

# The Kidney-Genetics Documentation

Bernt Popp, Nina Rank, Constantin Wolff, Jan Halbritter

2023-05-22

## Contents

<b>Preface</b>	<b>1</b>
Objective . . . . .	2
Methods . . . . .	2
0.1 Results . . . . .	2
0.2 Conclusion . . . . .	2
0.3 Outlook . . . . .	2
<b>1 Analyses result tables</b>	<b>2</b>
1.1 Main table: Merged analyses sources . . . . .	3
1.2 Result table: PanelApp . . . . .	3
1.3 Result table: Literature . . . . .	3
1.4 Result table: Diagnostic panels . . . . .	4
1.5 Result table: HPO in rare disease databases . . . . .	4
1.6 Result table: PubTator . . . . .	5
<b>2 Analyses plots</b>	<b>5</b>
2.1 UpSet plot of merged analyses sources . . . . .	6
2.2 Bar plot of PanelApp results . . . . .	6
2.3 Bar plot of Literature results . . . . .	7
2.4 Bar plot of Diagnostic panels results . . . . .	8
2.5 Bar plot of HPO in rare disease databases results . . . . .	9
2.6 Bar plot of PubTator results . . . . .	10

## Preface

---

This documentation is intended to describe the Kidney-Genetics<sup>1</sup> project.

<sup>1</sup><https://github.com/halbritter-lab/kidney-genetics>

## Objective

How can we address the lack of a unified and standardized database of kidney disease-associated genes, which hampers diagnosis, treatment, and research comparability in the field of kidney diseases?

## Methods

To create a comprehensive and standardized database of kidney-related genes, we employed the following methods:

1. Utilized data from Genomics England and Australia PanelApp.
2. Conducted a comprehensive literature review of published gene lists.
3. Collected information from clinical diagnostic panels for kidney disease.
4. Performed a Human Phenotype Ontology (HPO)-based search in rare disease databases (OMIM, Orphanet).
5. Employed a PubTator API-based automated literature extraction from PubMed.

We also developed an evidence-scoring system to differentiate highly confirmed disease genes from candidate genes.

### 0.1 Results

The “Kidney-Genetics” database currently includes detailed information on 2,906 kidney-associated genes. Notably, 439 genes (15.1%) are present in three or more of the analyzed information sources, indicating high confidence and their potential for diagnostic use.

To ensure currency, Kidney-Genetics will be regularly and automatically updated. We will also provide phenotypic and functional clustering results to facilitate gene grouping.

### 0.2 Conclusion

Kidney-Genetics is a comprehensive and freely accessible database that researchers can use to analyze genomic data related to kidney diseases. The database is regularly updated through a standardized pipeline and an automated system, ensuring it remains up-to-date with the latest advancements in kidney research and diagnostics.

By utilizing Kidney-Genetics, clinicians and researchers can enhance their understanding of the genetic aspects of kidney disorders.

### 0.3 Outlook

Future goals include manual curation and the assignment of diagnostic genes to specific nephrological disease groups, such as syndromic vs. isolated, adult- vs. pediatric-onset, and cystic vs. nephrotic, among others.

## 1 Analyses result tables

---

## 1.1 Main table: Merged analyses sources

This table shows the merged results of all analyses files as a wide table with summarized information.

approved_symbol	hgnc_id	evidence_count	list_count	01_PanelApp	02_Literature	03_DiagnosticPanels	04_HPO	05_PubTator
All	All	All	All	All	All	All	All	All

## 1.2 Result table: PanelApp

This table shows results of the first analysis searching kidney disease associated genes from the PanelApp project in the UK and Australia.

approved_symbol	hgnc_id	gene_name_reported	source	source_count	source_evidence
All	All	All	All	All	All

## 1.3 Result table: Literature

This table shows results of the second analysis searching kidney disease associated genes from various publications.

approved_symbol	hgnc_id	gene_name_reported	source	source_count	source_evidence
All	All	All	All	All	All

## 1.4 Result table: Diagnostic panels

This table shows results of the third analysis searching kidney disease associated genes from clinical diagnostic panels for kidney disease.

approved_symbol	hgnc_id	gene_name_reported	source	source_count	source_evidence
All	All	All	All	All	All

## 1.5 Result table: HPO in rare disease databases

This table shows results of the fourth analysis searching kidney disease associated genes from a Human Phenotype Ontology (HPO)-based search in rare disease databases (OMIM, Orphanet).

approved_symbol	hgnc_id	gene_name_reported	source	source_count	source_evidence
All	All	All	All	All	All

## 1.6 Result table: PubTator

This table shows results of the fifth analysis searching kidney disease associated genes from a PubTator API-based automated literature extraction from PubMed.

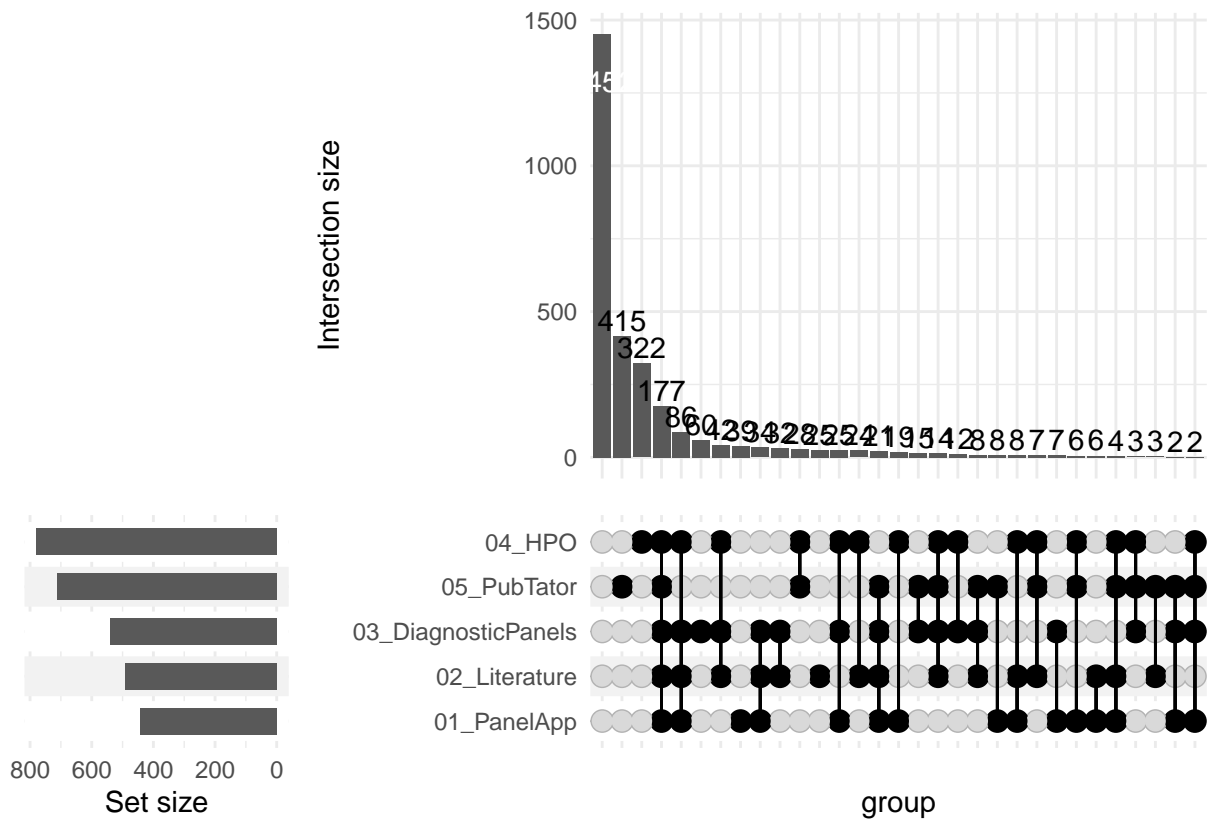
approved_symbol	hgnc_id	gene_name_reported	source	source_count	source_evidence
All	All	All	All	All	All

## 2 Analyses plots

---

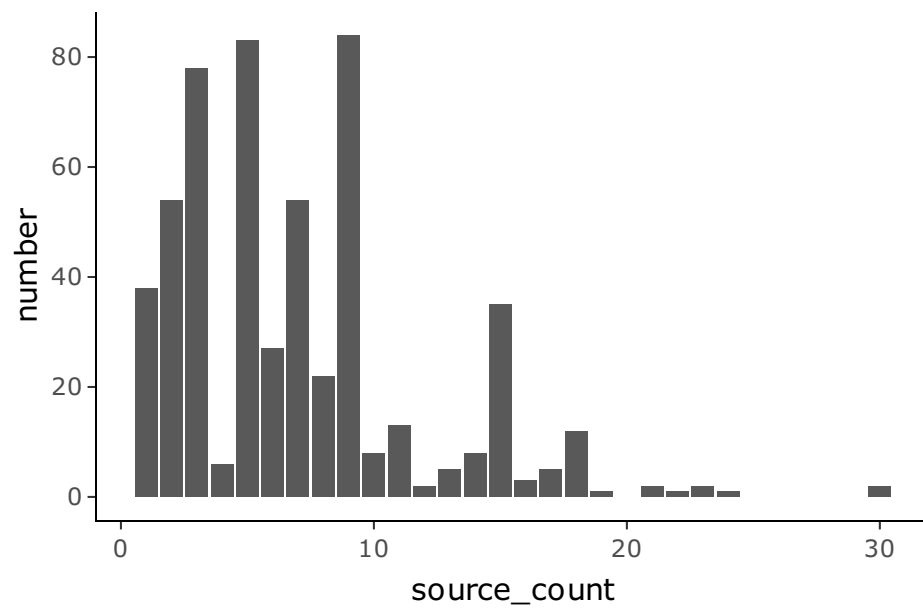
2.1 UpSet plot of merged analyses sources

UpSet plot of the merged analyses.



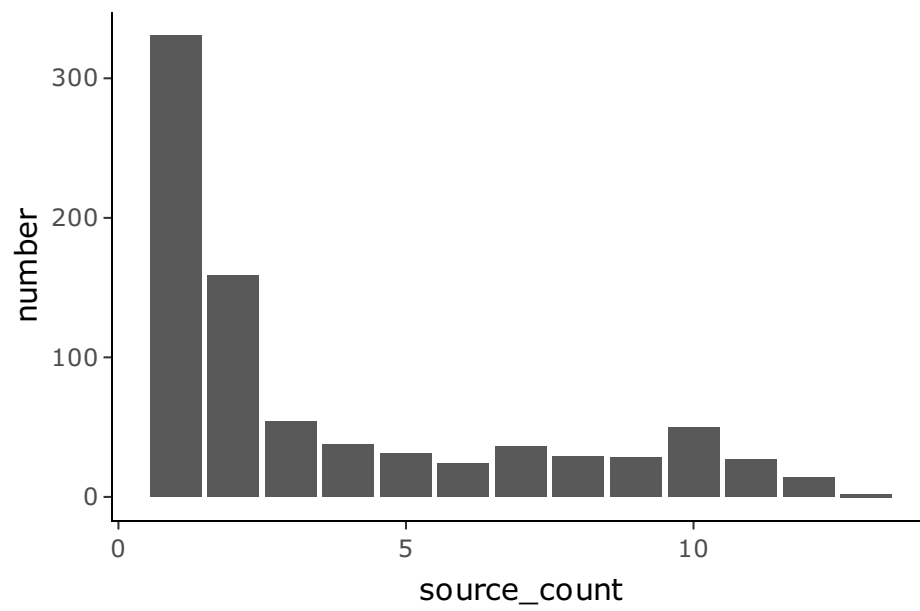
2.2 Bar plot of PanelApp results

Bar plot of the PanelApp analysis.



### 2.3 Bar plot of Literature results

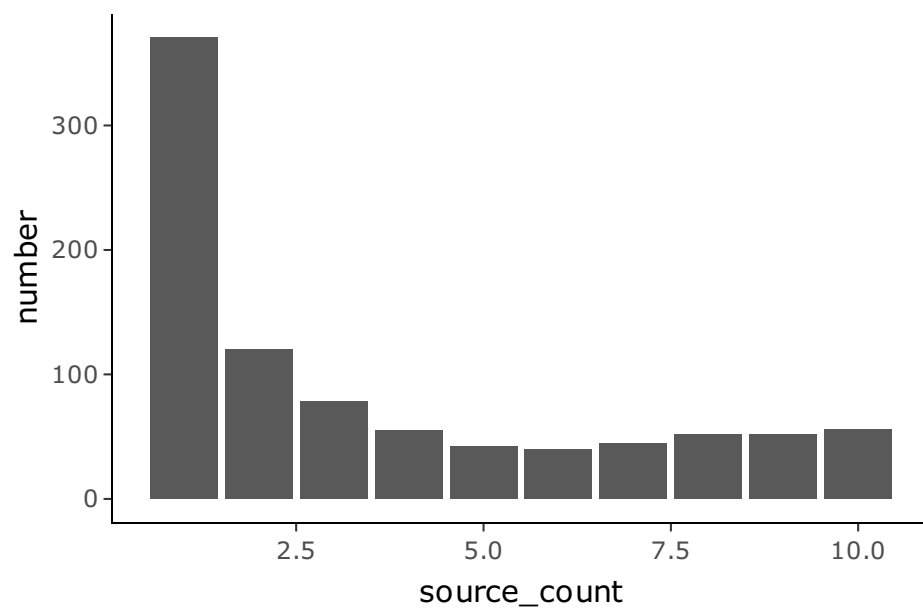
Bar plot of the Literature analysis.



## 2.4 Bar plot of Diagnostic panels results

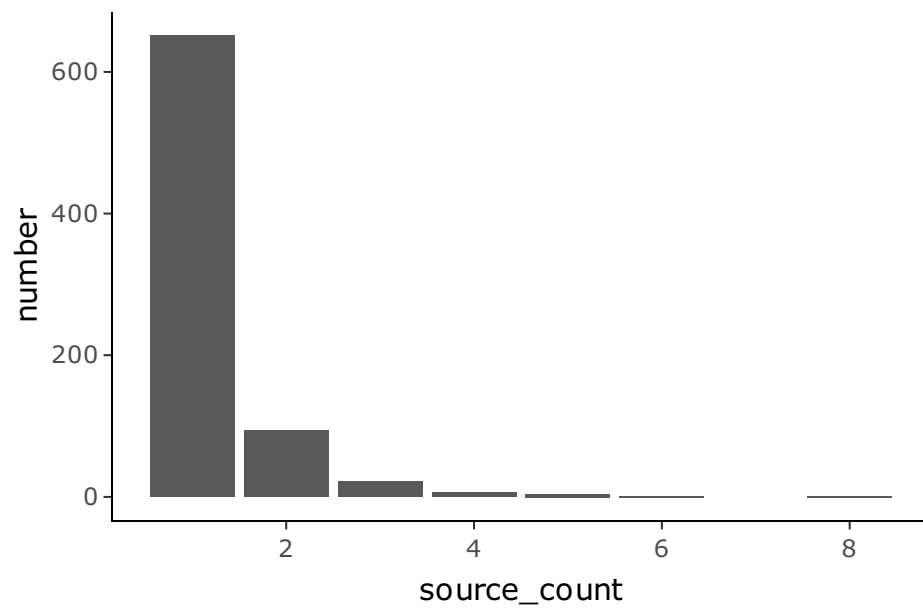
Bar plot of the Diagnostic panels analysis.





## 2.5 Bar plot of HPO in rare disease databases results

Bar plot of the HPO in rare disease databases analysis.



## 2.6 Bar plot of PubTator results

Bar plot of thePubTator analysis.

