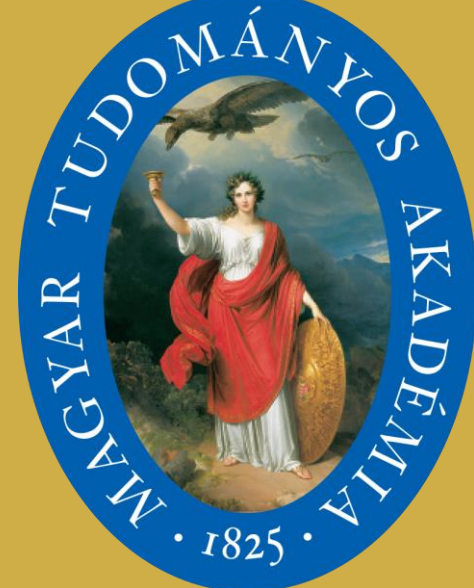




Interallelic interactions in autosomal recessive disorders: a population-genetic and a molecular approach

Ágnes Mikó^{1,2}, Ambrus Kaposi¹, Gusztáv Schay³, Eszter Balogh^{1,2}, Dániel Seidl^{1,2}, Karolina Schnabel^{1,2}, Dóra K. Menyhárd⁴, Kálmán Tory^{1,2}

¹MTA-SE Lendület Nephrogenetic Laboratory, Hungarian Academy of Sciences, Budapest, Hungary; ²First Department of Pediatrics, Semmelweis University, Budapest, Hungary;

³Semmelweis University, Department of Biophysics and Radiation Biology, Budapest, Hungary; ⁴MTA-ELTE Protein Modeling Research Group, Budapest, Hungary

INTRODUCTION

- Autosomal recessive (AR) disorders are typically completely penetrant.
- We identified by population-genetic approach the first variant (*NPHS2*, p.R229Q), that is pathogenic only when trans-associated to specific mutations [1].

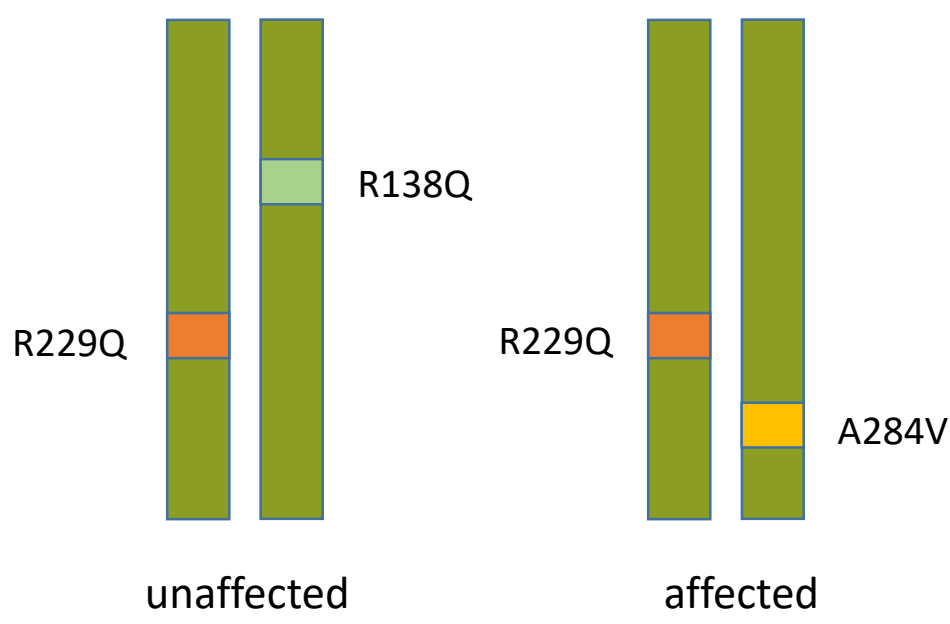


Figure 1. Incomplete penetrance secondary to interallelic interactions. The *NPHS2* R229Q variant is pathogenic only when trans-associated with specific 3' mutations, as A284V, but not with R138Q. Both R138Q and A284V are null mutations.

AIM

To identify other incompletely penetrant variants in AR disorders.

METHODS

I. **Genes of interest** were selected on the basis of four frequency criteria:

- the cumulative allele frequency (AF) of the LOF mutations in the European population is >0.02% in gnomAD
- >100 different mutations are reported in HGMD
- the prevalence of the associated disorder is reported to be >1:200.000
- >100 patients have been published

Figure 2. The 17 selected frequent AR disorders

	Genes	Disease	9	GALT	Galactosemia
1	ASL	Argininosuccinic aciduria	10	IDUA	Hurler- Scheie syndrome (Mucopolysaccharidosis type I)
2	ATP7B	Wilson disease	11	MUT	Methylmalonic acidemia
3	CAPN3	Juvenile limb-girdle muscular dystrophy	12	NPHS1	Congenital nephrotic syndrome
4	CFTR	Cystic fibrosis	13	NPHS2	Nephrotic syndrome
5	CTNS	Cistinosis	14	PAH	Phenylketonuria
6	DHCR7	Smith-Lemli-Opitz syndrome	15	PKHD1	Autosomal recessive polycystic kidney disease
7	GAA	Pompe disease (Glycogen storage disease type II)	16	PMM2	Congenital disorder of glycosylation
8	GALNS	Morquio syndrome (Mucopolysaccharidosis type IV)	17	SLC26A4	Pendred syndrome

II. Data collection

Clinical and genotype data of **12 056 patients** with biallelic HGMD mutations in one of the 17 selected genes have been collected from the medical literature.

The majority of the patients were of European origin (Fig. 3). Therefore, we subsequently analyzed the penetrance of the variants found in the European patient population.

III. Analysis of the described 1989 variants

1. Exclusion of the non-pathogenic variants

HGMD variants with a lower AF in the patient population as compared to the general population (like *CFTR*, R75Q, Fig. 4) were considered non-pathogenic and excluded from further analyses. Thus, 28 of 1989 variants were excluded.

2. Assessment of the penetrance

The penetrance of the pathogenic variants was assessed based on their relative allele frequency in the Caucasian patient population (RAF_{pt}) and in the general European population of gnomAD (RAF_{co}) as compared to that of the LOF mutations (Fig. 4, 5). All calculations were performed by a computerized algorithm written in R language.

Figure 4. Proportional representation of a completely penetrant variant (*CFTR*, F508del), LOF variants, a non-pathogenic (R75Q), and an incompletely penetrant variant (R117H) in the patient- and the general population.

While the completely penetrant F508del and the incompletely penetrant R117H variants are both enriched in the patient population, the R75Q variant is not, indicating its non-pathogenic nature.

3. Detection of interallelic interactions

To determine whether the incomplete penetrance results from interallelic interactions, we investigated whether an incompletely penetrant variant is associated to specific mutation(s) in the patient population, i.e. specific associations appear more frequently than expected in the patient population.

4. Characterization of the binding capacity of R229Q and wt podocin

HA-tagged podocin variants (n=11) were expressed in HEK293 cells, immunoprecipitated by anti HA-antibody and eluted by competitive peptide elution with HA peptides. Podocin variants were stained either by Alexa Fluor 488 C5 Maleimide (donor dye) or Alexa Fluor 555 C2 Maleimide (acceptor dye) molecules (Invitrogen, Thermo Fisher Scientific) and were incubated overnight at 4°C. Förster type Resonance Energy Transfer (FRET) was measured between two differently stained podocin variants in a final volume of 100 µl, containing each podocin with a concentration of 4 µM. The fluorescence lifetime of the donor dye (Alexa 488 C5) was measured with a Chronos BH Time Correlated Single Photon Counting lifetime spectrometer (ISS Inc., Champaign, IL) using an excitation wavelength of 488 nm with a 3 nm band pass and an emission wavelength of 524 nm with a 5 nm band pass.

RESULTS

Identification of incompletely penetrant variants

We found 29 variants (Fig. 5, 6) to have a significantly ($p < 0.0004$) reduced penetrance (<30%) as compared to the LOF mutations. The incompletely penetrant variants were found hypomorphic on the basis of the associated clinical phenotype (Fig. 7).

	General pop.	Patient pop.	penetrance
R117H	311	88	19% $p = 8 \times 10^{-37}$
LOF	352	518	

Figure 5. Penetrance of the *CFTR*, R117H variant

The R117H variant in the patient population was significantly underrepresented ($p = 8 \times 10^{-37}$) as compared to the LOF alleles.

Figure 6. Incompletely penetrant variants

	Gene	Nucleotid	Aminoacid	Penetrance	p-value
1	ASL	c.35G>A	p.R12Q	1.8%	9.78×10^{-33}
2		c.392C>T	T131M	1.6%	1.55×10^{-19}
3		c.2393C>A	p.A798E	10.1%	8.2×10^{-98}
4	CAPN3	c.1355-6G>A		5.20%	0.00022
5		c.1435A>G	p.S479G	12%	0.00024
6		c.2991G>C	p.L997F	5.7%	5.77×10^{-67}
7		c.350G>A	p.R117H	19.3%	3.5×10^{-36}
8	CFTR	c.349C>T	p.R117C	11.4%	1.53×10^{-11}
9		c.617T>G	p.L206W	26.4%	0.00011
10		c.2657+2_2657+3insA		13%	0.0002
11	GAA	c.-32-13T>G		18%	4.38×10^{-26}
12	NPHS2	c.686G>A	p.R229Q	1.4%	5×10^{-103}
13		c.734T>C	p.V245A	4.6%	1.20×10^{-45}
14		c.1139C>T	p.T380M	6.5%	2.24×10^{-19}
15		c.1241A>G	p.Y414C	26.5%	1.12×10^{-12}
16		c.194T>C	p.I65T	24.1%	4.67×10^{-12}
17		c.688G>A	p.V230I	9.3%	5.67×10^{-12}
18	PAH	c.1208C>T	p.A403V	29.9%	2.36×10^{-11}
19		c.261C>A	p.S87R	5.9%	2.32×10^{-99}
20		c.805A>C	p.I269L	5%	2.72×10^{-99}
21		c.896T>G	p.F299C	13.6%	2.93×10^{-98}
22		c.533A>G	p.E178G	14.7%	4.32×10^{-97}
23		c.117C>G	p.F39L	21.9%	6.30×10^{-96}
24		c.745C>T	p.L249F	8.55%	0.00033
25		c.7264T>G	p.C2422G	3.4%	2.61×10^{-19}
26		c.10036T>C	p.C3346R	4.6%	2.65×10^{-10}
27	PKHD1	c.5498C>T	p.S1833L	4.8%	9.73×10^{-10}
28		c.6992T>A	p.I2331K	15.2%	3.93×10^{-98}
29		c.4870C>T	p.R1624W	15.9%	1.97×10^{-95}

Figure 7. Incompletely penetrant variants are typically hypomorphic

variants	hypomorphic	null	no data
incompletely penetrant	22	3	4
completely penetrant (>70%)	179	737	673

While the majority of the incompletely penetrant variants are hypomorphic, the variants with a penetrance >70% are typically null ($p = 8.58 \times 10^{-13}$), supporting the specificity of the population-genetic approach in the assessment of incomplete penetrance.

Incomplete penetrance secondary to interallelic interactions

Among the incompletely penetrant variants, only the **R229Q** was associated significantly more frequently to specific mutations in the patient population.

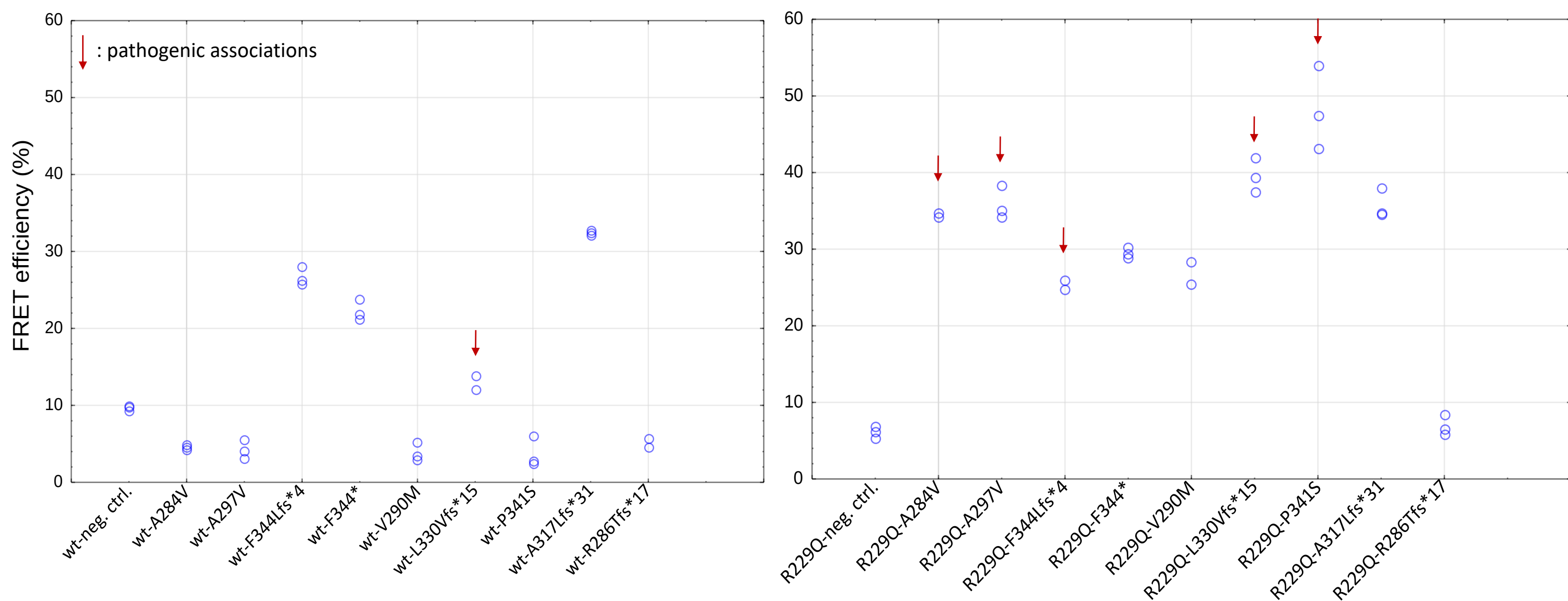
These were all 3' mutations with a known or suggested dominant negative effect (Fig. 8) [2].

Figure 8. Mutations that are significantly enriched (i.e. pathogenic) in patients with R229Q as compared to patients without R229Q.

	Nucleotid	Aminoacid	p-value
1	c.851C>T	A284V	3.55×10^{-11}
2	c.862G>A	A288T	3.09×10^{-05}
3	c.890C>T	A297V	2.51×10^{-04}
4	c.928G>A	E310L	4.54×10^{-07}
5	c.929A>T	E310V	2.51×10^{-04}
6	c.983A>G	Q328R	2.02×10^{-03}
7	c.1032delT	F344Lfs*4	3.77×10^{-06}

Figure 9. Increased binding capacity of R229Q podocin explains its pathogenicity in specific associations

We found a significantly higher FRET efficiency in the R229Q (n=7) (34% [6-32.8%], median [quartiles]) than in the wt heterooligomers (n=7) (5.8% [2.4%-32.8%], $p = 1 \times 10^{-06}$), indicating a higher binding capacity of R229Q podocin. We also found a significantly higher FRET efficiency in the pathogenic R229Q (n=5) (35% [24-54%], median [quartiles]) than in the non-pathogenic R229Q associations (n=4) (28.6% [5.9%-37.9%], $p = 0.01$), indicating a higher binding capacity of the pathogenic R229Q associations.



CONCLUSIONS

- We found 29 frequent variants to be incompletely penetrant in 17 AR disorders: incomplete penetrance is more common than generally expected in AR disorders.
- We found only the *NPHS2* R229Q variant to be incompletely penetrant secondary to interallelic interactions.
- The increased binding capacity of R229Q podocin is expected to contribute to the formation of pathogenic heterooligomers.

REFERENCES

- Tory K, Menyhárd DK, Woerner S, Nevo F, Gribouval O, Kerti A, Straner P, Arrondel C, Cong EH, Tulassay T, Mollet G, Perczel A, Antignac C Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome. Nat Genet 2014;46:299-304.
- Stráner P, Balogh E, Schay G, Arrondel C, Mikó Á, L'Auné G, Benmerah A, Perczel A, K. Menyhárd D, Antignac C, Mollet G, Tory K: C-terminal oligomerization of podocin mediates interallelic interactions. Biochim Biophys Acta Mol Basis Dis. 2018;1864:2448-2457
- Mikó Á, K. Menyhárd D, Kaposi A, Antignac C, Tory K: The mutation-dependent pathogenicity of *NPHS2* R229Q: a guide for clinical assessment, Hum Mutat, 2018