ORIGINAL ARTICLE

Mapping the functional landscape of frequent phenylalanine hydroxylase (PAH) genotypes promotes personalised medicine in phenylketonuria

Marta K Danecka,1 Mathias Woidy,1 Johannes Zschocke,2 François Feillet,3 Ania C Muntau,4 Sören W Gersting1

ABSTRACT

Background In phenylketonuria, genetic heterogeneity, frequent compound heterozygosity, and the lack of functional data for phenylalanine hydroxylase genotypes hamper reliable phenotype prediction and individualised treatment.

Methods A literature search revealed 690 different phenylalanine hydroxylase genotypes in 3066 phenylketonuria patients from Europe and the Middle East. We determined phenylalanine hydroxylase function of 30 frequent homozygous and compound heterozygous genotypes covering 55% of the study population, generated activity landscapes, and assessed the phenylalanine hydroxylase working range in the metabolic (phenylalanine) and therapeutic (tetrahydrobiopterin) space.

Results Shared patterns in genotype-specific functional landscapes were linked to biochemical and pharmacological phenotypes, where (1) residual activity below 3.5% was associated with classical phenylketonuria unresponsive to pharmacological treatment; (2) lack of defined peak activity induced loss of response to tetrahydrobiopterin; (3) a higher cofactor need was linked to inconsistent clinical phenotypes and low rates of tetrahydrobiopterin response; and (4) residual activity above 5%, a defined peak of activity, and a normal cofactor need were associated with pharmacologically treatable mild phenotypes. In addition, we provide a web application for retrieving country-specific information on genotypes and genotype-specific phenylalanine hydroxylase function that warrants continuous extension, updates, and research on demand.

Conclusions The combination of genotype-specific functional analyses with biochemical, clinical, and therapeutic data of individual patients may serve as a powerful tool to enable phenotype prediction and to establish personalised medicine strategies for dietary regimens and pharmacological treatment in phenylketonuria.

INTRODUCTION

The large number of phenylalanine hydroxylase (PAH) mutations (625 entries (public total), http://www.hgmd.org), the high incidence of compound heterozygosity, and the variability in distribution of common mutations between ethnic groups and geographical areas makes phenylketonuria (PKU, OMIM #261600) a genetic disease with pronounced allelic heterogeneity. PKU caused by deficiency of the PAH enzyme1 2 has an overall incidence of 1:10 000 in European descendants, which varies considerably among different populations.3 4

The clinical phenotype of PKU is driven by mutation-induced loss of PAH function. A continuum of residual in vivo PAH enzyme activity determines the clinical picture, ranging from severe classical PKU via mild PKU to mild hyperphenylalaninaemia (MHP) with increasing activity and decreasing blood phenylalanine concentrations from severe to mild.5 6 In addition to standard treatment using a phenylalanine-restricted diet, a significant proportion of patients is amenable to pharmacological treatment with the PAH cofactor tetrahydrobiopterin (BH4).7 8 This constitutes a superordinate phenotype that requires residual PAH activity9 10 and further adds to the phenotypic variability of PAH deficiency.11

Mutations in the PAH gene can induce PAH loss-of-function by different molecular mechanisms. Splicing mutations, nonsense mutations, and out-of-frame indels lead to a complete loss of the PAH protein, whereas missense mutations as well as some in-frame indels still lead to the production of variant PAH proteins. Loss of function of variant PAH is a consequence of loss of specific enzymatic activity and/or loss of effective intracellular PAH protein amount caused by protein misfolding.12 16

Databases such as PAHdb (http://www.pahdb.mcgill.ca) and BIOPKU (http://www.biopku.org) provide general information on PAH mutations and associated phenotypes; however, there is no database access on occurrence and frequency of genotypes in individual countries. A significant gap of knowledge concerns the enzymatic function associated with PAH genotypes. Available data on residual PAH enzyme activity is limited to single mutations expressed in different cellular systems using diverse enzyme assay protocols.5 13 16 19

Non-invasive functional assays, for example, in blood cells, are not feasible because metabolic PAH function is restricted to the liver. Therefore, PAH activity is commonly determined by means of experimental in vitro systems. Since calculated enzyme activities based on heterogeneous data on single mutations do not reliably reflect residual activity in the patient, a consistent model system to assess PAH enzyme function arising from two different alleles reflecting the most common genotypes is needed. In addition, previous work has revealed that PAH function depends on the metabolic phenotype and a potential pharmacologic
administration of BH₄.²⁰ Genotype-specific data on PAH function assessed at different substrate and cofactor concentrations may thus assist physicians in managing the dietary treatment and in predicting the probability of pharmacological response in clinical practice. In order to provide this information, we (1) mapped the most common reported PAH genotypes in Europe and Middle East; (2) determined PAH residual function of frequent homozygous and compound heterozygous genotypes in a cell based model; (3) generated activity landscapes; and (4) assessed the PAH working range.

This study provides the first comprehensive dataset on PAH function arising from the most common genotypes carried by PKU patients in Europe and the Middle East. PAH enzyme activity assessed at a wide range of phenylalanine and BH₄ concentrations mirrors metabolic and therapeutic conditions and thus provides information on the optimal range of substantial PAH enzyme function in PKU patients. This may allow for genotype-specific optimisation of phenylalanine restricted dietary regimens and treatment with BH₄.

**METHODS**

**Data sources for PAH genotypes and phenotypes**

Data from 24 publications reporting PAH genotypes in 19 countries in Europe and the Middle East and unpublished data from German and French medical centres were analysed (see online supplementary table S1). Types and frequencies of mutations were obtained from the Human Gene Mutation Database (http://www.hgmd.org), the PAHdb (http://www.pahdb.mcgill.ca) and the BIOPKU database (http://www.biopku.org). Data on PKU clinical phenotypes and BH₄ responsiveness were obtained from BIOPKU.

**Assay of PAH activity**

In this study, we analysed 30 PAH genotypes consisting of 18 PAH missense mutations and one intronic mutation c.1066–11G>A (previously known as IVS10–11G>A),¹⁶ which results in the insertion of three amino acids (p.Gln355_Tyr356insGlyLeuGln) (see online supplementary table S2). Constructs carrying PAH mutations were co-expressed in COS-7 cells reflecting PAH genotypes. The multi-well PAH activity assay for determination of activity landscapes from eukaryotically expressed PAH was performed as previously described for recombinantly expressed and purified PAH²⁰ with modifications. Varying volumes of a 5 mM l-phenylalanine (Sigma-Aldrich) solution and 22.35 mM NaHEPES, pH 7.3, were injected (FLUOstar OPTIMA, BMG Labtech) into all wells of a 96-well plate. This resulted in 12 columns of varying l-phenylalanine concentrations (0–4000 μM). A reaction buffer containing 1 mg/mL catalase (Sigma-Aldrich), 10 μM ferrous ammonium sulfate (Sigma-Aldrich), and 20 μL protein lysate per well was subsequently injected in all 96 wells. After pre-incubation for 5 min, the reaction was triggered by the addition of variable concentrations of BH₄ (6R-erythro-5,6,7,8-tetrahydrobipterin, Schircks Laboratories) (0–500 μM, rows of the 96-well plate) stabilised in 100 mM dithiothreitol (DTT, Sigma-Aldrich). The PAH activity assay was incubated at 25°C for 15 min and stopped by acetic acid. All concentrations mentioned refer to the final concentration in a 200 μL reaction mixture. The formation of the l-Tyr product was quantified by high performance liquid chromatography (HPLC) using a Hypersil ODS-2 column (ThermoScientific). All PAH activity landscapes were assayed in three independent experiments and the data were combined. Genotypes were divided into three groups based on PAH activity determined at the area of peak activity (group 1, PAH residual activity <3.5%; group 2, 3.5% and <5%; group 3 ≥5%). If no area of peak activity was delimitable, activity at standard conditions (l-phenylalanine, 1 mM; BH₄, 75 μM) was used for classification.

**PAH activity landscapes**

PAH activity landscapes were generated as previously described²⁰ with modifications. The dataset obtained from the multi-well PAH activity assay in a 12×8 matrix corresponding to 12 different l-phenylalanine concentrations (0–4000 μM; columns) at eight BH₄ concentrations (0–500 μM; rows) was loaded into an analysis software (GraphPad Prism 3.0a). Non-linear regression analysis allowed for expanding the dataset for BH₄ concentrations from eight measured to 200 interpolated values following a substrate inhibition curve. This resulted in a 12×200 matrix of activity values. For the generation of landscapes, the data matrix was exported to the free software package R (http://www.r-project.org). In order to draw a smooth surface of the landscape, the function interp.loess from an additional R package mgcv (http://cran.r-project.org/web/packages/mgcv/index.html) was applied, which interpolates between two data points by using local polynomial regression fitting to find a function between them. This resulted in a matrix of 400×400 data points, which was then depicted as a smooth landscape plot using the function image.plot from package fields (http://cran.r-project.org/web/packages/fields/index.html). To facilitate calculation of landscapes, a web application (http://pah-activitylandscapes.org/calcLandscapes.php) was set up accepting comma-separated files and automatically colouring landscapes depending on the measured and interpolated enzyme activity values.

**Web application**

An interactive website provides country- and region-specific information on PAH genotypes and links genotypes to graphical representation of activity landscapes (http://pah-activitylandscapes.org). The site will be updated and users can request the generation of landscapes of additional genotypes.

More detailed information on material and methods is available in the online supplementary data.

**RESULTS**

**PAH genotypes in Europe and the Middle East**

We performed a comprehensive PubMed literature search to determine the frequency and region-specific distribution of PAH genotypes. The study focused on Europe and the Middle East owing to the shared ethnic background in this region and data from 20 countries spanning from Portugal to Iran were selected for evaluation (figure 1 and see online supplementary table S1). The 3066 PKU patients of this study population displayed 690 different genotypes based on 264 mutations (http://pah-activitylandscapes.org).

The type and frequency of single PAH mutations constituting the genotypes in our study cohort matched well with the frequency documented in gene databases (table 1), and mutations considered as prevalent in the European population covered 80.3% of alleles in our cohort (see online supplementary table S2). The predominant variant p.Arg408Trp (37%) occurred in all countries except Lebanon and Cyprus, while the second most common variant, p.Gln355_Tyr356insGlyLeuGln (10.2%), did not appear in Croatian, Slovenian, and Swedish populations. A region-specific distribution was also observed for other mutations (see online supplementary figure S1).

The most common full PAH genotypes for eight countries are shown in figure 1. Of all genotypes analysed, 9% were homozygous, 91% compound heterozygous, and 49% of PKU patients...
carried missense mutations on both alleles. The 30 most frequent genotypes encoding expressible PAH proteins from both alleles (table 1) covered 55% of patients in our study population (BIOPKU, 33%; PAHdb, 18%). The frequency of these genotypes was high in Croatia (88%), Turkey (87%), and France (61%) and lower in Slovenia (11%) and Spain (10%). The genotype p.[Arg408Trp];[Arg408Trp] occurred in 21% of patients and was predominant in Northern and Eastern Europe (Lithuania, Norway, Poland, Slovakia), followed by homozygous p.Gln355_Tyr356insGlyLeuGln (5.2%), which was frequent in the Middle East (Iran, Turkey, Israel, Lebanon, Armenia). Genotype frequencies in five geographical regions (Northern Europe, Southern Europe, Western Europe, Eastern Europe, and Middle East) and country-specific frequencies are summarised in online supplementary figure S2 and are available at http://pah-activitylandscapes.org.

Residual PAH function of wild-type PAH and frequent PAH genotypes
Ex vivo assessment of PAH enzyme activity in liver tissue requires unjustifiable invasive procedures. Thus, we aimed to provide a cellular model that quantifies genotype-specific residual PAH function and takes into account PAH protein homeostasis and the metabolic environment. PAH activity of

Figure 1  Phenylalanine hydroxylase (PAH) genotypes in Europe and Middle East. Countries included in the study are marked in blue. The graphs below show the frequencies of the most common PAH genotypes in individual countries for which data from more than 100 patients were available. In case the effect of intronic mutation on protein level is not known, the substitution at DNA level was used. The map was created using http://www.stepmap.de.

cells expressing variant PAH encoded by both alleles of individual PAH genotypes was determined at the physiological to pathological range of phenylalanine concentrations (25–4000 μM) as well as a physiological to supratherapeutic range of BH4 cofactor concentrations (9–500 μM).

The activity landscape of the wild-type enzyme showed a peak maximum enzyme activity of 72 658 pmol L-phenylalanine x min x mg protein at 431 μM L-phenylalanine and 130 μM BH4 (figure 2, see online supplementary figure S3 and table 2). This was in line with previous results for purified wild-type PAH protein expressed in Escherichia coli.20 The optimal working range for L-phenylalanine spanned from 135 μM (K) to 747 μM (K). Thus, the PAH enzyme displayed very low activity at L-phenylalanine concentrations <120 μM and a broader range of considerable activity at higher L-phenylalanine concentrations, even >600 μM. In addition, the wild-type enzyme showed high activity at cofactor concentrations between 30 μM (C) and 294 μM (K). At L-phenylalanine concentrations above the therapeutic threshold of 360 μM21 the enzyme showed a need for increased BH4 concentrations to maintain the same level of activity.

Residual PAH activity of eukaryotically expressed genotypes (table 2, see online supplementary tables S3 and S4) ranged from 1.7% (p.Arg408Trp;p.Arg408Trp) to 11% (p.Arg261Gln;p.Arg261Gln) of wild-type PAH, which fits well with the threshold of 15% PAH activity related to metabolic decompensation.2 Genotypes were divided into three groups: those with PAH residual activity <3.5% (group 1, n=9); those

<table>
<thead>
<tr>
<th>Table 1</th>
<th>PAH genotypes in PKU patients from Europe and the Middle East and associated frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele 1</td>
<td>Allele 2</td>
</tr>
<tr>
<td>1</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>3</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>4</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>5</td>
<td>p.Arg261Gln</td>
</tr>
<tr>
<td>6</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>7</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>8</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>9</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>10</td>
<td>p.Pro281Leu</td>
</tr>
<tr>
<td>11</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>12</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>13</td>
<td>p.Leu48Ser</td>
</tr>
<tr>
<td>14</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>15</td>
<td>p.Arg261Gln</td>
</tr>
<tr>
<td>16</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>17</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>18</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>19</td>
<td>p.Arg252Trp</td>
</tr>
<tr>
<td>20</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>21</td>
<td>p.Pro281Leu</td>
</tr>
<tr>
<td>22</td>
<td>p.Ala300Ser</td>
</tr>
<tr>
<td>23</td>
<td>p.Tyr414Cys</td>
</tr>
<tr>
<td>24</td>
<td>c.441+5G&gt;T</td>
</tr>
<tr>
<td>25</td>
<td>p.Leu48Ser</td>
</tr>
<tr>
<td>26</td>
<td>p.Leu48Ser</td>
</tr>
<tr>
<td>27</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>28</td>
<td>p.Leu48Ser</td>
</tr>
<tr>
<td>29</td>
<td>p.Ile65Thr</td>
</tr>
<tr>
<td>30</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>31</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>32</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>33</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>34</td>
<td>p.Arg261Gln</td>
</tr>
<tr>
<td>35</td>
<td>p.Arg408Trp</td>
</tr>
<tr>
<td>36</td>
<td>p.Arg408Trp</td>
</tr>
</tbody>
</table>

The table shows the 36 most frequent genotypes in Europe and the Middle East and their relative frequencies. *Frequencies of genotypes retrieved from the literature (PubMed search) and obtained via personal communications. †Frequencies of genotypes retrieved from public databases: BIOPKU (query 23 April 2014) and PAHdb (query 19 April 2014). ‡30 PAH genotypes encoding expressible PAH proteins from both alleles selected for experimental studies. PAH, phenylalanine hydroxylase; PKU, phenylketonuria.
with activities ≥3.5% and <5% (group 2, n=3); and those ≥5% (group 3, n=18) (table 3, figures 3–5). Group 3 contained all genotypes with mutations known to be associated with mild clinical phenotypes (p.Val243Ala, p.Ala300Ser, p.Glu390Gly, p.Ala403Val). Although p.Gln355_Tyr356insGlyLeuGln is commonly considered a null mutation,16 it reached up to 5% of activity in the homozygous state, suggesting preservation of some residual function in patients. Taking into consideration the 30 most frequent genotypes, 37% of PKU patients in this cohort carried a genotype associated with low residual PAH function <5%, and 19% of patients can be expected to have a considerable residual enzyme function with activities ≥5%. Armenia, Croatia, Iran, Lithuania, Poland, Slovakia and Turkey showed a predominance of severe genotypes, whereas France, Italy, and the Czech Republic are countries where milder biochemical phenotypes are more frequent than severe ones (see online supplementary table S5).

Peak enzyme activity and optimal working range of PAH genotypes

We aimed to identify the metabolic context for best PAH function. Peak PAH enzyme activity—that is, the relation of optimal enzyme activity to substrate and cofactor concentrations—constitutes an important functional feature in phenylalanine metabolism and is of clinical relevance. We defined the working optimum for the PAH enzyme as the range of activity between [S]0.5 and (K) for the substrate and between [C]0.5 and (K) for the cofactor—that is, the ranges of all activities above 50% (table 2 and see online supplementary figure S4). Functional activity landscapes from group 1 displayed very low activity with no defined peak (figure 3). In group 2, enzyme activities were still low but an area of peak activity mapping to the area of wild-type PAH was delimitable (figure 4). Group 3 genotypes showed considerable residual activities and delimitable peak areas with the exceptions of p.[Pro281Leu];[Gln355_Tyr356insGlyLeuGln] and homozygous p.[Gln355_Tyr356insGlyLeuGln];[Gln355_Tyr356insGlyLeuGln] (figure 5). However, the position and shape of areas of peak enzyme activity showed large variation within this group. The substrate concentration at peak activity ranged from 231–882 μM for genotypes that displayed a defined peak area. For p.[Ile306Val];[Arg408Trp], p.[Ala403Val];[Arg408Trp], and p.[Tyr414Cys];[Arg408Trp] peak activity mapped to significantly lower phenylalanine concentrations than observed for the wild-type. These left-shifts were associated with higher apparent affinity of the enzymes towards the substrate, as reflected by lower values for [S]0.5. On the other hand, p.[Arg261Gln];[Arg261Gln], p.[Arg261Gln];[Arg408Trp], p.[Arg261Gln];[Gln355_Tyr356insGlyLeuGln], and p.[Arg297His];[Arg408Trp] led to peak activities at higher phenylalanine concentrations associated with lower apparent affinity of the enzyme towards the substrate, as reflected by higher values for [S]0.5. All but one of these genotypes contained p.Arg261Gln. All other genotypes displayed peak activities at L-phenylalanine concentrations in the wild-type range.

The working range of wild-type PAH with respect to the substrate covered an interval of >600 μM. Broadening of the working range of the PAH enzyme (interval >900 μM), denoting high residual enzyme activity over a wide range of phenylalanine concentrations, was observed for p.[Leu48Ser];[Arg261Gln], p.[Arg261Gln];[Arg261Gln], p.[Gln355_Tyr356insGlyLeuGln], and p.[Arg297His];[Arg408Trp]. Three of these genotypes contained the allele p.Arg261Gln and 4.1% of the patients in our study population carried a genotype associated with this activity landscape feature. On the other hand, a narrow working range leading to considerable PAH activity only in a tight window of substrate concentrations (interval <460 μM) was observed for p.[Leu48Ser];[Leu48Ser], p.[Leu48Ser];[Arg158Gln], p.[Leu48Ser];[Arg408Trp], p.[Leu48Ser];[Gln355_Tyr356insGlyLeuGln], p.[Ala403Val];[Gln355_Tyr356insGlyLeuGln], p.[Ile306Val];[Arg408Trp], p.[Ala403Val];[Arg408Trp], and p.[Tyr414Cys];[Arg408Trp], with p.Leu48Ser and p.Arg408Trp being the leading mutations in this group. These occurred in 7.7% of the patients. Genotypes showing a left-shift always showed a smaller area of enzyme activity, but not all genotypes with narrow working ranges were left-shifted. Thus, high affinity towards the substrate and a smaller area of phenylalanine-dependent activity range may be due to independent mechanisms.

The cofactor concentration at peak activity ranged from 94–409 μM. All genotypes with a residual activity ≥3.5% and peak enzyme activities at lower cofactor concentrations than the wild-type (p.[Leu48Ser];[Arg261Gln], p.[Arg261Gln];[Arg261Gln], p.[Arg261Gln];[Arg408Trp], p.[Arg297His];[Arg408Trp], p.[Glu490Ala];[Arg408Trp], p.[Ala403Val];[Arg408Trp]) included the alleles p.Arg261Gln or p.Arg408Trp. Interestingly, a shift towards lower cofactor concentrations was not associated with an increase in the apparent affinity of the enzymes towards the cofactor as reflected by comparable values for [C]0.5. Genotypes p.[Leu48Ser];[Arg408Trp], p.[Leu48Ser];[Gln355_Tyr356insGlyLeuGln], p.[Ile306Thr];[Gln355_Tyr356insGlyLeuGln], p.[Arg408Trp];[Gln355_Tyr356insGlyLeuGln], p.[Ala300Ser];[Gln355_Tyr356insGlyLeuGln], p.[Arg408Trp];[Arg158Gln], and p.[Leu48Ser];[Arg408Trp] led to peak activities at higher BH4 concentrations. It is of note that the presence of the allele p.Gln355_Tyr356insGlyLeuGln correlated with a shift of peak activity towards higher BH4 concentrations when residual activity is ≥5%. Most of these genotypes contained a mutation previously classified as being inconsistently linked to clinical phenotypes (p.[Leu48Ser], p.[Ile65Thr], p.[Ala300Ser]) (see online supplementary figure S5). Patients carrying these genotypes may need higher cofactor concentrations for optimal enzyme activity. All other genotypes induced peak activities at similar BH4 concentrations than the wild-type.

Impact of PAH function on clinical phenotypes

Next, we linked functional data to clinical phenotypes retrieved from BIOPKU (see online supplementary figure S5). Group 1 genotypes with very low PAH activities lacking a well-defined peak of activity were consistently associated with BH4 non-responsive classical PKU. One exception, p.[Arg241His];[Arg408Trp], showed a...
leading phenotype of mild PKU (70% of patients), yet responsiveness to BH4 was only assessed for two cases. Group 2 genotypes, all displaying intermediate residual activity and a defined area of peak activity, showed phenotypic heterogeneity with both mild and classical PKU phenotypes and a significant share of patients responding to BH4. Among group 3, two subgroups were analysed. Group 3 genotypes without area of peak activity (p.[Pro281Leu]; [Gln355_Tyr356insGlyLeuGln]), leading to a total loss of landscape architecture, were mainly associated with classical PKU and a low response rate to BH4 (10–25%). These also displayed rather low residual activities within this group (5.1%, 5.7%). Yet other genotypes with residual activities in the same range, but showing a defined area of peak activity, displayed higher response rates suggesting that a loss of landscape architecture is linked to a loss of BH4 responsiveness.

Group 3 genotypes with peak activity were predominantly associated with MHP or mild PKU and high response rates to cofactor treatment. This was not the case for p.[Leu48Ser]; [Gln355_Tyr356insGlyLeuGln], p.[Ile65Thr];[Gln355_Tyr356insGlyLeuGln], p.[Arg261Gln];[Arg408Trp], p.[Arg261Gln];[Arg261Gln], p.[Pro281Leu];[Arg408Trp], p.[Val245Ala];[Arg408Trp] and p.[Ala403Val];[Arg408Trp], with the predominance of classical PKU and unresponsiveness to pharmacological treatment. Interestingly, three of these genotypes contained p.Gln355_Tyr356insGlyLeuGln and three of these genotypes were associated with peak activities at increased cofactor concentrations. Therefore, dosages used in standard response tests may not be sufficient to meet the high need for BH4 of patients carrying one of these genotypes.

Taken together, the results from this study indicate that residual PAH activity alone is not a reliable predictor of the clinical phenotype but other factors may be taken into account.
Our analyses led to the recognition of the following orienteering rules for the implementation of personalised medicine strategies. (1) Low residual activity is associated with mostly BH₄ unresponsive classical PKU. (2) Lack of a defined area of peak activity leads to a loss of response to BH₄. (3) The occurrence of inconsistent clinical phenotypes with low rates of BH₄ response is linked to a higher need for BH₄. (4) Residual activity >5%, a defined peak of activity, and normal need for BH₄ are associated with mostly BH₄ responsive mild PKU or MHP. Moreover, the occurrence of the allele p.Arg261Gln is often associated with inconsistent phenotypes, whereas the allele p.Gln355_Tyr356insGlyLeuGln is linked to a higher cofactor need.

**DISCUSSION**

The healthcare practitioner’s view on genetic disease underwent change during the last decade. Genotyping has become easier and databases provide access to a wealth of data related to genetic variation. In addition, an increasing number of therapies evolved from an improved understanding of the molecular basis of maladies. After diagnosis of an inherited disease, families request reliable prediction of the clinical phenotype and the long-term health benefits of therapy. Figure 3 shows the activity landscapes for Group 1 genotypes, which are associated with a PAH residual activity of <3.5% of the wild-type. Figure 4 depicts the activity landscapes for Group 2 genotypes, which have a residual activity of ≥3.5% and <5% of the wild-type.
Figure 5  Group 3 phenylalanine hydroxylase (PAH) activity landscapes. The interpolated residual enzyme activities after expression of homozygous and compound heterozygous PAH genotypes were colour-coded and given as a function of different L-phenylalanine and BH4 concentrations. Group 3 refers to genotypes associated with a PAH residual activity ≥5% of the wild-type.
term outcome. In the specific case of PKU, the approval of a drug providing the first pharmacological therapy for the disorder\(^7\) added a new dimension to diagnosis and treatment. Clinical genetics at the service of patients faces the challenge of combining empirical evidence buried in databases with functional evidence from basic research. To this end, an experimental basis to assess molecular consequences of genetic alteration and to generate solid functional data considering full genotypes is required.

In this study, we aimed to provide a tool for the thorough analysis of genotype-related PAH function for PKU patients in Europe and the Middle East. A comprehensive literature search revealed that the 30 most frequent genotypes represent 55% of patients. Thus, the experimental investigation of a convenient number of genotypes can cover significant shares of a given population and thus provide meaningful information to physicians in clinical practice.

Residual enzyme activities reported for a number of single PAH mutations\(^16\) do not reflect the patient’s situation since values well above 50% are not expected to produce a biochemical and clinical phenotype. A study investigating in vivo PAH function provided evidence for a threshold, where enzyme activities >15% were linked to normal biochemical parameters and thus normal function.\(^2\) Moreover, it is well known that consequences arising from full genotypes consisting of both alleles differ from the average induced by the two mutations involved and that proteins arising from different alleles influence each other—a phenomenon termed interallelic complementation.\(^26\)–\(^28\) Furthermore, although residual PAH enzyme activity assessed at fixed substrate and cofactor conditions for single mutations shows a general correlation with clinical severity and BH\(_4\) responsiveness,\(^29\) there is still significant inconsistency that hampers solid phenotype prediction in clinical care. This prompted us to develop a model system that mimics the cellular situation in carriers of both homozygous and compound heterozygous genotypes. Our system was calibrated at the equilibrium situation in carriers of both homozygous and compound heterozygous genotypes. The carrier of such a genotype may need to keep blood phenylalanine concentrations in a tight range to optimise phenylalanine flux and could thus benefit from avoidance of phenylalanine fluctuations.

In the liver, the PAH enzyme is subjected to fluctuations of phenylalanine and, under treatment, of BH\(_4\). In order to reflect the situation in vivo, we analysed PAH activity landscapes for full genotypes at a wide range of substrate and cofactor concentrations. These conditions assessed PAH function in the context of metabolic and therapeutic states and mirrored major molecular mechanisms underlying missense mutation-induced PKU; however, data obtained in vitro may still differ from residual activity in patients.

We observed several patterns of changes in the architecture of PAH activity landscapes and the combined evidence from functional analyses with biochemical, clinical, and therapeutic data delivered means to generate hypotheses about the mechanisms behind individual genotypes. Changes in PAH activity landscapes as regards to substrate concentrations may impact dietary management. Some genotypes (p.[Arg261Gln]; [Arg261Gln], p.[Arg261Gln];[Arg408Trp], p.[Arg261Gln]; [Gln355_Tyr356insGlyLeuGln], p.[Arg297His];[Arg408Trp]) induced a right-shift of the activity landscape. As a consequence, these genotypes are associated with very low enzyme activity at low phenylalanine concentrations. On the one hand, the enzyme is inactive at phenylalanine concentrations <200 \(\mu\)M. On the other hand, the right-shift results in good metabolic PAH function at higher phenylalanine concentrations (peak activity, ~800 \(\mu\)M phenylalanine). Therefore, internationally accepted target phenylalanine concentrations <360 \(\mu\)M in the patient may in these cases be difficult to achieve. In turn, phenylalanine tolerance may be disproportionately higher at phenylalanine concentrations slightly above the target, lowering the burden of diet. This has to be confirmed by clinical studies with appropriate analysis of dietary phenylalanine intake and blood phenylalanine concentrations. Notably, different dietary regimens leading to different metabolic states may at least in part explain the significant phenotypic variability of homozygous p.Arg261Gln and p.[Gln355_Tyr356insGlyLeuGln];[Arg408Trp] with respect to disease severity and BH\(_4\) responsiveness (http://www.biopku.org) (see online supplementary figure S5).\(^20\) In the experimental dataset presented here, the presence of the mutation p.Arg261Gln in the genotype often induced broadening of the PAH working range, with rather high enzyme activities at a wide range of phenylalanine concentrations. This may be associated with fewer phenylalanine fluctuations, known to be a major factor of brain damage.\(^33\) Eight genotypes showed a narrow range of \(\Delta\) phenylalanine concentrations associated with considerable PAH activity. In three cases this was accompanied by a shift of the area of peak PAH activity below 300 \(\mu\)M \(\Delta\) phenylalanine. The carrier of such a genotype may need to keep blood phenylalanine concentrations in a tight range to optimise phenylalanine flux and could thus benefit from avoidance of phenylalanine fluctuations.

Another group of genotypes induced changes in PAH activity landscapes as regards to cofactor concentrations that may have implications not only for pharmacological treatment with BH\(_4\). Some genotypes showed PAH activity peaks at lower BH\(_4\) concentrations. Among these, the presence of the p.Arg261Gln mutation was associated with inconsistent BH\(_4\) response, whereas the mutation p.Arg408Trp led to consistent BH\(_4\) response. An activity peak at increased BH\(_4\) concentrations was associated with the occurrence of p.Gln355_Tyr356insGlyLeuGln. Moreover, genotypes in this group also harboured mutations typically associated with inconsistent clinical phenotypes (p.Leu48Ser, p.Ile65Thr, p.Ala300Ser). Patients carrying these genotypes may need more BH\(_4\) to achieve their optimal enzyme function. As a consequence BH\(_4\) dosages used in standard BH\(_4\) response tests may be insufficient to achieve an increase of enzyme activity and decrease of phenylalanine concentrations, the endpoint in response tests. Interestingly, genotypes with loss of landscape architecture p.[Gln355_Tyr356insGlyLeuGln];[Gln355_Tyr356insGlyLeuGln], p.[Pro281Leu];[Gln355_Tyr356insGlyLeuGln] also contained p.Gln355_Tyr356insGlyLeuGln and are associated with severe BH\(_4\) non-responsive phenotypes. The apparent affinity of PAH to BH\(_4\) in eukaryotic cells is \(~30\) \(\mu\)M and peak PAH activity was reached at 130 \(\mu\)M in our setting, which is in line with values obtained from recombinant PAH.\(^30\) In addition, liver cells keep BH\(_4\) at a concentration of about 9 \(\mu\)M.\(^34\) Therefore, the PAH system suffers from a constant shortage of BH\(_4\). Oxidative stress\(^35\) or other environmental factors with negative impact on BH\(_4\) availability may trigger PAH deficiency and thus may further aggravate clinical phenotypes. Taken together, alterations in BH\(_4\) demand may add to the molecular mechanisms underlying phenotypic variability related to both disease severity and BH\(_4\) responsiveness.

Our results confirm the frequent observation that residual PAH enzyme activity is not only the determinant as to the severity of the biochemical phenotype but also a prerequisite for BH\(_4\) response and thus correlates with it.\(^36\) However, this correlation is not absolute. For instance, in group 3 genotypes with residual enzyme activity >5%, we observed a significant proportion of genotypes associated with little or no BH\(_4\) response. On the one hand, they were associated with relatively low residual activity
within this group; on the other hand, lack of response to cofactor treatment was linked to a higher cofactor need. The definition of three groups provides a classification of genotypes with different overall PAH enzyme activity that allows for a global estimation of the severity of the disease. Beyond that, our study revealed that PAH activity landscapes not only provide information on the quantitative value of residual activity but also on the position of enzyme activity in the metabolic and therapeutic space, and that this is another important determinant of the phenotype.

Inconsistent phenotypes in individuals carrying identical genotypes are well known in clinical practice. In Leuders et al,36 five patients carrying the genotype p.Leu48Ser;[Leu48Ser] all showed responsiveness at a low BH4 dosage. In databases, four cases with non-responsive PAH deficiency (9%) and four cases with slow response to BH4 (9%) are reported for individuals carrying this genotype (see online supplementary figure S3). These observations, however, are not necessarily in contradiction to each other. We observed a higher need for BH4 to achieve peak activity for three genotypes containing the mutation p.Leu48Ser. Notably, Leuders et al36 tested BH4 response after 7 days of treatment. Repeated administration of BH4 may lead to constantly elevated concentrations of cellular BH4 and thus provide the variant PAH protein with a sufficient amount of cofactor, which may explain the consistent positive response. On the other hand, inconsistent BH4 responsiveness has previously been reported for genotypes containing the p.Arg261Gln mutation.37 38 Following our observation of a right-shifted area of peak activity with respect to phenylalanine values for p.[Arg261Gln];[Arg408Trp], the variability in the phenotype of patients carrying this genotype may, at least in part, be due to fluctuations in substrate concentrations, particularly with regard to blood phenylalanine concentrations at the time of the BH4 loading test. As a conclusion for these observations, the assessment of PAH activity landscapes testing PAH function revealed how varying metabolic and therapeutic conditions may contribute to variability in the clinical phenotype. Nevertheless, other genetic and non-genetic factors39 may also influence phenylalanine clearance and thus the clinical phenotype in PKU patients. Among these, differences in the disposal of excess phenylalanine by transamination when the hydroxylation reaction is blocked, variation in the carrier-mediated uptake of phenylalanine by the brain and the liver, and environmental factors such as oxidative stress are the best recognised modifiers to date.25 33 35 40–42

In conclusion, this work may exemplify how application of molecular genetics translates into medical practice. It gives new significance to gene analysis of the PAH locus in clinical routine and shows the importance of considering full genotypes. The analysis of activity landscapes in the metabolic and therapeutic space provided clinically relevant new insights into genotype-related impaired PAH function beyond the known link between residual activity and clinical phenotype. In the presence of specific genotypes we observed different patterns of activity landscape architecture and established orienteering rules for their interpretation. In addition, the work provides a web-based tool to assist clinicians in clinical care of PKU patients (http://pah-activitylandscapes.org). The web application allows for retrieving country- and region-specific information on genotypes and their related PAH function. Its sustainability and increasing representativeness will be ensured by continuous extension and updates as well as research on demand delivering new PAH activity landscapes upon request.

The analysis of individual genotype-related PAH activity landscapes together with the information provided on genotype-related clinical phenotypes may permit improved long term phenotype prediction and the implementation of personalised medicine strategies for dietary regimens and pharmacological treatment.

Acknowledgements We thank Heidi Noll-Puchta, Isabell Schönwasser, and Christoph Siegel for fruitful scientific support. We wish to thank Michael Staudigl for fruitful scientific discussion.

Contributors ACM, FF and IZ contributed data on genotypes. ACM, MKD and SWG designed the study. MKD and MW performed functional analyses and generated activity landscapes. ACM, MKD and SWG analysed data and wrote the manuscript. IZ contributed to the final version of the manuscript.

Funding This work was supported by the Bavarian Genome Research Project (Junior Project #3).

Competing interests ACM and FF are members of Merck Serono advisory boards and ACM is currently conducting research sponsored by this company. SWG is currently conducting research sponsored by Merck Serono. The research presented in this specific manuscript was not sponsored by the company.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


