



**University of Dundee**

## **Identification of prolactin receptor variants with diverse effects on receptor signalling**

Gorvin, Caroline M; Newey, Paul J; Thakker, Rajesh V

*Published in:*  
Journal of Molecular Endocrinology

*DOI:*  
[10.1530/JME-22-0164](https://doi.org/10.1530/JME-22-0164)

*Publication date:*  
2022

*Licence:*  
CC BY

*Document Version*  
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

*Citation for published version (APA):*  
Gorvin, C. M., Newey, P. J., & Thakker, R. V. (2022). Identification of prolactin receptor variants with diverse effects on receptor signalling. *Journal of Molecular Endocrinology*. <https://doi.org/10.1530/JME-22-0164>

### **General rights**

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Identification of prolactin receptor variants with diverse effects on receptor signalling**

(79 characters, 85 characters allowed)

Caroline M. Gorvin<sup>1,2,3,4</sup>, Paul J. Newey<sup>1,5</sup>, Rajesh V. Thakker<sup>1,2</sup>

**Affiliations:**

<sup>1</sup>Academic Endocrine Unit, Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, OX3 7LJ, UK

<sup>2</sup>Oxford NIHR Biomedical Research Centre, University of Oxford, Churchill Hospital, Oxford, OX3 7LE, UK

<sup>3</sup>Institute of Metabolism and Systems Research (IMSR) & Centre for Endocrinology, Diabetes and Metabolism (CEDAM), Birmingham Health Partners, University of Birmingham, Birmingham, B15 2TT, UK

<sup>4</sup>Centre of Membrane Proteins and Receptors (COMPARE), University of Birmingham, Birmingham, B15 2TT, UK

<sup>5</sup>Division of Molecular & Clinical Medicine (MCM), University of Dundee, Jacqui Wood Cancer Centre, James Arnott Drive, Dundee, DD1 9SY, UK

\*Address correspondence to Professor R.V. Thakker at the Radcliffe Department of Medicine, University of Oxford, Oxford Centre for Diabetes, Endocrinology and Metabolism (OCDEM), Churchill Hospital, Oxford, OX3 7LJ, United Kingdom, or at [rajesh.thakker@ndm.ox.ac.uk](mailto:rajesh.thakker@ndm.ox.ac.uk)

**Short title** (45 characters, 46 characters allowed, including spaces): Prolactin receptor variants affect signalling

**Keywords (4 allowed):** Hyperprolactinaemia, Prolactinoma, Proliferation, JAK-STAT5

**Word count (5000 word limit, excluding references & figure legends):** 4958 words

**Abstract (248 words, 250 words allowed)**

The prolactin receptor (PRLR) signals predominantly through the JAK2-STAT5 pathway regulating multiple physiological functions relating to fertility, lactation, and metabolism. However, the molecular pathology and role of PRLR mutations and signalling is incompletely defined, with progress hampered by a lack of reported disease-associated PRLR variants. To date, two common germline *PRLR* variants are reported to demonstrate constitutive activity, with one, Ile146Leu, overrepresented in benign breast disease, whilst a rare activating variant, Asn492Ile, is reported to be associated with an increased incidence of prolactinoma. In contrast, an inactivating germline heterozygous *PRLR* variant (His188Arg) was reported in a kindred with hyperprolactinaemia, whilst an inactivating compound heterozygous *PRLR* variant (Pro269Leu/Arg171Stop) was identified in an individual with hyperprolactinaemia and agalactia. We hypothesised that additional rare germline *PRLR* variants, identified from large-scale sequencing projects (ExAC and GnomAD) may be associated with altered *in vitro* PRLR signalling activity. We therefore evaluated >300 previously uncharacterised non-synonymous, germline *PRLR* variants and selected ten variants for *in vitro* analysis based on protein prediction algorithms, proximity to known functional domains and structural modelling. Five variants, including extracellular and intracellular domain variants were associated with altered responses when compared to the wild-type receptor. These altered responses included loss- and gain-of-function activities related to STAT5 signalling, Akt and FOXO1 activity, as well as cell viability and apoptosis. These studies provide further insight into PRLR structure-function and indicate that rare germline *PRLR* variants may have diverse modulating effects on PRLR signalling, although the pathophysiologic relevance of such alterations remain to be defined.

## Introduction

The prolactin receptor (PRLR), and its ligand, the hormone prolactin (PRL) are reported to have diverse roles that include induction and maintenance of lactation in the peripartum and postpartum periods (Ben-Jonathan, et al. 2008), parental behaviour, immune function, reproduction and metabolic functions such as pregnancy-related increases in  $\beta$ -cell mass, and regulation of lipid content and body temperature (Banerjee, et al. 2016; Freemark, et al. 2002; Huang, et al. 2009; Schuff, et al. 2002; Smiley, et al. 2022; Viengchareun, et al. 2008; Viengchareun, et al. 2004). PRL binds to the PRLR, a class I cytokine receptor, which is functionally active as a homodimer (Brooks, et al. 2014; Gadd and Clevenger 2006; Qazi, et al. 2006). Each mature PRLR monomer has a multi-domain structure consisting of a highly-conserved ligand-binding extracellular domain (ECD, 1-210), a transmembrane  $\alpha$ -helix (TM, residues 211-234) and an intracellular domain (ICD, residues 235-598). Structural analysis of the ECD has revealed two subdomains, designated D1 and D2, which are important in ligand binding and subsequent receptor activation (Brooks 2012; Broutin, et al. 2010; Rao and Brooks 2011; Svensson, et al. 2008; van Agthoven, et al. 2010), and the WSxWS motif, which acts as a molecular switch during activation (Broutin et al. 2010; Dagil, et al. 2012; van Agthoven et al. 2010).

The ICD interacts with the Janus kinase-2 (JAK2) protein, that associates with a highly-conserved structural motif named Box 1 (residues 243-251), and potentially with a second motif, Box 2 (residues 287-296) within the ICD (Lebrun, et al. 1995). Moreover, recent structural studies have revealed both PRLR and the related growth hormone receptor (GHR), harbour conserved regions that interact with lipids, referred to as lipid-interacting domains or LIDs within the inner plasma membrane leaflet thereby allowing a greater surface area, and potentially simultaneous interaction with multiple signalling kinases (Bugge, et al. 2016; Haxholm, et al. 2015). Hormone binding to the ECD activates conformational changes within the TMD and ICD, allowing separation of the ICDs and initiates phosphorylation cascades downstream of JAK2 (Brooks et al. 2014; Brown, et al. 2005; Bugge et al. 2016; Haxholm et al. 2015). JAK2 activates complex signalling pathways, predominantly via interaction with the signal activator of transcription 5 (STAT5) pathway (Brooks 2012); but also by the phosphatidylinositol 3-kinase (PI3K)/Akt and mitogen-activated protein kinase pathways (Amaral, et al. 2004; Brooks 2012; Fresno Vara, et al. 2000). These signal pathways lead to transcription of target

genes that regulate proliferation, differentiation and cell survival (Amaral et al. 2004; Brooks 2012; Fresno Vara et al. 2000). Despite these findings, the function of individual residues in receptor activation and signal transduction is poorly understood.

A number of human studies have highlighted residues that are important for receptor function and subsequent signal transduction. The PRLR variants Ile76Val and Ile146Leu were reported to be gain-of-function variants with constitutive activity that occur in 15% of women with breast fibroadenomas (Bogorad, et al. 2008; Courtillot, et al. 2010). However, more recent studies did not detect such correlations (Glasow, et al. 2001; Lee, et al. 2007; Nyante, et al. 2011; Vaclavicek, et al. 2006), or marked changes in signalling activity (Bernard, et al. 2016; Gorvin, et al. 2018b), although the Ile146 residue has been shown to be important for receptor folding and stability (Dagil et al. 2012; Zhang, et al. 2015). A loss-of-function pathogenic germline *PRLR* variant (His188Arg), which affected a highly conserved His188 residue within the D2 domain that is important for hormone binding was described in a family with hyperprolactinaemia (Kulkarni, et al. 2010; Newey, et al. 2013). Subsequently, an individual with hyperprolactinaemia and agalactia was reported with germline compound heterozygous nonsense (Arg171Stop) and missense (Pro269Leu) *PRLR* variants (Kobayashi, et al. 2018), and recently, a germline Asn492Ile *PRLR* variant that increases receptor activity via the PI3K-Akt pathway was reported to be associated with a higher incidence of prolactinoma (Gorvin et al. 2018b).

Both the Ile146Leu and Asn492Ile variants are present in recently described population databases such as the Exome Aggregation Consortium (ExAC) and the Genome Aggregation Database (GnomAD)(Karczewski, et al. 2020), and we hypothesised that a further examination of these population-based databases, could yield important structural and functional insights for individual *PRLR* residues and provide information on activation of specific signalling pathways. Indeed, similar studies have previously identified residues within the adaptor protein-2 sigma subunit that are important for calcium homeostasis (Gorvin, et al. 2018a), and residues in  $\alpha$ -N-acetylglucosaminidase that contribute to the rare lysosomal storage disease Sanfilippo Type-B (Clark, et al. 2018). We therefore examined the ExAC/GnomAD datasets with the aim of identifying missense coding variants in the *PRLR*, which could be characterised for their functional consequences.

## Materials and Methods

### Protein sequence alignment and three-dimensional modelling of PRLR structure

The population frequencies of germline non-synonymous *PRLR* single nucleotide variants (SNVs) were evaluated using ExAC and Genome Aggregation Databases (both datasets (ExAC and GnomAD v.2.1) now reported at GnomAD, <https://gnomad.broadinstitute.org/>)(Karczewski et al. 2020)). SIFT, MutationTaster, Polyphen-2 and REVEL were used to predict the effect of amino acid substitutions (Adzhubei, et al. 2010; Ioannidis, et al. 2016; Kumar, et al. 2009; Schwarz, et al. 2014). Amino acid conservation was examined in PRLR mammalian orthologs using *ClustalW2* (Larkin, et al. 2007). The crystal structure of the two chains of the rat PRLR extracellular domain in complex with PRL (Protein Data Bank (PDB) accession code 3NPZ) and the NMR structure of the human PRLR ECD D2 domain (PDB:2LFG)(Dagil et al. 2012; van Agthoven et al. 2010) was used to predict the effect of ECD variants on PRLR structure. The NMR structure of the single-pass transmembrane domain of PRLR (PDB:2N7I)(Bugge et al. 2016) was used to predict the structural effect of TMD variants. Figures were prepared using the PyMOL Molecular Graphics System (Schrodinger, LLC).

### Cell culture and transfection

PRLR variants were introduced into the wild-type (WT) pdEYFP-PRLR construct by site-directed mutagenesis using the Quikchange Lightning Kit (Agilent Technologies, Santa Clara, CA) and gene-specific primers (Sigma, Gillingham, UK), and confirmed, as previously described (Newey et al. 2013). Expression constructs were transiently transfected into HEK293 cells, maintained in DMEM-Glutamax media (Gibco) with 10% fetal bovine serum (Gibco, Waltham, MA) at 37°C, 5% CO<sub>2</sub>, using Lipofectamine 2000 (LifeTechnologies), as described (Newey et al. 2013), and functional studies performed using poly-L-lysine treated plates. Western blot analysis was used to assess expression of transfected PRLR and endogenous  $\alpha$ -tubulin as a loading control, using anti-PRLR (1:1000, SantaCruz Biotechnology, SantaCruz, CA) and anti- $\alpha$ -tubulin (1:1000, Abcam, Cambridge, UK) antibodies, as described (Newey et al. 2013).

### Confocal microscopy

Confocal imaging was performed as previously described (Gorvin et al. 2018b). Cells were plated in six-well plates containing poly-L-lysine-treated coverslips and cultured at 37°C. Cells were transiently transfected with 1000 ng of either WT or variant PRLR expression constructs. After 24 hours cells were fixed in 4% paraformaldehyde/PBS (SigmaAldrich), permeabilised in 1% Triton-X100/PBS (Thermo Scientific), and immunostained with primary anti-PRLR (1:200, Santa Cruz) and secondary antibody Alexa Fluor-488 (1:1000, Molecular Probes, Waltham, MA). Cells were mounted in Prolong Gold Antifade reagent (Invitrogen, Waltham, MA). Images were captured using a Zeiss LSM780 confocal microscope with a Plan-Apochromat x63/1.2/water DIC objective. An argon laser (488 nm) was used to excite Alexa Fluor-488.

### **AlphaScreen SureFire assays**

AlphaScreen assays were performed as previously described (Gorvin et al. 2018b; Newey et al. 2013). Cells were transiently transfected in 48-well plates with 200 ng of either WT or variant PRLR vectors. After 30 hours, cells were incubated in serum-free media for 12 hours prior to treatment with human recombinant PRL (PromoCell, Heidelberg, Germany) for 20 minutes at concentrations ranging from 0-1000 ng/mL. Cells were lysed in Surefire lysis buffer and AlphaScreen Surefire pSTAT5 or pAkt assays (PerkinElmer, Beaconsfield, UK), performed according to manufacturer's instructions (Binder, et al. 2008). The fluorescence signal was measured using a PHERAstar FS microplate reader (BMG Labtech, Aylesbury, UK). A minimum of 4 independently transfected replicates were used for each construct within each experiment, and each experiment performed on 4-5 separate occasions with different cell passages. Data was plotted as fold-change responses relative to the response at 0ng/mL in cells expressing the WT PRLR expression construct. Statistical analyses were performed using two-way ANOVA with Dunnett's or Tukey's multiple-comparisons tests for pSTAT5 studies and by one-way ANOVA with Sidak's multiple-comparisons tests for pAkt studies

### **Luciferase reporter assays**

The Forkhead box O1 (*FOXO1*) promoter region was PCR amplified from human genomic DNA using previously described primers (Essaghir, et al. 2009) and cloned into the pGL4.10 vector

(Promega). The sequence of the insert was confirmed by Sanger DNA sequencing (Source Bioscience, Nottingham, UK). The pGL4.10 vector containing the cytokine inducible SH2-containing protein (*CISH*) reporter has been described previously (Newey et al. 2013). HEK293 cells were transiently co-transfected in 24-well plates with 100 ng of pGL4.10-*CISH* reporter gene construct, 10ng of pRL (renilla) control vector, and 100ng of WT or variant PRLR vectors. Following transfection, cells were incubated in serum-free media overnight. Cells were then treated with 0-500ng/mL PRL for 24 hours in serum-free media. Cells were lysed and assayed for luciferase activity using a Turner Biosystems (Promega, Southampton, UK) or Centro LB960 (Berthold Technologies, Harpenden, UK) luminometer, and the Dual-Luciferase Reporter assay system (Promega). The firefly luciferase activity was adjusted for Renilla luciferase activity (Firefly/Renilla ratio) and ratios expressed as a fold-change relative to cells treated with 0ng/mL PRL within each group. A minimum of 4 independently transfected replicates were performed in each experiment, and each experiment was performed on 4-7 separate occasions with different cell passages. Statistical analysis was performed by two-way ANOVA with Sidak's or Dunnett's multiple-comparisons test for *CISH* and by Kruskal-Wallis with Dunn's test or one-way ANOVA with Dunnett's test for *FOXO1*.

### **Cell viability assay**

Cells were plated in 96-well plates and transfected with 50ng WT or variant PRLR per well. Following 24-hours, cells were treated with 200 ng/mL PRL and cell viability assessed 96 hours later using the CellTiter Blue kit (Promega)(Gorvin et al. 2018b). The cell count for day one (i.e. time 0 before PRL was added) was set as 1 and each cell count was expressed relative to this original cell count. Plates were read on a PHERAstar *FS* microplate reader (BMG Labtech). A minimum of 4 independently transfected replicates were performed in each experiment, and each experiment was performed on 4 separate occasions with different cell passages. Statistical analysis was performed by one-way ANOVA with Dunnett's test or Kruskal-Wallis with Dunn's test.

### **Apoptosis assay**



Cells were plated in 96-well plates and transfected with 50ng WT or variant PRLR vectors per well. Following 24-hours, cells were treated with 0 ng/mL or 200 ng/mL PRL and apoptosis assessed at 0 hours and 96 hours post-PRL treatment using the Caspase-Glo-3/7 kit (Promega) (Gorvin et al. 2018b). Plates were read on a Centro LB960 luminometer. A minimum of 4 independently transfected replicates were performed in each experiment, and each experiment was performed on 4-6 separate occasions with different cell passages. Statistical analysis was performed by one-way ANOVA with Dunnett's multiple comparisons test.

### **Statistics**

The number of experimental replicates denoted by n is indicated in each figure legend. Data was plotted and statistical analyses performed in Graphpad Prism 7. Normality tests (Shapiro-Wilk or D'Agostino-Pearson) were performed on all datasets to determine whether parametric or non-parametric tests were appropriate. A p-value of <0.05 was considered statistically significant. A minimum of 4 independently transfected replicates was performed in all cell-based assays, and each experiment was performed on separate occasions with different passages of cells. Specific details of each test are outlined in the figure legends and within the relevant methods section.

## Results

### Identification of non-synonymous PRLR variants and their predicted effects on protein function

An analysis of the ExAC (v1.0) and GnomAD (v2.1.1) databases was performed to identify non-synonymous, missense *PRLR* variants in the full-length, membrane-expressed (i.e. excluding the 24 amino acid signal peptide) protein. These analyses revealed 310 non-synonymous missense *PRLR* variants comprising 85 ECD variants, 10 TMD variants and 215 ICD variants. The distribution of variants between the ECD, TMD and ICD was 27.4%, 3.2% and 69.4% respectively. This was significantly different to that expected based on the size of each region ( $p < 0.05$ ,  $\chi^2$ -test), with fewer variants observed in the ECD than expected (37.7%), indicating this region may be less tolerant to variation.

The predicted deleteriousness/pathogenicity of each *PRLR* variant was determined by assessing their population frequency; their effect on protein function using four online prediction software packages (SIFT, Polyphen-2, MutationTaster and REVEL (Adzhubei, et al. 2013; Ioannidis et al. 2016; Kumar et al. 2009; Schwarz et al. 2014)), and the evolutionary conservation of each residue in mammalian species. Following exclusion of variants that have previously been functionally expressed (e.g. Gly57Ser, Ile76Val, Ile146Leu, Gly376Gln, Asn492Ile, and Glu554Gln (Bernard et al. 2016; Bogorad et al. 2008; Gorvin et al. 2018b; Newey et al. 2013)), examination of the minor allele frequencies (MAF) of all the remaining variants revealed them to be rare (defined as a MAF of  $< 1\%$  (Agarwala, et al. 2013)). Variants that were predicted benign or tolerated by all the prediction programs, and those that affected residues that were conserved in fewer than 2 mammalian orthologues, were excluded from further analyses. Thus, 42 ECD, 5 TMD and 93 ICD variants (Table 1) were predicted to be potentially deleterious/pathogenic by these criteria, indicating they may have functional consequences on *PRLR* signalling.

### Structural characterisation of the ECD *PRLR* variants

Three-dimensional modelling using the crystal structure of the PRL-bound homodimeric ECD and the NMR structure of the ligand-free D2 domain (Dagil et al. 2012; van Agthoven et al. 2010) was performed to determine the locations of the ECD *PRLR* variants and predict their effects on structural

integrity. This revealed that 23 of the total 42 variants were within the D1 lobe and the remaining 19 variants were within the D2 lobe, of which 11 variants (6 in D1 and 5 in D2) were predicted to have structural effects on the PRLR protein (Table 1, Figure 1A). Three variants (Pro7Ser, Gly30Arg, Cys51Tyr) are predicted to gain contacts with adjacent residues, six variants (Lys6Asn, Tyr99His, Thr141Met, Glu155Lys, Arg183His, Asp187Glu) are predicted to lose contacts with adjacent residues, and two variants (Ser88Arg, Glu145Asp) are predicted to both lose and gain contacts with adjacent PRLR residues (Figure 1B-D and 2, Table 1). These predicted changes may affect ECD flexibility and disrupt PRLR activation.

Several variants were predicted to affect important structural components in the C-terminal region of D1 and in D2 of the PRLR ECD (Figure 2). The Tyr99His variant, in D1, is predicted to retain a contact with the PRL ligand, but at a more distal site (i.e further away from the PRL  $\alpha$ -helix), and therefore is likely to increase the distance between the PRL and PRLR molecules, which may affect agonist binding or PRLR activation (Figure 2A, Table 1). Two D2 variant residues, His183 and Asp145, are predicted to disrupt contacts between the wild-type Arg183 and Glu145 residues. The Glu145 and Arg183 PRLR residues are located in adjacent  $\beta$ -strands, and form four direct contacts (Figure 2B). The variants Asp145 and His183 are predicted to disrupt two of these contacts and may increase the flexibility of the D2 lobe (Figure 2B). Additionally, the Asp145 variant is predicted to lose a contact with the neighbouring Ile146, a residue previously demonstrated to be important for PRLR folding and stability (Dagil et al. 2012; Zhang et al. 2015) (Figure 2B); while His183 loses a contact with Ala193, a residue that forms part of the highly-conserved WSxWS motif (Figure 2C). Arg183 is one of five highly-conserved residues (Ile146, Glu151, Glu155, Tyr178, Arg183) that together with residues of the Trp-Arg ladder undergo conformational changes to switch PRLR to an active state (Dagil et al. 2012). Rare variants were identified to affect two more of these five residues (Glu151Lys, Glu155Lys) (Table 1). Moreover, the variant Lys155 loses a contact with Lys114 on the adjacent PRLR protomer, which may affect receptor dimerisation and protein stability (Figure 2D, Table 1). The Asp187Glu variant lies close to the PRL binding site and is predicted to lose a contact with the adjacent His188 residue, which has a critical role in ligand binding (Kulkarni et al. 2010) and is mutated in some individuals with

hyperprolactinaemia (Newey et al. 2013) (Figure 2E, Table 1). Thus, the PRLR D2 domain mutations are predicted to affect PRL binding, flexibility of the PRLR structure and PRLR activation.

### **Structural characterisation of the TMD and ICD PRLR variants**

The NMR structure of the single-pass transmembrane domain of the PRLR (PDB:2N7I (Bugge et al. 2016)) was used to study the structural effects of the 5 highly-conserved TMD PRLR variants (Ala222Gly, Ile227Thr, Trp230Cys, Val232Ala, Val232Met) that are predicted to be potentially deleterious/pathogenic (Figure 3A). Three-dimensional modeling did not predict any changes in interactions between the TMD located PRLR variants. However, the side chains of each amino acid project into the plasma membrane bilayer, and therefore the variant residues may affect interactions with lipids, which cannot be predicted using the current structural models. In addition to the TMD, plasma membrane interactions also occur with three regions of the ICD (LID1-3) (Bugge et al. 2016) (Figure 3B). Forty-five PRLR rare variants affect residues within these LID regions, 31 in LID1, 3 in LID2 and 11 in LID3 (Table 1). The ICD contains two other structural features, Box 1 and 2, the binding sites for JAK2 (Lebrun et al. 1995), and a conserved degradation motif, DSGxxS (located at residues 324-329) (Plotnikov, et al. 2009) (Figure 3C). Seven rare variants were identified in residues within Box 1, and two variants within Box 2, while three variants within the same residue were identified in the degradation motif.

### **Functional analysis of the PRLR rare variants**

Based on their predicted effects on pathogenicity, evolutionary conservation and structure we chose to assess ten PRLR rare variants. Five ECD variants (Tyr99His, Glu145Asp, Glu155Lys, Arg183His, Asp187Gly) (Figure 2A-E) were selected including ones within the PRL binding region, in the homodimeric interface and close to the WSxWS motif, and five ICD variants including, two residues within LID1 located between Box 1 and 2 (Phe255Ser, Gly263Asp); two residues close to, or within the degradation motif (Asp320Tyr, Arg327Gln); and one distal rare variant close to LID3 (Val535Met) (Figure 3C).

### ***Effect of rare variants on PRLR expression***

Initially, PRLR protein expression was assessed by Western blot analyses in HEK293 cells transiently transfected with WT or variant PRLR (Figure 4A-B). The PRLR protein was expressed at equivalent levels in cells transfected with the five ECD and five ICD rare variants and WT PRLR constructs (Figure 4A-B), indicating it is unlikely that any of the variants affect protein folding. Although structural analysis had predicted that the Glu155Lys variant may affect contacts between the two PRLR protomers, concentrations of PRLR dimers were similar in cells transfected with wild-type or the ECD variants (Figure 4A). Therefore, the Glu155Lys variant may not have a major impact on homodimer stability.

The cellular expression of each of the PRLR variants was then assessed using confocal microscopy. As previously observed, the WT PRLR protein is located within the cytoplasm and at the plasma membrane (Gorvin et al. 2018b). A similar expression pattern was observed for the ten PRLR rare variants (Figure 4C). Therefore, it is unlikely that the variant residues affect protein expression or trafficking of the PRLR to the cell surface.

### ***Effect of the PRLR rare variants on STAT5 signalling***

Previous studies have demonstrated PRLR to predominantly signal via the STAT5 pathway (Brooks 2012) and we therefore assessed this pathway by measuring immediate signalling by phospho-STAT5 (pSTAT5), and later downstream effects on transcription by measuring the STAT5 target gene *CISH*. The effects on PRLR signaling were assessed together with that of the His188Arg mutant PRLR that has been reported to result in a loss-of-function in association with familial hyperprolactinaemia (Bernard et al. 2016; Newey et al. 2013). We first assessed the ECD variants and demonstrated that increasing concentrations of PRL led to an increase in pSTAT5 and CISH luciferase reporter in a similar dose-dependent manner in cells expressing the PRLR variants Glu145Asp and Arg183His, and cells expressing WT PRLR (Figure 5A-B). However, responses were significantly reduced in cells expressing the PRLR ECD variants Tyr99His, Glu155Lys and Asp187Glu, when compared to those expressing WT PRLR (Figure 5A-B, Table 2).

The effects of the five PRLR ICD rare variants were next assessed on STAT5 signalling. Increasing concentrations of PRL led to an increase in pSTAT5 and CISH luciferase reporter in a similar dose-dependent manner in cells expressing the PRLR rare variants Gly263Asp, Asp320Tyr and Val535Met, and cells expressing WT PRLR (Figure 6A-B). In contrast, the Phe255Ser rare ICD PRLR variant significantly reduced pSTAT5 and CISH reporter responses, while the Arg327Gln variant had significantly elevated pSTAT5 and CISH reporter responses, when compared to wild-type expressing cells (Figure 6A-B, Table 2). There was consistently no response in cells expressing the His188Arg mutant protein to increasing concentrations of PRL (Figure 6, Table 2).

### ***Effect of the PRLR rare variants on Akt signalling***

PRLR can also signal by the Akt pathway, and we have previously demonstrated that some PRLR rare variants affect signalling by this pathway (Gorvin et al. 2018b). To assess the effects of the PRLR variants on Akt signalling, we investigated PRL-induced responses of phospho-Akt by AlphaScreen analysis and luciferase reporter activity by the Akt-target gene *FOXO1* (Essaghir et al. 2009). Exposure of four of the PRLR ECD variants (Tyr99His, Glu145Asp, Glu155Lys, Arg183His) to 200 ng/mL PRL led to an increase in pAkt activity, which was not significantly different to that observed in WT PRLR expressing cells (Figure 7A, Table 2). However, cells expressing the Asp187Glu rare variant were unable to induce increases in p-Akt in response to 200 ng/mL PRL (Figure 7A).

Akt phosphorylates FOXO proteins, resulting in their exclusion from the nucleus and subsequent degradation. Thus, PRL activation of the PRLR, which induces increases in Akt signalling, will reduce FOXO1 transcription (Essaghir et al. 2009). Assessment of *FOXO1* luciferase activity in cells expressing the WT or PRLR ECD variants showed that all cells could reduce *FOXO1* luciferase reporter activity in response to 200 ng/mL PRL (Figure 7B). However, cells expressing the Glu155Lys variant had lower basal expression of *FOXO1* luciferase activity (Figure 7B). In contrast, the His188Arg mutant, which has previously been shown to reduce pAkt activity (Gorvin et al. 2018b), was unable to reduce *FOXO1* luciferase reporter activity following exposure to 200 ng/mL PRL (Figure 7B), consistent with impaired pAkt activity.

Assessment of the ICD PRLR rare variants showed that four variants (Phe255Ser, Gly263Asp, Asp320Tyr, Val535Met) had similar PRL-induced pAkt responses to wild-type PRLR expressing cells (Figure 7C, Table 2). However, the Arg327Gln variant did not increase pAkt in response to 200 ng/mL PRL. This may have been a consequence of constitutively high basal pAkt concentrations in Arg327Gln expressing cells (Figure 7C, Table 2). None of the five ICD variants had a significant effect on PRL-induced FOXO1 responses compared to the wild-type PRLR (Figure 7D).

### ***Effect of the PRLR rare variants on cell viability and apoptosis***

Both the STAT5 and Akt signalling pathways lead to transcription of target genes that regulate proliferation and cell survival (Amaral et al. 2004; Brooks 2012; Fresno Vara et al. 2000), and previous studies have demonstrated that the Asn492Ile PRLR variant increases proliferation, while the His188Arg mutation increases apoptosis (Gorvin et al. 2018b). We therefore assessed the effect of the ten PRLR rare variants on cell viability using the CellTiter Blue assay, and on apoptosis using a Caspase-Glo-3/7 assay (Gorvin et al. 2018b). This demonstrated that all five ECD variants (Tyr99His, Glu145Asp, Glu155Lys, Arg183His, Asp187Glu) and three ICD variants (Gly263Asp, Asp320Tyr, Val535Met) had a similar effect on cell viability when compared to cells expressing WT PRLR, following exposure to 200 ng/mL PRL for 96 hours (Figure 8A-B, Table 2). The ICD Phe255Ser and Arg327Gln PRLR variants were associated with a significantly increased numbers of viable cells after 96 hours PRL treatment. Assessment of apoptosis was performed in wild-type and variant PRLR expressing cells after 96 hours of exposure to 200 ng/mL PRL. The ECD His188Arg loss-of-function mutation increased apoptosis in cells treated with 200 ng/mL PRL (Figure 8C), consistent with our previous report (Gorvin et al. 2018b). However, none of the other ECD (Figure 8C) or ICD (Figure 8D) PRLR variants had a significant effect on apoptosis. Therefore, both loss-of-function and gain-of-function mutations in the PRLR increased the number of viable cells, although none of these investigated rare variants affected apoptosis.

## Discussion

Evaluating the clinical significance of rare coding variants within genes associated with Mendelian disorders and complex traits represents a significant challenge, and consequently a range of *in silico* methods have been developed to facilitate the identification of potentially deleterious variants resulting in altered protein function. Our analysis of PRLR variants demonstrated that *in silico* tools could not accurately predict those that affected PRLR function (Table 2). This is consistent with previous studies that have shown that algorithms are only 65-80% accurate in predicting known disease variants (Thusberg, et al. 2011) as pathogenic, and often over-predict missense changes as deleterious, while they are unreliable in predicting variants with milder effects (Choi, et al. 2012). As such the American College of Medical Genetics and Genomics (ACMG) recommends that protein prediction software is not used as the sole source of information to make clinical decisions (Richards, et al. 2015). For the PRLR, only one variant, Phe255Ser, was correctly predicted deleterious in all four *in silico* methods (Table 1 and 2); while the gain-of-function Arg327Gln variant was predicted benign in 3 of 4 tools examined. This is consistent with previous studies of missense variants in the *ABCC8*, *GCK* and *KCNJ11* genes that showed SIFT and Polyphen to be better at predicting inactivating than gain-of-function mutations (Flanagan, et al. 2010). Furthermore, most *in silico* tools use the evolutionary conservation of the affected residue as a parameter to predict deleteriousness/pathogenicity (Richards et al. 2015). Such reliance on evolutionary conservation may be poorly predictive for PRLR variants, as the receptor has a specific role in lactation, and thus a lack of conservation with non-lactating species may be unimportant and could account for the poor predictive capability of *in silico* tools for PRLR.

Our functional studies identified five of ten PRLR germline variants that were associated with altered signalling, and these variants were all located in regions of the PRLR that have known receptor functions (Table 2). Thus, the three ECD variants that reduced PRLR function were predicted to affect ligand binding, receptor activation and homodimerisation, while the ICD variants are located close to the JAK2 binding site and a known degradation motif (Figure 1-3). Examination of the ExAc/GnomAD databases identified significantly fewer ECD variants than predicted to occur in a region of this size, and that there are more singleton variants (i.e. those identified in a single individual), which have previously been reported to have a higher probability of being functionally damaging and typically have



occurred recently in evolutionary terms (Tennesen, et al. 2012). These findings indicate that the ECD may be less tolerant of genetic variation, due to its critical roles in ligand binding and receptor activation, and that identification of variants in known functional domains is a reasonable predictor of possible pathogenicity. It is of note that several PRLR variants from other regions had similar responses to wild-type PRLR or were associated with, at most, modest effects on receptor function that may only be identified at supra-physiological concentrations of PRL (e.g. pSTAT5 responses for Gly263Asp and Val535Met).

The Glu155Lys variant is predicted to disrupt a contact formed across the homodimer interface and was associated with a partial loss-of-function for both STAT5 and pAkt signalling pathways (Figure 2, 5 and 7, Table 2). However, analysis of protein expression in Glu155Lys expressing cells showed no discernible difference in PRLR dimer or monomer concentrations (Figure 4), indicating that loss of this homodimeric interaction is not sufficient to impair dimer formation. However, the partial loss-of-function associated with this Glu155Lys variant indicates that this residue may have an important role in facilitating conformational changes across the dimer interface that are necessary for receptor activation. The Glu155 residue lies in proximity to the WSXWS motif, a highly conserved motif of cytokine receptors, that holds the receptor in an 'off-state', until ligand interaction occurs, inducing formation of a Trp-Arg ladder to activate the receptor (Dagil et al. 2012). The Glu155Lys variant may disrupt these conformational changes, with a consequent reduction in signalling as observed in functional studies (Table 2). The elucidation of the full-length PRLR structure in the active and inactive states could help resolve whether the Glu155Lys variant has such functional effects.

Within the ICD, the Phe255 residue is located in the region between the two JAK2 binding sites, within a series of residues that are unnecessary for JAK2 phosphorylation, but critical for downstream transcription of the beta-casein reporter gene (Ali and Ali 1998; Lebrun et al. 1995). We therefore hypothesise that Phe255 is involved in interaction with downstream signalling partners of PRLR that are necessary for pSTAT5 signalling, but not for pAkt signalling, which was unaffected by the variant (Figure 6 and 7, Table 2). Interestingly, the Phe255Ser variant was associated with increased numbers of viable cells, which we have previously observed for the gain-of-function Asn492Ile ICD variant (Gorvin et al. 2018b), and showed for the activating Arg327Gln variant in this study (Figure 8,

Table 2). Thus, both inactivating and activating ICD PRLR variants may be associated with increased cell survival and proliferation. It is likely that increased cell survival by the three receptor variants involves different mechanisms. The Asn492Ile variant had no effect on PRL-induced pSTAT5 signalling, but increased Akt signalling, which could be rectified by treatment with a PI3K inhibitor (Gorvin et al. 2018b). Thus, it is likely this variant increases cell survival by activating the Akt-PI3K signalling pathway. The Arg327Gln variant does not increase PRL-induced Akt signalling, but has constitutive Akt activity (Figure 7, Table 2), which may contribute to increased numbers of viable cells. Additionally, the Arg327Gln variant significantly increases STAT5 signalling, which has previously been shown to promote mammary cell proliferation (Iavnilovitch, et al. 2002). Thus, proliferation may be increased by both enhanced PRL-induced STAT5 signalling and constitutive Akt signalling. Furthermore, the Arg327 residue lies within the PRLR degradation motif (Plotnikov et al. 2009) and the Arg327Gln variant may impair receptor degradation. However, western blot analyses did not show enhanced protein expression, indicating that Arg327Gln is unlikely to affect protein turnover. The mechanism by which the Phe255Ser variant increases cell viability is unknown. It is possible that this residue enhances binding of negative regulators of PRLR signalling, such as the suppressors of cytokine signalling (SOCS) proteins (Tomic, et al. 1999), or may activate signalling pathways that are yet to be identified. Examination of proteins that regulate the cell cycle, including expression of cyclin D1 and transcription factors such as c-Myc, which are controlled by prolactin-mediated JAK-STAT and Akt signalling (Acosta, et al. 2003; Brockman, et al. 2002), could provide more insights into the different effects of PRLR variants on cell survival. The use of pathway-specific inhibitors, such as the Akt1/2 inhibitor previously used to examine the prolactinoma-associated Asn492Ile PRLR variant, may be required to further elucidate these mechanisms. Moreover, studies of additional PRLR ICD variants within the LID1 region may identify other residues with similar effects on signalling and proliferation.

Although these studies identified several inactivating PRLR variants, some activity was retained by all the variants, in contrast to the hyperprolactinaemia-associated His188Arg variant (Newey et al. 2013), which abolishes signalling. This is in keeping with the observation that the His188 residue occurs within the high-affinity ligand-binding interface and has a functional role in ligand binding and receptor activation (Kulkarni et al. 2010). The retention of some signalling activity may

also explain why only the His188Arg variant is associated with enhanced apoptosis (Gorvin et al. 2018b). It is unclear whether the partial loss-of-activity associated with the PRLR variants examined in this study would affect PRLR physiological activities, as all the variants have been identified in the heterozygous state. Previous *in vitro* studies of the Pro269Leu PRLR variant identified in a compound heterozygote individual with hyperprolactinaemia, showed that the variant impaired STAT5 phosphorylation, but had no effect on STAT5 signalling when expressed with wild-type PRLR (Kobayashi et al. 2018). Thus, it is possible that the 10 variants characterised in this study may have minimal effect on PRLR function in heterozygous individuals, unless expressed with other PRLR variants. Further investigation of the inactivating and activating PRLR variants in large well-characterised populations is required to determine their physiological consequences.

This study and our previous analysis of PRLR variants associated with prolactinoma demonstrated that both JAK-STAT and Akt signalling can be impaired by genetic variants of the receptor (Gorvin, et al. 2018a). The PRLR has also been described to activate other signalling pathways including Ras-Raf-mediated mitogen-activated protein kinase (MAPK) signalling (Bole-Feysot, et al. 1998), and can activate Src family kinases independently of JAK2 phosphorylation (Fresno Vara et al. 2000) to increase focal-adhesion kinase/MAPK and PI3K-Akt signalling, upregulate c-Myc and cyclin d1 mRNA expression, enhance cell proliferation and accelerate receptor internalisation (Acosta et al. 2003; Piazza, et al. 2009). It is possible that the PRLR variants studied in this manuscript may also affect these other signalling pathways and the observed effects on Akt could be mediated by JAK2-independent Src signalling. Src family kinases and MAPK signalling proteins are expressed in HEK293 cells (Della Rocca, et al. 1997) and future studies of PRLR variants could expand the screening pipeline to include examination of these proteins.

In summary, these studies give further insight into PRLR structure-function and highlight that rare PRLR variants are associated with alterations in receptor signalling. Future studies of rare coding variants will require a combination of molecular, *in vivo* and epidemiological approaches to appropriately classify the significance of such variants.

## References

- Acosta JJ, Munoz RM, Gonzalez L, Subtil-Rodriguez A, Dominguez-Caceres MA, Garcia-Martinez JM, Calcabrini A, Lazaro-Trueba I & Martin-Perez J 2003 Src mediates prolactin-dependent proliferation of T47D and MCF7 cells via the activation of focal adhesion kinase/Erk1/2 and phosphatidylinositol 3-kinase pathways. *Mol Endocrinol* **17** 2268-2282.
- Adzhubei I, Jordan DM & Sunyaev SR 2013 Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* **Chapter 7** Unit7 20.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS & Sunyaev SR 2010 A method and server for predicting damaging missense mutations. *Nat Methods* **7** 248-249.
- Agarwala V, Flannick J, Sunyaev S, Go TDC & Altshuler D 2013 Evaluating empirical bounds on complex disease genetic architecture. *Nat Genet* **45** 1418-1427.
- Ali S & Ali S 1998 Prolactin receptor regulates Stat5 tyrosine phosphorylation and nuclear translocation by two separate pathways. *J Biol Chem* **273** 7709-7716.
- Amaral ME, Cunha DA, Anhe GF, Ueno M, Carneiro EM, Velloso LA, Bordin S & Boschero AC 2004 Participation of prolactin receptors and phosphatidylinositol 3-kinase and MAP kinase pathways in the increase in pancreatic islet mass and sensitivity to glucose during pregnancy. *J Endocrinol* **183** 469-476.
- Banerjee RR, Cyphert HA, Walker EM, Chakravarthy H, Peiris H, Gu X, Liu Y, Conrad E, Goodrich L, Stein RW, et al. 2016 Gestational Diabetes Mellitus From Inactivation of Prolactin Receptor and MafB in Islet beta-Cells. *Diabetes* **65** 2331-2341.
- Ben-Jonathan N, LaPensee CR & LaPensee EW 2008 What can we learn from rodents about prolactin in humans? *Endocr Rev* **29** 1-41.
- Bernard V, Bouilly J, Beau I, Broutin I, Chanson P, Young J & Binart N 2016 Germline Prolactin Receptor Mutation Is Not a Major Cause of Sporadic Prolactinoma in Humans. *Neuroendocrinology* **103** 738-745.
- Binder C, Lafayette A, Archibeque I, Sun Y, Plewa C, Sinclair A & Emkey R 2008 Optimization and utilization of the SureFire phospho-STAT5 assay for a cell-based screening campaign. *Assay Drug Dev Technol* **6** 27-37.
- Bogorad RL, Courtilot C, Mestayer C, Bernichtein S, Harutyunyan L, Jomain JB, Bachelot A, Kuttann F, Kelly PA, Goffin V, et al. 2008 Identification of a gain-of-function mutation of the prolactin receptor in women with benign breast tumors. *Proc Natl Acad Sci U S A* **105** 14533-14538.
- Bole-Feysot C, Goffin V, Edery M, Binart N & Kelly PA 1998 Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocr Rev* **19** 225-268.
- Brockman JL, Schroeder MD & Schuler LA 2002 PRL activates the cyclin D1 promoter via the Jak2/Stat pathway. *Mol Endocrinol* **16** 774-784.
- Brooks AJ, Dai W, O'Mara ML, Abankwa D, Chhabra Y, Pelekanos RA, Gardon O, Tunny KA, Blucher KM, Morton CJ, et al. 2014 Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. *Science* **344** 1249783.
- Brooks CL 2012 Molecular mechanisms of prolactin and its receptor. *Endocr Rev* **33** 504-525.
- Broutin I, Jomain JB, Tallet E, van Agthoven J, Raynal B, Hoos S, Kragelund BB, Kelly PA, Ducruix A, England P, et al. 2010 Crystal structure of an affinity-matured prolactin complexed to its dimerized receptor reveals the topology of hormone binding site 2. *J Biol Chem* **285** 8422-8433.
- Brown RJ, Adams JJ, Pelekanos RA, Wan Y, McKinsty WJ, Palethorpe K, Seeber RM, Monks TA, Eidne KA, Parker MW, et al. 2005 Model for growth hormone receptor activation based on subunit rotation within a receptor dimer. *Nat Struct Mol Biol* **12** 814-821.
- Bugge K, Papaleo E, Haxholm GW, Hopper JT, Robinson CV, Olsen JG, Lindorff-Larsen K & Kragelund BB 2016 A combined computational and structural model of the full-length human prolactin receptor. *Nat Commun* **7** 11578.
- Choi Y, Sims GE, Murphy S, Miller JR & Chan AP 2012 Predicting the functional effect of amino acid substitutions and indels. *PLoS One* **7** e46688.

- Clark WT, Yu GK, Aoyagi-Scharber M & LeBowitz JH 2018 Utilizing ExAC to assess the hidden contribution of variants of unknown significance to Sanfilippo Type B incidence. *PLoS One* **13** e0200008.
- Courtillot C, Chakhtoura Z, Bogorad R, Genestie C, Bernichtein S, Badachi Y, Janaud G, Akakpo JP, Bachelot A, Kuttann F, et al. 2010 Characterization of two constitutively active prolactin receptor variants in a cohort of 95 women with multiple breast fibroadenomas. *J Clin Endocrinol Metab* **95** 271-279.
- Dagil R, Knudsen MJ, Olsen JG, O'Shea C, Franzmann M, Goffin V, Teilum K, Breinholt J & Kragelund BB 2012 The WSXWS motif in cytokine receptors is a molecular switch involved in receptor activation: insight from structures of the prolactin receptor. *Structure* **20** 270-282.
- Della Rocca GJ, van Biesen T, Daaka Y, Luttrell DK, Luttrell LM & Lefkowitz RJ 1997 Ras-dependent mitogen-activated protein kinase activation by G protein-coupled receptors. Convergence of Gi- and Gq-mediated pathways on calcium/calmodulin, Pyk2, and Src kinase. *J Biol Chem* **272** 19125-19132.
- Essagher A, Dif N, Marbehant CY, Coffey PJ & Demoulin JB 2009 The transcription of FOXO genes is stimulated by FOXO3 and repressed by growth factors. *J Biol Chem* **284** 10334-10342.
- Flanagan SE, Patch AM & Ellard S 2010 Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet Test Mol Biomarkers* **14** 533-537.
- Freemark M, Avril I, Fleenor D, Driscoll P, Petro A, Opara E, Kendall W, Oden J, Bridges S, Binart N, et al. 2002 Targeted deletion of the PRL receptor: effects on islet development, insulin production, and glucose tolerance. *Endocrinology* **143** 1378-1385.
- Fresno Vara JA, Carretero MV, Geronimo H, Ballmer-Hofer K & Martin-Perez J 2000 Stimulation of c-Src by prolactin is independent of Jak2. *Biochem J* **345 Pt 1** 17-24.
- Gadd SL & Clevenger CV 2006 Ligand-independent dimerization of the human prolactin receptor isoforms: functional implications. *Mol Endocrinol* **20** 2734-2746.
- Glasow A, Horn LC, Taymans SE, Stratakis CA, Kelly PA, Kohler U, Gillespie J, Vonderhaar BK & Bornstein SR 2001 Mutational analysis of the PRL receptor gene in human breast tumors with differential PRL receptor protein expression. *J Clin Endocrinol Metab* **86** 3826-3832.
- Gorvin CM, Metpally R, Stokes VJ, Hannan FM, Krishnamurthy SB, Overton JD, Reid JG, Breitwieser GE & Thakker RV 2018a Large-scale exome datasets reveal a new class of adaptor-related protein complex 2 sigma subunit (AP2sigma) mutations, located at the interface with the AP2 alpha subunit, that impair calcium-sensing receptor signalling. *Hum Mol Genet* **27** 901-911.
- Gorvin CM, Newey PJ, Rogers A, Stokes V, Neville MJ, Ntali G, Lees P, Morrison PJ, Singhellakis PN, Malandrinou FC, et al. 2018a Association of prolactin receptor (PRLR) variants with prolactinomas
- Gorvin CM, Newey PJ, Rogers A, Stokes V, Neville MJ, Lines KE, Ntali G, Lees P, Morrison PJ, Singhellakis PN, et al. 2018b Association of prolactin receptor (PRLR) variants with prolactinomas. *Hum Mol Genet*.
- Haxholm GW, Nikolajsen LF, Olsen JG, Fredsted J, Larsen FH, Goffin V, Pedersen SF, Brooks AJ, Waters MJ & Kragelund BB 2015 Intrinsically disordered cytoplasmic domains of two cytokine receptors mediate conserved interactions with membranes. *Biochem J* **468** 495-506.
- Huang C, Snider F & Cross JC 2009 Prolactin receptor is required for normal glucose homeostasis and modulation of beta-cell mass during pregnancy. *Endocrinology* **150** 1618-1626.
- Iavnilovitch E, Groner B & Barash I 2002 Overexpression and forced activation of stat5 in mammary gland of transgenic mice promotes cellular proliferation, enhances differentiation, and delays postlactational apoptosis. *Mol Cancer Res* **1** 32-47.
- Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, et al. 2016 REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* **99** 877-885.
- Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, et al. 2020 The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581** 434-443.
- Kobayashi T, Usui H, Tanaka H & Shozu M 2018 Variant Prolactin Receptor in Agalactia and Hyperprolactinemia. *N Engl J Med* **379** 2230-2236.
- Kulkarni MV, Tettamanzi MC, Murphy JW, Keeler C, Myszka DG, Chayen NE, Lolis EJ & Hodsdon ME 2010 Two independent histidines, one in human prolactin and one in its receptor, are critical for pH-dependent receptor recognition and activation. *J Biol Chem* **285** 38524-38533.

- Kumar P, Henikoff S & Ng PC 2009 Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **4** 1073-1081.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007 Clustal W and Clustal X version 2.0. *Bioinformatics* **23** 2947-2948.
- Lebrun JJ, Ali S, Goffin V, Ullrich A & Kelly PA 1995 A single phosphotyrosine residue of the prolactin receptor is responsible for activation of gene transcription. *Proc Natl Acad Sci U S A* **92** 4031-4035.
- Lee SA, Haiman CA, Burt NP, Pooler LC, Cheng I, Kolonel LN, Pike MC, Altshuler D, Hirschhorn JN, Henderson BE, et al. 2007 A comprehensive analysis of common genetic variation in prolactin (PRL) and PRL receptor (PRLR) genes in relation to plasma prolactin levels and breast cancer risk: the multiethnic cohort. *BMC Med Genet* **8** 72.
- Newey PJ, Gorvin CM, Cleland SJ, Willberg CB, Bridge M, Azharuddin M, Drummond RS, van der Merwe PA, Klenerman P, Bountra C, et al. 2013 Mutant prolactin receptor and familial hyperprolactinemia. *N Engl J Med* **369** 2012-2020.
- Nyante SJ, Faupel-Badger JM, Sherman ME, Pfeiffer RM, Gaudet MM, Falk RT, Andaya AA, Lissowska J, Brinton LA, Peplonska B, et al. 2011 Genetic variation in PRL and PRLR, and relationships with serum prolactin levels and breast cancer risk: results from a population-based case-control study in Poland. *Breast Cancer Res* **13** R42.
- Piazza TM, Lu JC, Carver KC & Schuler LA 2009 SRC family kinases accelerate prolactin receptor internalization, modulating trafficking and signaling in breast cancer cells. *Mol Endocrinol* **23** 202-212.
- Plotnikov A, Varghese B, Tran TH, Liu C, Rui H & Fuchs SY 2009 Impaired turnover of prolactin receptor contributes to transformation of human breast cells. *Cancer Res* **69** 3165-3172.
- Qazi AM, Tsai-Morris CH & Dufau ML 2006 Ligand-independent homo- and heterodimerization of human prolactin receptor variants: inhibitory action of the short forms by heterodimerization. *Mol Endocrinol* **20** 1912-1923.
- Rao GV & Brooks CL 2011 Functional epitopes for site 1 of human prolactin. *Biochemistry* **50** 1347-1358.
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al. 2015 Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17** 405-424.
- Schuff KG, Hentges ST, Kelly MA, Binart N, Kelly PA, Iuvone PM, Asa SL & Low MJ 2002 Lack of prolactin receptor signaling in mice results in lactotroph proliferation and prolactinomas by dopamine-dependent and -independent mechanisms. *J Clin Invest* **110** 973-981.
- Schwarz JM, Cooper DN, Schuelke M & Seelow D 2014 MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* **11** 361-362.
- Smiley KO, Brown RSE & Grattan DR 2022 Prolactin action is necessary for parental behavior in male mice. *J Neurosci*.
- Svensson LA, Bondensgaard K, Norskov-Lauritsen L, Christensen L, Becker P, Andersen MD, Maltesen MJ, Rand KD & Breinholt J 2008 Crystal structure of a prolactin receptor antagonist bound to the extracellular domain of the prolactin receptor. *J Biol Chem* **283** 19085-19094.
- Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, et al. 2012 Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337** 64-69.
- Thusberg J, Olatubosun A & Vihinen M 2011 Performance of mutation pathogenicity prediction methods on missense variants. *Hum Mutat* **32** 358-368.
- Tomic S, Chughtai N & Ali S 1999 SOCS-1, -2, -3: selective targets and functions downstream of the prolactin receptor. *Mol Cell Endocrinol* **158** 45-54.
- Vaclavicek A, Hemminki K, Bartram CR, Wagner K, Wappenschmidt B, Meindl A, Schmutzler RK, Klaes R, Untch M, Burwinkel B, et al. 2006 Association of prolactin and its receptor gene regions with familial breast cancer. *J Clin Endocrinol Metab* **91** 1513-1519.
- van Agthoven J, Zhang C, Tallet E, Raynal B, Hoos S, Baron B, England P, Goffin V & Broutin I 2010 Structural characterization of the stem-stem dimerization interface between prolactin receptor chains complexed with the natural hormone. *J Mol Biol* **404** 112-126.

Viengchareun S, Servel N, Feve B, Freemark M, Lombes M & Binart N 2008 Prolactin receptor signaling is essential for perinatal brown adipocyte function: a role for insulin-like growth factor-2. *PLoS One* **3** e1535.

Viengchareun S, Bouzinba-Segard H, Laigneau JP, Zennaro MC, Kelly PA, Bado A, Lombes M & Binart N 2004 Prolactin potentiates insulin-stimulated leptin expression and release from differentiated brown adipocytes. *J Mol Endocrinol* **33** 679-691.

Zhang C, Cherifi I, Nygaard M, Haxholm GW, Bogorad RL, Bernadet M, England P, Broutin I, Kragelund BB, Guidotti JE, et al. 2015 Residue 146 regulates prolactin receptor folding, basal activity and ligand-responsiveness: potential implications in breast tumorigenesis. *Mol Cell Endocrinol* **401** 173-188.

## **ACKNOWLEDGEMENTS**

This work was supported by the United Kingdom Medical Research Council (MRC) programme grant G1000467/2010 (RVT); Chief Scientist Office, UK Fellowship SCAF/15/01 (PJM); start-up funds from the University of Birmingham and COMPARE (CMG); a Wellcome Trust Senior Investigator Award (RVT); National Institute for Health Research (NIHR) Oxford Biomedical Research Centre Programme (RVT); and NIHR Senior Investigator Award (RVT).

## **Conflict of interest statement**

The authors declare that there is no conflict of interest.



## FIGURES AND TABLES

### Figure 1 Location of the PRLR ECD variants and structural characterisation of rare variants located in D1

(A) Crystal structure showing the two monomers (PRLR1 in pink and PRLR2 in light brown) of the rat PRLR extracellular domain (ECD) in complex with the PRL hormone (purple) (PDB 3NPZ (van Agthoven et al. 2010)). Each PRLR ECD monomer is comprised of two subdomains, designated D1 (residues 1-101) and D2 (residues 109-210), which are important in ligand binding and subsequent receptor activation (Brooks 2012; Broutin et al. 2010; Rao and Brooks 2011; Svensson et al. 2008; van Agthoven et al. 2010). Side chains of the potential deleterious variant residues from ExAC and GnomAD are shown in red. The His188 residue that is mutated in hyperprolactinaemia is shown in green. Residue 49 is labelled as Tyr in the rat structure, in humans this is His49. Residue 169 is labelled as Val in the rat structure, in human this is Ile169. Rare variants are shown in only one monomer. (B) Lys6, Pro7 and Ser88 are located in close proximity within the  $\beta 6$ - $\beta 7$ - $\beta 8$  region (Top). The variant Asn6 loses a contact with Asp91 on the  $\beta 8$  strand, Ser7 gains a contact with Val79 on the  $\beta 6$  strand and Arg88 loses a contact with Pro4 and gains a new contact with the adjacent Ser87 (all blue, bottom). (C) The Gly30 residue, located in the  $\beta 2$ - $\beta 3$  loop, forms a polar contact with the Ser2 residue (Top). When mutated to Arg30 (blue, bottom) the longer side chain is able to retain this contact and form a new hydrogen bond with Asn83 in the  $\beta 6$ - $\beta 7$  loop. (D) The Cys51 residue is located in the  $\beta 4$ - $\beta 5$  loop. Mutation in Tyr51 (blue, bottom) adds a more bulky residue within this tightly packed region, which is predicted to result in gain of two contacts with Asp53 and Phe63 of the  $\beta 4$ - $\beta 5$  loop.

### Figure 2 Structural characterisation of rare variants located in the C-terminal D1 and D2 of PRLR ECD

(A) The Tyr99 residue is located close to the PRL binding site with the PRLR. The hydroxyl group of the wild-type Tyr99 contacts Arg176 on the PRL protein (left). Mutation to His99 (blue, right) predicts retention of the contact, but at a more distal site, further away from the PRL  $\alpha$ -helix, and therefore is likely to increase the distance between the PRL and PRLR molecules which may affect binding and activation. (B) The Glu145 and Arg183 PRLR residues are located in adjacent  $\beta$ -strands and form polar

contacts with each other. Mutation to His183 (middle) and Asp145 (bottom) is predicted to disrupt two of these contacts. Additionally, Asp145 is predicted to lose a contact with the neighbouring, Ile146, a residue previously demonstrated to be important for PRLR folding and stability (Dagil et al. 2012; Zhang et al. 2015). (C) The His183 also loses a contact with Ala193, which forms part of the highly-conserved WSxWS motif (red). (D) The wild-type Glu155 forms a contact with Lys114 on the opposite PRLR protomer. The Lys155 variant loses this contact and may disrupt homodimeric structural stability. (E) The wild-type Asp187 residue lies close to the PRL binding site and forms a contact with the His188 residue of PRLR, which has a critical role in ligand binding (Kulkarni et al. 2010) (Left). Mutation to Glu187 (blue, right) leads to loss of this contact, which may affect activation of the PRLR protein.

### Figure 3 Structural characterisation of the PRLR variants located in the TMD and ICD

(A) A single  $\alpha$ -helix forms the transmembrane domain (TMD) of the PRLR. Five rare PRLR variants (Ala222Gly, Ile227Thr, Trp230Cys, Val232Ala, Val232Met) are located within the TMD, and four of these are present within the published NMR structure of the PRLR TMD (PDB:2N7I (Bugge et al. 2016)). The four WT residues (black) are located at the cytoplasmic end of the TMD and each forms backbone contacts with adjacent residues within the  $\alpha$ -helix. The mutant residues (blue) are not predicted to affect these backbone contacts. (B) Cartoon depicting the PRLR structure with the two monomers shown in brown and pink. The extracellular domain (ECD) contains two domains (D1 and D2) and is connected to the intracellular domain (ICD) via the TMD. The ICD is predicted to interact with the plasma membrane in at least three regions known as lipid interacting domains (LID1-3). The ICD also interacts with the JAK2 proteins that activate signalling downstream of the PRLR. (C) Cartoon showing the known functional domains of the PRLR ICD, with the amino acid residues (236-300, 243-251, 324-329, 350-383 and 547-598) involved shown in parentheses. The LIDs are shown in green. Two regions, Box 1 and Box 2 (brown shaded regions) are binding sites for JAK2 (Lebrun et al. 1995), and the DSGxxS region (blue shaded region) acts as a degradation motif (Plotnikov et al. 2009). The PRLR residues investigated in this study are shown in brown above the cartoon and residues investigated in previous studies are shown in red below the cartoon (Bernard et al. 2016; Gorvin et al.

2018a). The location of the Asn492Ile gain-of-function PRLR variant that is associated with prolactinoma is indicated in orange.

#### **Figure 4 Expression of the PRLR ECD and ICD variants**

Western blot analyses of HEK293 cells expressing: **(A)** PRLR ECD rare variants and **(B)** PRLR ICD rare variants. Lysates show approximately equal expression levels of PRLR in cells transfected with each rare variant and the WT PRLR. Tubulin was used as a loading control. **(C)** Confocal microscopy images of the PRLR WT and variant proteins in transfected HEK293 cells. Bar indicates 10µm.

#### **Figure 5 Functional characterisation of the JAK-STAT signalling pathway by PRLR ECD variants**

**(A)** pSTAT5 responses following prolactin (PRL) treatment in cells expressing wild-type (WT), mutant His188Arg, or ECD variants Tyr99His, Glu145Asp, Glu155Lys, Arg183His, Asp187Glu. PRL-induced pSTAT5 production was abolished in His188Arg and Asp187Glu expressing cells and significantly reduced in Tyr99His and Glu155Lys expressing cells compared to cells expressing WT. **(B)** CISH luciferase reporter activity in cells transfected with WT or the five ECD variant PRLRs. CISH reporter activity was significantly reduced in Tyr99His, Glu155Lys and Asp187Glu expressing cells compared to WT cells. Mean from 4-5 independent assays for all panels. Statistical analyses show comparisons between WT and the five ECD PRLR variants (black) and WT and the His188Arg mutant (grey) by two-way ANOVA with Dunnett's or Sidak's multiple comparisons tests. \*\*\*\*p<0.0001, \*\*\*p<0.001, \*\*p<0.01, \*p<0.05.

#### **Figure 6 Functional characterisation of the JAK2-STAT5 signalling pathway by PRLR ICD variants**

**(A)** pSTAT5 responses following prolactin (PRL) treatment in cells expressing wild-type (WT), mutant His188Arg, or ICD variants Phe255Ser, Gly263Asp, Asp320Tyr, Arg327Gln, Val535Met. PRL-

induced pSTAT5 production was abolished in His188Arg expressing cells, significantly reduced in Phe255Ser and significantly increased in Arg327Gln expressing cells when compared to WT cells. Additionally, pSTAT5 was reduced in Gly263Asp and Val535Met expressing cells when treated with high (1000 ng/mL) PRL. **(B)** CISH luciferase reporter activity in cells transfected with WT, mutant His188Arg, or the five ICD variant PRLRs. CISH reporter activity was abolished in His188Arg expressing cells, significantly reduced in Phe255Ser and significantly increased in Arg327Gln expressing cells when compared to WT cells. Mean from 4-5 independent assays for all panels. Statistical analyses show comparisons between WT and the five ICD PRLR variants (black) and WT and the His188Arg mutant (grey) by two-way ANOVA with Dunnett's or Tukey's multiple comparisons tests. \*\*\*\* $p < 0.0001$ , \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

**Figure 7 Functional characterisation of the Akt signalling pathway by PRLR rare variants**

**(A)** pAkt responses following PRL treatment in cells transfected with wild-type (WT) or the five extracellular domain (ECD) variant PRLRs. Asp187Glu had impaired pAkt responses when compared to WT expressing cells. **(B)** *FOXO1* luciferase reporter activity in cells transfected with WT, mutant His188Arg, or the five ECD variant PRLRs. Addition of 200 ng/mL PRL reduces *FOXO1* luciferase activity in WT and four variant cell-lines. However, no response was observed in His188Arg cells and Glu155Lys had reduced basal *FOXO1* activity and impaired PRL-induced responses. **(C)** pAkt responses following PRL treatment in cells transfected with WT or the five intracellular domain (ICD) variant PRLRs. Arg327Gln expressing cells had elevated basal pAkt activity. **(D)** *FOXO1* luciferase activity in cells transfected with WT or the five ICD variant PRLRs. Addition of 200 ng/mL prolactin to WT and variant PRLRs reduces *FOXO1* luciferase activity similarly in all cell-lines. Data in all panels was expressed relative to WT cells treated with 0 ng/mL PRL. Mean from 4-6 independent assays for pAkt assays. Mean (panel D) or median (panel B) from 7 independent assays for luciferase reporter assays. Statistical analyses performed by one-way ANOVA with Sidak's or Dunnett's multiple comparisons tests for panels A, C and D. Statistical analysis performed by Kruskal-Wallis test with Dunn's multiple comparisons tests for panel B. Comparisons show 0 vs. 200 nM PRL (asterisks) or between WT and variant (#). \*\*\*\* $p < 0.0001$ , \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ , NS = not significant.

**Figure 8 Effect of the PRLR rare variants on cell viability and apoptosis**

(A-B) Effect of PRL (200 ng/mL) on viability in cells expressing WT, mutant His188Arg, or (A) the extracellular domain (ECD) or (B) the intracellular domain (ICD) variant PRLRs. Cell viability was increased in cells expressing the Phe255Ser and Arg327Gln variant PRLRs at 96 hours post-treatment with PRL, when compared to WT cells. (C-D) Effect of PRL (200 ng/mL) on apoptosis in cells expressing WT, mutant His188Arg, or (C) the ECD or (D) the ICD variant PRLRs. Each point shows one biological replicate (derived from the mean of four technical replicates) performed on independent occasions. Statistical analyses performed by one-way ANOVA with Dunnett's multiple comparisons tests for panels A, C and D. Statistical analysis performed by Kruskal-Wallis test with Dunn's multiple comparisons tests for panel B. Comparisons show WT vs. variant (asterisks). \*\*\*\* $p < 0.0001$ , \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

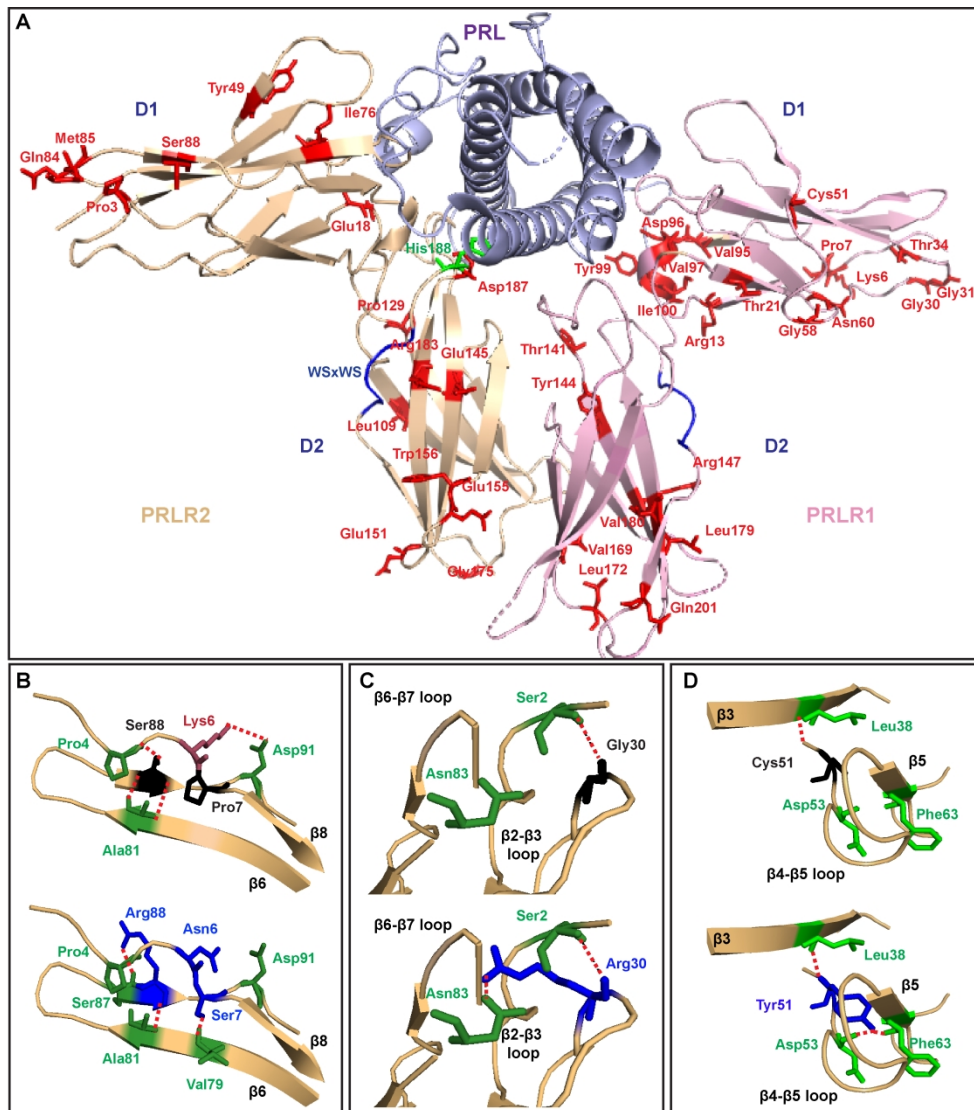


Figure 1

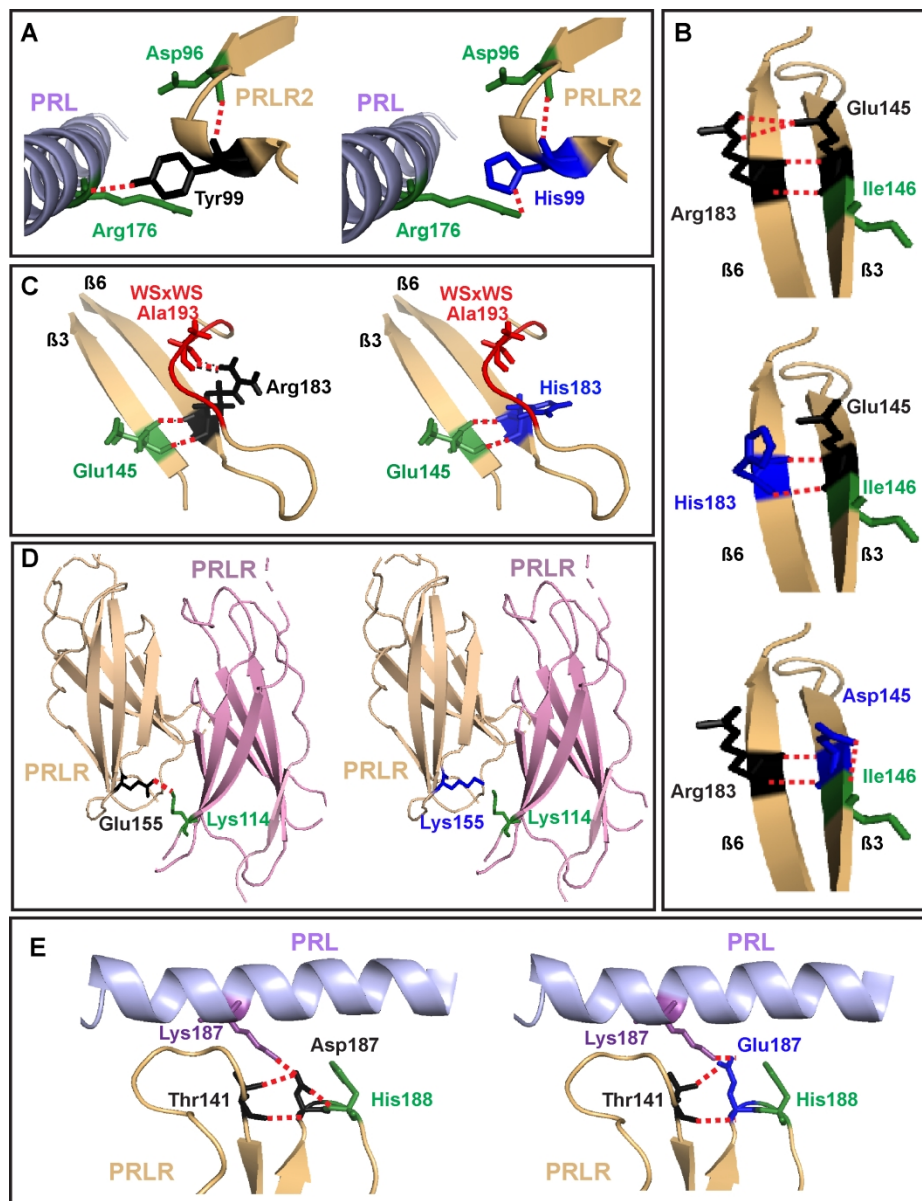


Figure 2

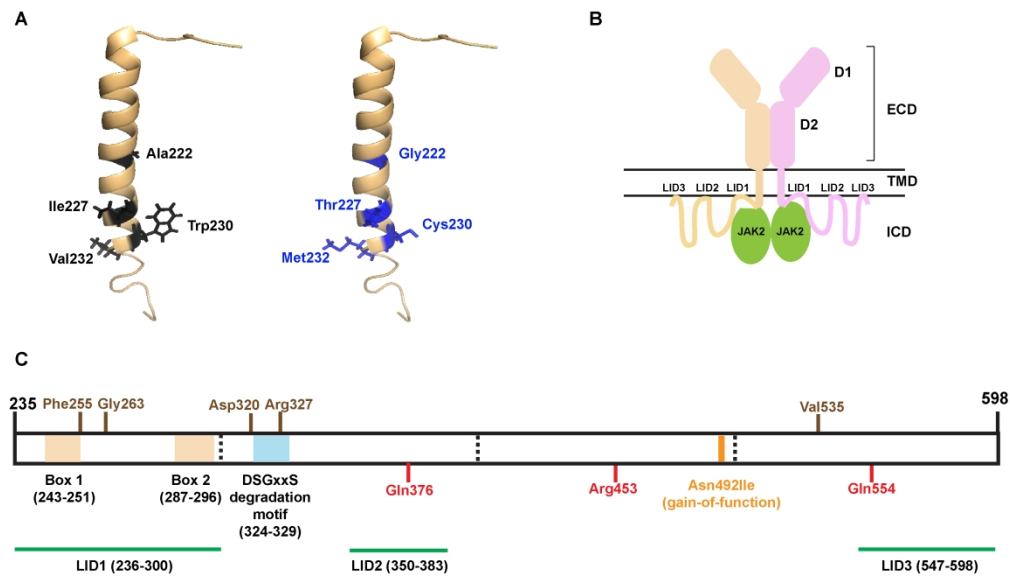


Figure 3



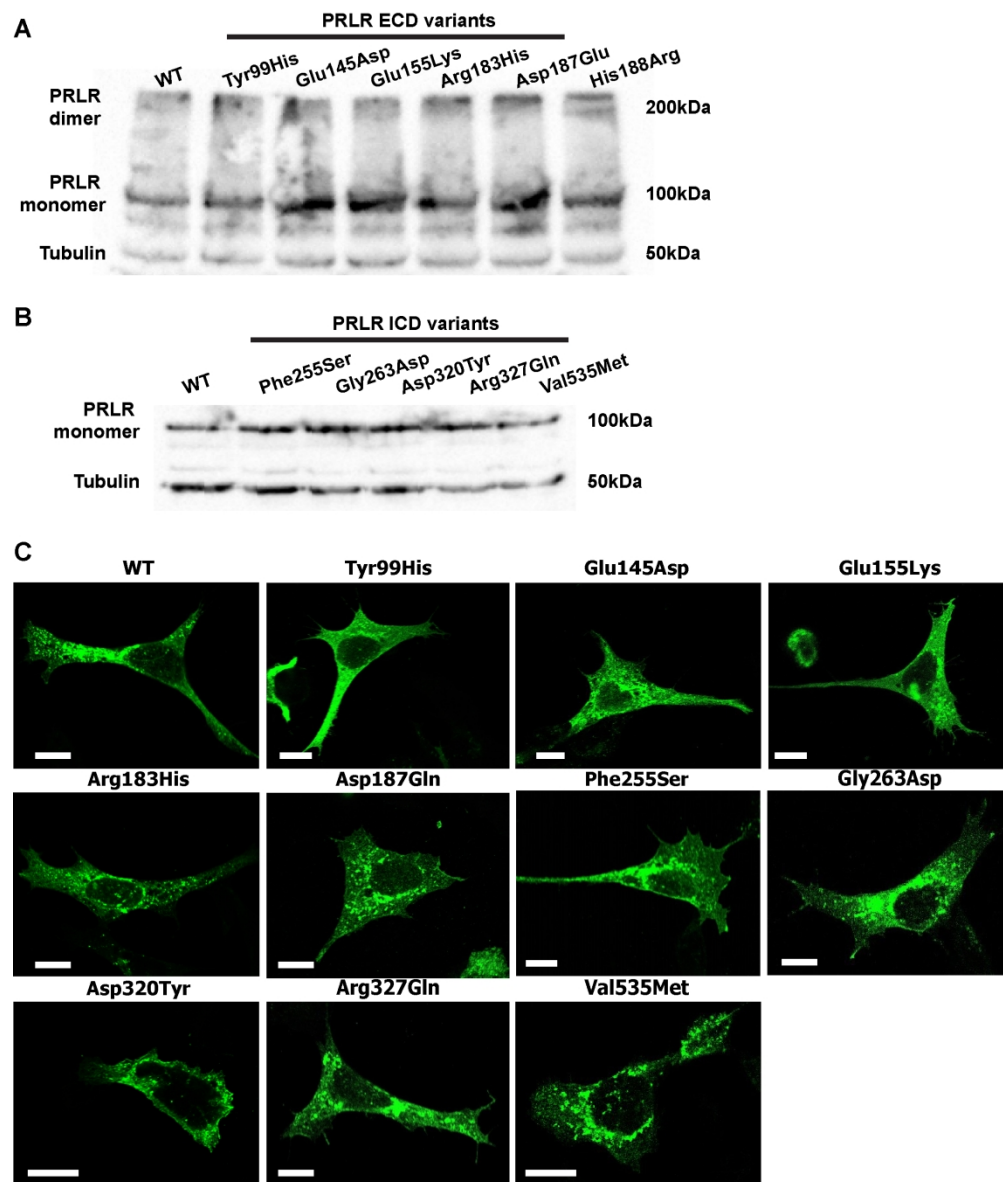


Figure 4

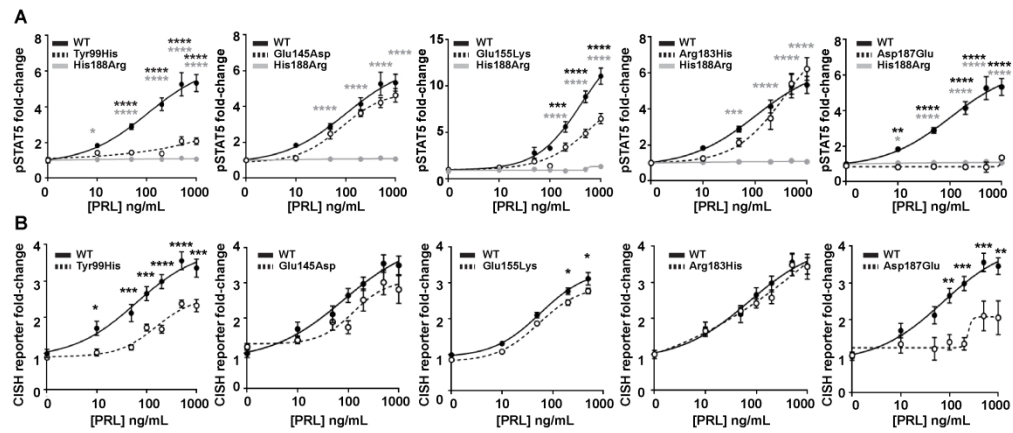


Figure 5

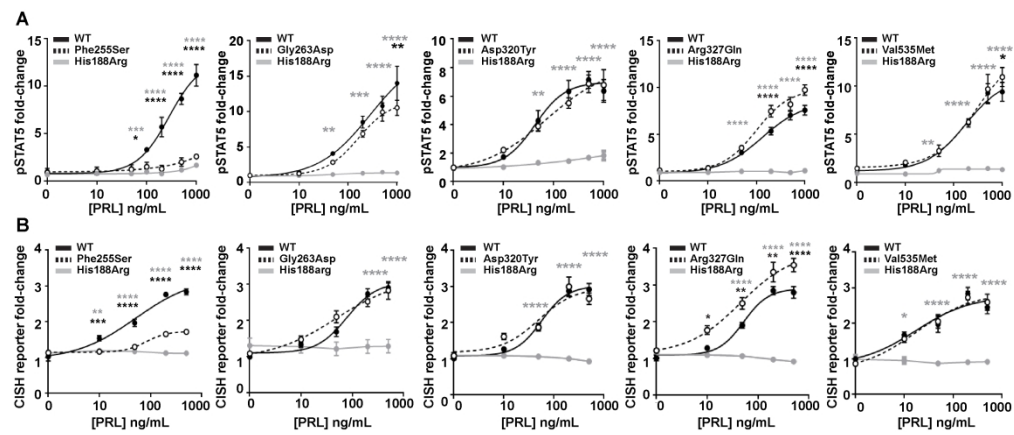


Figure 6

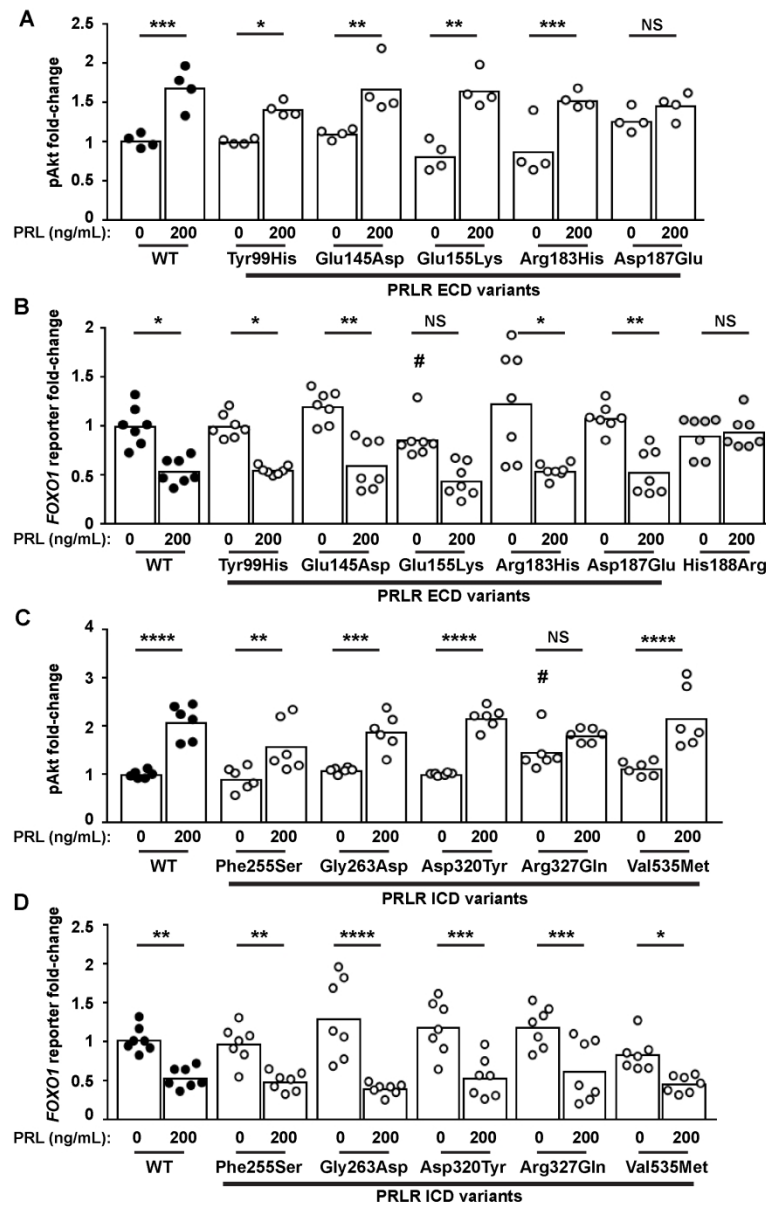


Figure 7

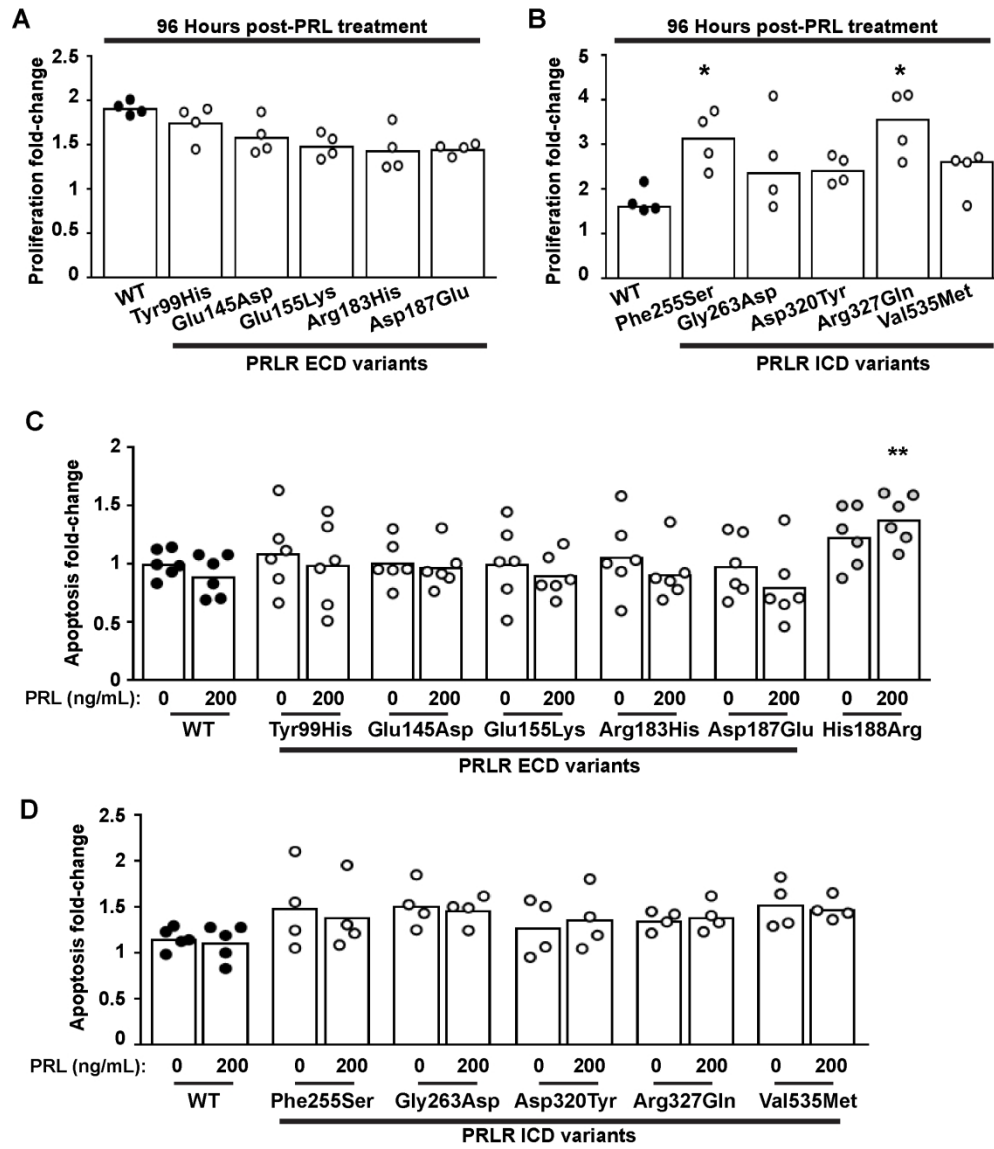


Figure 8

1 **Table 1 PRLR rare variants predicted to be potentially deleterious/pathogenic in the**  
 2 **GnomAD database**

	Amino acid change	Residue	Protein predictions <sup>a</sup>				Evolutionary Conservation <sup>b</sup>	Predicted Structural Effect
			SIFT	Polyphen	Mutation Taster	REVEL		
ECD D1	Pro>Leu	3 (27)	tolerated	probably damaging	disease causing	benign	4	No change
	Pro>His	3 (27)	deleterious	probably damaging	polymorphism	benign	4	No change
	Lys>Asn	6 (30)	deleterious	probably damaging	disease causing	benign	4	Loss of contact with Asp91
	Pro>Ser	7 (31)	deleterious	probably damaging	disease causing	disease causing	4	Gains contact with Val79
	Arg>His	13 (37)	deleterious	probably damaging	disease causing	disease causing	4	No change
	Glu>Lys	18 (42)	deleterious	probably damaging	disease causing	benign	4	No change
	Thr>Ile	21 (45)	deleterious	probably damaging	disease causing	disease causing	4	No change
	Gly>Arg	30 (54)	deleterious	probably damaging	disease causing	benign	4	Gains contact with Asn83
	Gly>Arg	31 (55)	deleterious	probably damaging	disease causing	disease causing	4	No change
	Thr>Ser	34 (58)	deleterious	benign	disease causing	benign	4	No change
	His>Arg	49 (73)	deleterious	possibly damaging	polymorphism	benign	2	No change
	Cys>Tyr	51 (75)	deleterious	probably damaging	disease causing	disease causing	4	Gains contact with Phe63
	Gly>Asp	58 (82)	deleterious	probably damaging	disease causing	benign	4	No change
	Asn>Ser	60 (84)	deleterious	possibly damaging	disease causing	benign	4	No change
	Ile>Met	76 (100)	tolerated	possibly damaging	polymorphism	benign	3	No change
	Gln>His	84 (108)	deleterious	possibly damaging	disease causing	benign	3	No change
	Met>Ile	85 (109)	tolerated	probably damaging	disease causing	benign	4	No change
	Ser>Arg	88 (112)	deleterious	possibly damaging	polymorphism	benign	4	Gains contact with Ser87 & loses contact with Pro4
	Val>Leu	95 (119)	deleterious	possibly damaging	disease causing	benign	4	No change
	Asp>Glu	96 (120)	tolerated	probably damaging	disease causing	benign	3	No change
Val>Met	97 (121)	deleterious	probably damaging	disease causing	benign	4	No change	
<b>Tyr&gt;His</b>	<b>99 (123)</b>	<b>deleterious</b>	<b>probably damaging</b>	<b>disease causing</b>	<b>benign</b>	<b>4</b>	<b>Loses contact with PRL</b>	
Ile>Val	100 (124)	deleterious	probably damaging	disease causing	benign	4	No change	
ECD D2	Leu>Met	109 (133)	deleterious	possibly damaging	polymorphism	benign	4	No change
	Pro>Ala	129 (153)	deleterious	possibly damaging	disease causing	likely benign	4	No change
	Pro>Leu	129 (153)	deleterious	possibly damaging	disease causing	disease causing	4	No change
	Thr>Met	141 (165)	deleterious	benign	disease causing	benign	3	Loses contact with Asp187
	Tyr>Asn	144 (168)	deleterious	probably damaging	disease causing	disease causing	4	No change
	<b>Glu&gt;Asp</b>	<b>145 (169)</b>	<b>deleterious</b>	<b>probably damaging</b>	<b>disease causing</b>	<b>benign</b>	<b>4</b>	<b>Loses 3 of 4 contacts with Arg183, gains contact with Ile146</b>
	Arg>Gln	147 (171)	tolerated	benign	disease-causing	benign	4	No change
	Glu>Lys	151 (175)	deleterious	benign	disease causing	benign	3	No change
	<b>Glu&gt;Lys</b>	<b>155 (179)</b>	<b>deleterious</b>	<b>possibly damaging</b>	<b>disease causing</b>	<b>benign</b>	<b>3</b>	<b>Loses contact with Lys114 on opposite PRLR protomer</b>
	Trp>Leu	156 (180)	deleterious	possibly damaging	disease causing	benign	4	No change
	Ile>Val	169 (193)	tolerated	benign	disease causing	likely benign	2	No change
	Leu>Val	172 (196)	tolerated	benign	disease-causing	benign	4	No change
	Gly>Glu	175 (199)	deleterious	probably damaging	disease causing	benign	4	No change
	Leu>Phe	179 (203)	tolerated	benign	disease-causing	benign	4	No change
Val>Ile	180 (204)	tolerated	benign	disease causing	benign	4	No change	

	<b>Arg&gt;His</b>	<b>183 (207)</b>	<b>tolerated</b>	<b>probably damaging</b>	<b>disease causing</b>	<b>benign</b>	<b>4</b>	<b>Loses 2 of 4 contacts with Glu145; loses contact with Ala193 of WSxWS motif</b>
	<b>Asp&gt;Glu</b>	<b>187 (211)</b>	<b>deleterious</b>	<b>possibly damaging</b>	<b>disease causing</b>	<b>benign</b>	<b>4</b>	<b>Loses contact with His188</b>
	Gln>Arg	201 (225)	deleterious	benign	polymorphism	benign	2	No change
	Asp>Asn	205 (229)	tolerated	benign	disease-causing	benign	4	No change
<b>TMD</b>	Ala>Gly	222 (246)	deleterious	benign	disease causing	benign	4	No change
	Ile>Thr	227 (251)	tolerated	possibly damaging	disease causing	disease causing	4	No change
	Trp>Cys	230 (254)	deleterious	possibly damaging	disease causing	disease causing	4	No change
	Val>Ala	232 (256)	deleterious	benign	disease causing	disease causing	4	No change
	Val>Met	232 (256)	tolerated	benign	possibly damaging	benign	4	No change
	Lys>Arg	235 (259)	deleterious	probably damaging	disease causing	disease causing	4	Unknown
	Gly>Ser	236 (260)	tolerated	possibly damaging	disease causing	benign	4	LID1, Box 1
	Met>Leu	239 (263)	tolerated	benign	possibly damaging	benign	4	LID1, Box 1
	Cys>Arg	242 (266)	tolerated	probably damaging	disease causing	disease causing	4	LID1, Box 1
	Ile>Val	243 (267)	tolerated	benign	disease causing	Likely benign	4	LID1, Box 1
	Pro>Leu	245 (269)	deleterious	probably damaging	disease causing	disease causing	4	LID1, Box 1
	Pro>Ser	246 (270)	deleterious	probably damaging	disease causing	disease causing	4	LID1, Box 1
	Lys>Asn	251 (275)	deleterious	probably damaging	disease causing	disease causing	4	LID1, Box 1
	Lys>Arg	253 (277)	tolerated	probably damaging	disease causing	benign	4	LID1
	Lys>Glu	253 (277)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	<b>Phe&gt;Ser</b>	<b>255 (279)</b>	<b>deleterious</b>	<b>probably damaging</b>	<b>disease causing</b>	<b>disease causing</b>	<b>4</b>	<b>LID1</b>
	Asp>Ala	256 (280)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	Leu>Trp	260 (284)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	Gly>Ser	263 (287)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	<b>Gly&gt;Asp</b>	<b>263 (287)</b>	<b>deleterious</b>	<b>probably damaging</b>	<b>disease causing</b>	<b>disease causing</b>	<b>4</b>	<b>LID1</b>
	Lys>Gln	264 (288)	deleterious	probably damaging	disease causing	benign	4	LID1
<b>ICD</b>	Ser>Cys	265 (289)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	Glu>Gln	266 (290)	deleterious	probably damaging	disease causing	benign	4	LID1
	Ser>Thr	270 (294)	tolerated	benign	disease causing	benign	3	LID1
	Ser>Arg	270 (294)	deleterious	benign	disease causing	benign	3	LID1
	Ala>Val	271 (295)	deleterious	possibly damaging	disease causing	benign	4	LID1
	Leu>Ser	272 (296)	deleterious	possibly damaging	disease causing	disease causing	4	LID1
	Gly>Val	273 (297)	tolerated	probably damaging	disease causing	benign	3	LID1
	Asp>Tyr	276 (300)	deleterious	possibly damaging	disease causing	benign	4	LID1
	Pro>His	278 (302)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	Ser>Phe	281 (305)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	Ser>Tyr	281 (305)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	Asp>Asn	285 (309)	deleterious	probably damaging	disease causing	disease causing	4	LID1
	Asp>Glu	285 (309)	tolerated	benign	disease causing	likely benign	4	LID1
	Val>Ile	293 (317)	tolerated	possibly damaging	disease causing	benign	4	LID1, Box 2
	Asp>Glu	294 (318)	tolerated	possibly damaging	polymorphism	benign	4	LID1, Box 2
Asp>Asn	298 (322)	deleterious	benign	disease causing	benign	4	LID1	
Leu>Ile	301 (325)	deleterious	possibly damaging	disease causing	benign	4		
Met>Ile	302 (326)	tolerated	benign	possibly damaging	benign	4		

Pro>Arg	316 (340)	deleterious	possibly damaging	polymorphism	benign	4	
<b>Asp&gt;Tyr</b>	<b>320 (344)</b>	<b>deleterious</b>	<b>probably damaging</b>	<b>polymorphism</b>	<b>benign</b>	<b>4</b>	
<b>Arg&gt;Gln</b>	<b>327 (351)</b>	<b>tolerated</b>	<b>benign</b>	<b>disease causing</b>	<b>benign</b>	<b>2</b>	<b>Degradation motif</b>
Arg>Pro	327 (351)	deleterious	probably damaging	disease causing	disease causing	2	Degradation motif
Arg>Trp	327 (351)	deleterious	probably damaging	disease causing	disease causing	2	Degradation motif
Asp>Asn	331 (355)	deleterious	probably damaging	disease causing	benign	4	
Pro>Arg	333 (357)	deleterious	possibly damaging	disease causing	benign	2	
Ser>Phe	334 (358)	tolerated	probably damaging	disease causing	benign	4	
Cys>Tyr	340 (364)	tolerated	benign	disease causing	benign	4	
Glu>Lys	342 (366)	deleterious	possibly damaging	disease causing	benign	4	
Pro>Thr	353 (377)	tolerated	probably damaging	polymorphism	benign	4	LID2
Glu>Gly	376 (400)	deleterious	possibly damaging	polymorphism	benign	2	LID2
Tyr>Cys	381 (405)	tolerated	benign	disease causing	benign	2	LID2
Gly>Glu	386 (410)	tolerated	possibly damaging	polymorphism	benign	2	
Gln>Glu	396 (420)	deleterious	probably damaging	polymorphism	benign	2	
Arg>Ile	403 (427)	deleterious	possibly damaging	polymorphism	benign	4	
Asp>Val	411 (435)	deleterious	possibly damaging	disease causing	disease causing	4	
Asp>Tyr	411 (435)	deleterious	possibly damaging	polymorphism	disease causing	4	
Cys>Arg	413 (437)	tolerated	possibly damaging	polymorphism	benign	4	
Cys>Tyr	413 (437)	tolerated	possibly damaging	polymorphism	benign	4	
Thr>Pro	425 (449)	tolerated	possibly damaging	polymorphism	likely benign	2	
Asp>Asn	463 (487)	tolerated	benign	disease-causing	benign	2	
Gln>His	472 (496)	tolerated	possibly damaging	polymorphism	benign	4	
Lys>Glu	481 (505)	tolerated	possibly damaging	polymorphism	benign	3	
Asp>Asn	484 (508)	deleterious	possibly damaging	disease causing	disease causing	3	
Val>Met	486 (510)	deleterious	probably damaging	disease causing	disease causing	4	
Glu>Ala	487 (511)	deleterious	probably damaging	disease causing	disease causing	4	
Glu>Lys	487 (511)	deleterious	probably damaging	disease causing	disease causing	4	
Ile>Phe	488 (512)	deleterious	probably damaging	disease causing	benign	4	
Ile>Thr	488 (512)	deleterious	probably damaging	disease causing	disease causing	4	
His>Tyr	489 (513)	deleterious	probably damaging	disease causing	disease causing	4	
Gly>Val	495 (519)	deleterious	probably damaging	disease causing	benign	4	
Leu>Phe	497 (521)	deleterious	probably damaging	polymorphism	benign	4	
Leu>Arg	500 (524)	deleterious	possibly damaging	polymorphism	disease causing	2	
Pro>Gln	501 (525)	deleterious	probably damaging	disease causing	benign	4	
Pro>Thr	501 (525)	tolerated	possibly damaging	disease causing	benign	4	
Pro>Leu	501 (525)	tolerated	benign	disease causing	benign	4	
Arg>Thr	504 (528)	deleterious	benign	polymorphism	benign	2	
Lys>Asn	520 (544)	deleterious	probably damaging	disease causing	benign	4	
Glu>Lys	521 (545)	deleterious	probably damaging	disease causing	disease causing	4	
Lys>Glu	524 (548)	deleterious	possibly damaging	disease causing	disease causing	4	
Asp>His	530 (554)	deleterious	probably damaging	disease causing	benign	4	
Asp>Gly	530 (554)	tolerated	possibly damaging	disease causing	benign	4	
Asn>Lys	531 (555)	deleterious	benign	polymorphism	benign	3	



Asn>Ser	532 (556)	tolerated	benign	disease causing	benign	4	
Leu>Val	534 (558)	deleterious	probably damaging	disease causing	disease causing	4	
<b>Val&gt;Met</b>	<b>535 (559)</b>	<b>deleterious</b>	<b>probably damaging</b>	<b>disease causing</b>	<b>disease causing</b>	<b>4</b>	
Asp>His	539 (563)	deleterious	probably damaging	polymorphism	benign	3	
Ala>Asp	542 (566)	tolerated	possibly damaging	polymorphism	benign	4	
Glu>Lys	549 (573)	deleterious	possibly damaging	polymorphism	benign	4	LID3
Glu>Gly	549 (573)	deleterious	possibly damaging	polymorphism	benign	4	LID3
Ala>Asp	552 (576)	deleterious	benign	polymorphism	disease causing	4	LID3
Lys>Arg	553 (577)	tolerated	possibly damaging	polymorphism	benign	4	LID3
Asn>Lys	562 (586)	tolerated	possibly damaging	polymorphism	benign	2	LID3
Asn>Tyr	562 (586)	deleterious	probably damaging	polymorphism	benign	2	LID3
Leu>Pro	580 (604)	tolerated	probably damaging	disease causing	disease causing	3	LID3
Asp>Tyr	586 (610)	deleterious	probably damaging	disease causing	disease causing	3	LID3
Asp>Gly	589 (613)	deleterious	probably damaging	disease causing	disease causing	3	LID3
Pro>Arg	590 (614)	deleterious	probably damaging	disease causing	disease causing	3	LID3
Cys>Arg	592 (616)	tolerated	possibly damaging	polymorphism	disease causing	3	LID3

3

4 <sup>a</sup>Protein predictions were made using SIFT, Polyphen-2, Mutation Taster and REVEL. <sup>b</sup>Evolutionary

5 conservation was based on the PRLR ortholog sequence in four species (cow, dog, mouse and rat) in

6 comparison to the human protein. The score shows the number of species (out of 4) in which the residue

7 is conserved. <sup>c</sup>Structural analyses were performed using the reported crystal structure of the ligand-

8 bound homodimeric ECD and the NMR structure of the D2 domain (Dagil, et al. 2012; van Agthoven,

9 et al. 2010) (Figures 1-3). The ten PRLR variants shown in bold are those that were functionally

10 characterised.

1 **Table 2: Summary of effects of rare variants on PRLR function.**

PRLR Variant	Predicted pathogenicity			STAT5 signalling		PI3K/ Akt signalling		Cell viability	Apoptosis
	Prediction programs <sup>a</sup>	Evolutionary conservation <sup>b</sup>	Structural predictions	pSTAT5	CISH	pAkt	FOXO1		
WT	0	N/A	N/A	++	++	++	++	++	++
His188Arg	4	4	Affects PRL binding	-	-	-	-	++	+++
ECD	Tyr99His	3	Affects PRL binding	+	+	++	++	++	++
	Glu145Asp	3	Loss and gain of contacts	++	++	++	++	++	++
	Glu155Lys	3	Disrupts homodimer region	+	+	++	+	++	++
	Arg183His	2	Loses contact with WSxWS motif	++	++	++	++	++	++
	Asp187Glu	3	Loses contact with His188	-	+	+	++	++	++
	Phe255Ser	4	4	LID1, close to JAK2 binding site	-	+	++	++	+++
ICD	Gly263Asp	4	4	LID1	++	++	++	++	++
	Asp320Tyr	2	4	Unknown	++	++	++	++	++
	Arg327Gln	1	2	Degradation motif	+++	+++	+	++	+++
	Val535Met	4	4	Unknown	++	++	++	++	++

2

3 <sup>a</sup>A score out of four based on protein predictions using SIFT, Polyphen-2, Mutation Taster and REVEL  
4 is given. If the variant was predicted to be probably damaging or damaging/ pathogenic it was classified  
5 as affected. <sup>b</sup>Evolutionary conservation was based on the PRLR ortholog sequence in four species (cow,  
6 dog, mouse and rat) in comparison to the human protein. The score shows the number of species in  
7 which the residue is conserved out of 4. +++, gain-of-function; ++, normal; +, impaired; -, abolished.  
8 \*Arg327Gln has significantly increased basal Akt activity, which likely accounts for the impaired PRL-  
9 induced activity. N/A, not applicable.

10

11