH gene		
	Pathogenic allele	*Cis* alleles	Mendelian (allelic heterogeneity)
	Mutant genotype	*Trans* acting factors	Mendelian (autosomal recessive)
Phenotype			
PAH protein	PAH polypeptide	Translation; chaperones; proteases	Complex trait?
Enzymic	PAH oligomeric enzyme	Complementation	–
	Hydroxylation cofactor BH_4	Synthesis, recycling	Locus heterogeneity
	Hydroxylating reaction	–	–
Metabolic	Phenylalanine homeostasis	Input (protein)	Multifactorial
		Disposal (metabolism)	Complex trait
Cognitive	Brain function	Blood–brain barrier	Complex trait

Abbreviation: BH_4, tetrahydrobiopterin (obligate cofactor for PAH enzyme).

mutant allelic expression in reductive *in vitro* expression systems and in transgenic animals. But genomes function *in vivo*, where much more than the major gene (here the *PAH* gene) is expressed and where the whole organismal phenotype is more than the sum of the parts; it is an emergent property[58]. The biological point of view has implications for the investigation, counselling and treatment of patients with genetic disease, because each person, perhaps excepting monozygotic twin pairs, with a nominal monogenic disorder is likely to have a particular phenotypic form of it.

The ideas expressed here are not really new. Only the ability to test the hypotheses and produce evidence is new. Lionel Penrose, in his inaugural address as Galton Professor in 1946 (Ref. 59) observed, among other things, that IQ was not predictable from genotype within PKU families, and opined that a multidisciplinary approach would benefit our understanding of this disease. His original paper has been reprinted[60] to remind us that there were giants on whose shoulders we stand to see a little further.

Acknowledgements

This article started its life as a Distinguished Speakers' Lecture at the Annual Meeting of the American Society of Human Genetics (San Francisco, 1996). Its further growth was stimulated by a European Science Foundation Workshop on protein folding and degradation as a disease-causing mechanism (Århus, Denmark, 1998).

References

1 Ridley, M. (1991) A survey of science. The edge of ignorance. *The Economist,* February 16, 1–22
2 McKusick, V.A. (1998) *Mendelian Inheritance in Man. A Catalog of Human Genes and Genetic Disorders,* The Johns Hopkins University Press
3 Bassett, D.E.J. *et al.* (1997) Genome cross-referencing and XREFdb: implications for the identification and analysis of genes mutated in human disease. *Nat. Genet.* 15, 339–344
4 Cotton, R.G.H. (1993) Current methods of mutation detection. *Mutat. Res.* 285, 125–144
5 Grompe, M. (1993) The rapid detection of unknown mutations in nucleic acids. *Nat. Genet.* 5, 111–117
6 Scriver, C.R. (1991) Phenylketonuria – genotypes and phenotypes. *New Engl. J. Med.* 324, 1280–1281
7 Okano, Y. *et al.* (1991) Molecular basis of phenotypic heterogeneity in phenylketonuria. *New Engl. J. Med.* 324, 1232–1238
8 Scriver, C.R. *et al.* (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (7th edn) (Scriver, C.R. *et al.*, eds), pp. 1015–1075, McGraw-Hill
9 Romeo, G. and McKusick, V.A. (1994) Phenotypic diversity, allelic series and modifier genes. *Nat. Genet.* 7, 451–453
10 Summers, K.M. (1996) Relationship between genotype and phenotype in monogenic diseases: relevance to polygenic diseases. *Hum. Mutat.* 7, 283–293
11 Estivill, X. (1996) Complexity in a monogenic disease. *Nat. Genet.* 12, 348–350
12 Scriver, C.R. *et al.* (1987) The Hartnup phenotype: mendelian transport disorder, multifactorial disease. *Am. J. Hum. Genet.* 40, 401–412
13 Wu, Y. *et al.* (1994) A lag in intracellular degradation of mutant α1-antitrypsin correlates with the liver disease phenotype in homozygous PiZZ α1-antitrypsin deficiency. *Proc. Natl. Acad. Sci. U. S. A.* 91, 9014–9018
14 Qu, D. *et al.* (1996) Degradation of a mutant secretory protein, α1-antitrypsin Z, in the endoplasmic reticulum requires proteasome activity. *J. Biol. Chem.* 271, 22791–22795
15 Scriver, C.R. (1994) Science, medicine and phenylketonuria. *Acta Pediatr.* (Suppl.) 407, 11–18
16 Folling, A. (1934) Uber Ausscheidung von Phenylbrenztraubensaure in den Harn als Stoffwechselanomalie in Verbindung mit Imbezillitat. *Hoppe-Seyler's Z. Physiol. Chem.* 277, 169–176
17 Scriver, C.R. *et al.* (1996) in *Variation in the Human Genome* (Chadwick, D. and Cardew, G., eds), pp. 73–96, John Wiley & Sons
18 Nowacki, P.M. *et al.* (1998) PAH mutation analysis consortium database: 1997. Prototype for relational locus-specific mutation databases. *Nucleic Acids Res.* 26, 220–225
19 Waters, P.J. *et al.* (1998) *In vitro* expression analysis of mutations in phenylalanine hydroxylase: linking genotype to phenotype and structure to function. *Hum. Mutat.* 11, 4–17
20 Erlandsen, H. *et al.* (1997) Crystal structure of the catalytic domain of human phenylalanine hydroxylase reveals the structural basis for phenylketonuria. *Nat. Struct. Biol.* 4, 995–1000
21 Fusetti, F. *et al.* (1998) Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. *J. Biol. Chem.* 273, 16962–16967
22 Weiss, K.M. (1996) Is there a paradigm shift in genetics? Lessons from the study of human diseases. *Mol. Phylogenet. Evol.* 5, 259–265
23 Meli, C. *et al.* (1998) Different clinical manifestations in siblings with identical phenylalanine hydroxylase gene mutations. *J. Inher. Metab. Dis.* (Suppl. 2) 21, 10
24 Di Silvestre, D. *et al.* (1991) Different clinical manifestations in three siblings with identical phenylalanine hydroxylase genes. *Am. J. Hum. Genet.* 48, 1014–1016
25 Treacy, E. *et al.* (1996) *In vivo* disposal of phenylalanine in phenylketonuria: A study of two siblings. *J. Inher. Metab. Dis.* 19, 595–602
26 Kayaalp, E. *et al.* (1997) Human phenylalanine hydroxylase mutations and hyperphenylalaninemia phenotypes: a metanalysis of genotype–phenotype correlations. *Am. J. Hum. Genet.* 61, 1309–1317
27 Guldberg, P. *et al.* (1998) A European multicenter study of phenylalanine hydroxylase deficiency: classification of 105 mutations and a general system for genotype-based prediction of metabolic phenotype. *Am. J. Hum. Genet.* 63, 71–79
28 Ramus, S.J. *et al.* (1993) Comparison of genotype and intellectual phenotype in untreated PKU patients. *J. Med. Genet.* 30, 401–405
29 Weglage, J. *et al.* (1998) *In vivo* NMR spectroscopy in patients with phenylketonuria. Clinical significance of interindividual differences in brain phenylalanine concentrations. *J. Inher. Metab. Dis.* 21, 81–83
30 Moller, H.E. *et al.* (1998) Blood–brain barrier phenylalanine transport and individual vulnerability in phenylketonuria. *J. Cereb. Blood Flow Metab.* 18, 1184–1191
31 Scriver, C.R. *et al.* (1985) Normal plasma free amino acid values in adults: the influence of some common physiological variables. *Metabolism* 34, 868–873
32 Kacser, H. and Burns, J.A. (1981) The molecular basis of dominance. *Genetics* 97, 639–666
33 Salter, M. *et al.* (1986) Quantification of the importance of individual steps in the control of aromatic amino acid metabolism. *Biochem. J.* 234, 635–647
34 Treacy, E.P. *et al.* (1997) Analysis of phenylalanine hydroxylase genotypes and hyperphenylalaninemia phenotypes using L-[1-^{13}C] phenylalanine oxidation rates *in vivo*: A pilot study. *Pediatr. Res.* 42, 430–435
35 Van Spronsen, F.J. *et al.* (1998) Phenylketonuria (PKU): the *in vivo* hydroxylation rate of phenylalanine into tyrosine is decreased. *J. Clin. Invest.* 101, 2875–2880
36 Scriver, C.R. (1998) An ongoing debate over phenylalanine hydroxylase deficiency in phenylketonuria. *J. Clin. Invest.* 101, 2613–2614
37 Kato, G.J. (1999) Human genetic diseases of proteolysis. *Hum. Mutat.* 13, 87–98
38 Betts, S. *et al.* (1997) Mutational effects on inclusion body formation. *Adv. Protein Chem.* 50, 243–264
39 Waters, P.J. *et al.* Missense mutations in the phenylalanine hydroxylase gene (*PAH*) can cause accelerated proteolytic turnover of PAH enzyme; a mechanism underlying phenylketonuria. *J. Inher. Metab. Dis.* (in press)
40 Eiken, H.G. *et al.* (1996) PKU mutation G46S is associated with increased aggregation and degradation of the phenylalanine hydroxylase enzyme. *Hum. Mutat.* 7, 228–238
41 Bjorgo, E. *et al.* (1998) Partial characterization and three-dimensional-structural localization of eight mutations in exon 7 of the human phenylalanine hydroxylase gene associated with phenylketonuria. *Eur. J. Biochem.* 257, 1–10
42 Waters, P.J. *et al.* (1998) Alterations in protein aggregation and degradation due to mild and severe missense mutations (A104D, R157N) in the human phenylalanine hydroxylase gene (*PAH*). *Hum. Mutat.* 12, 344–354
43 Bross, P. *et al.* (1998) Impaired folding and subunit assembly as disease mechanism: the example of medium-chain acryl-CoA-dehydrogenase deficiency. *Prog. Nucleic Acids Res. Mol. Biol.* 58, 303–337
44 Hufton, S.E. *et al.* (1998) Structure/function analysis of the domains required for the multimerisation of phenylalanine hydroxylase. *Biochim. Biophys. Acta* 1382, 295–304
45 Perez, B. *et al.* (1998) Effect of the S349L PKU mutation on the folding, stability, and activity of the PAH protein. *Eur. J. Hum. Genet.* (Suppl. 1) 6, 167
46 Fields, S. and Song, O.K. (1989) A novel genetic system to detect protein–protein interactions. *Nature* 340, 245–246
47 Hartl, F. *et al.* (1994) Molecular chaperones in protein folding: the art of avoiding sticky situations. *Trends Biochem. Sci.* 19, 20–25
48 Doskeland, A.P. and Flatmark, T. (1996) Recombinant human phenylalanine hydroxylase is a substrate for the ubiquitin-conjugating enzyme system. *Biochem. J.* 319, 941–945
49 Brooks, D.A. (1997) Protein processing: a role in the pathophysiology of genetic disease. *FEBS Lett.* 409, 115–120
50 Lai, K. *et al.* (1998) Duarte allele impairs biostability of human galactose-1-phosphate uridyltransferase in human lymphoblasts. *Hum. Mutat.* 11, 28–38
51 Kuznetsov, G. and Nigam, S.K. (1998) Folding of secretory and membrane proteins. *New Engl. J. Med.* 339, 1688–1695
52 Netzer, W.J. and Hartl, F.U. (1998) Protein folding in the cytosol: chaperonin-dependent and -independent mechanisms. *Trends Biochem. Sci.* 23, 68–73
53 Konecki, D.S. *et al.* (1992) Structural characterization of the 5′ region of the human phenylalanine hydroxylase gene. *Biochemistry* 31, 8363–8368
54 Svensson, E. *et al.* (1993) Three polymorphisms but no disease-causing mutations in the proximal part of the promoter of the phenylalanine hydroxylase gene. *Eur. J. Hum. Genet.* 1, 306–313
55 Strachan, T. *et al.* (1997) A new dimension for the human genome project: towards comprehensive expression maps. *Nat. Genet.* 16, 126–132
56 Plasterk, R.H.A. (1999) Hershey heaven and *Caenorhabditis elegans. Nat. Genet.* 21, 63–64
57 Scriver, C.R. (1985) in *Congenital Metabolic Diseases* (Wapnir, R.A., ed.), pp. 11–40, Marcel Dekker
58 Mayr, E. (1982) *The Growth of Biological Thought. Diversity, Evolution and Inheritance,* pp 63–67, The Belknap Press of Harvard University Press
59 Penrose, L.S. (1946) Phenylketonuria. A problem in eugenics. *Lancet* 1, 949–953
60 Penrose, L.S. (1998) Phenylketonuria. A problem in eugenics. *Ann. Hum. Genet.* 62, 193–202
61 Vogel, F. and Motulsky, A. (1986) *Human Genetics Problems and Approaches,* pp. 181–183, Springer-Verlag
62 Beaudet, A.L. *et al.* (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (7th edn) (Scriver, C.R. *et al.*, eds), pp. 53–118, McGraw-Hill

Reference added in press

63 Kobe, B. *et al.* (1999) Structural basis of autoregulation of phenylalanine hydroxylase. *Nat. Struct. Biol.* 6, 442–448