nature portfolio

https://doi.org/10.1038/d41573-022-00120-3

Supplementary information

Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs

In the format provided by the authors

Data and analysis

To obtain the gene–indication–drug triples, the list of 50 approved drugs provided by Mullard[1] was expanded, mapping the on-target gene products to their corresponding Ensembl genes using the ChEMBL database. After manual curation, 8 additional mechanisms of action (MoAs) not originally reported by ChEMBL were included in the set. The drug indications for the 50 approved drugs were also manually mapped to the Experimental Factor Ontology.

In order to extract the genetic support for the 50 gene–indication pairs, we leveraged 15 different genetic resources as integrated by the Open Targets Platform. The 15 resources cover different aspects of rare, common or somatic variation. In Supplementary Figure 1, evidence supporting the same gene–indication pair was aggregated using the Platform association score using the ontological expansion of evidence https://platform-docs.opentargets.org/associations. This information is represented in Supplementary Figure 1 with a blue-coloured background. Direct associations with ontological expansion account for 19 out of the 50 genetically supported approvals.

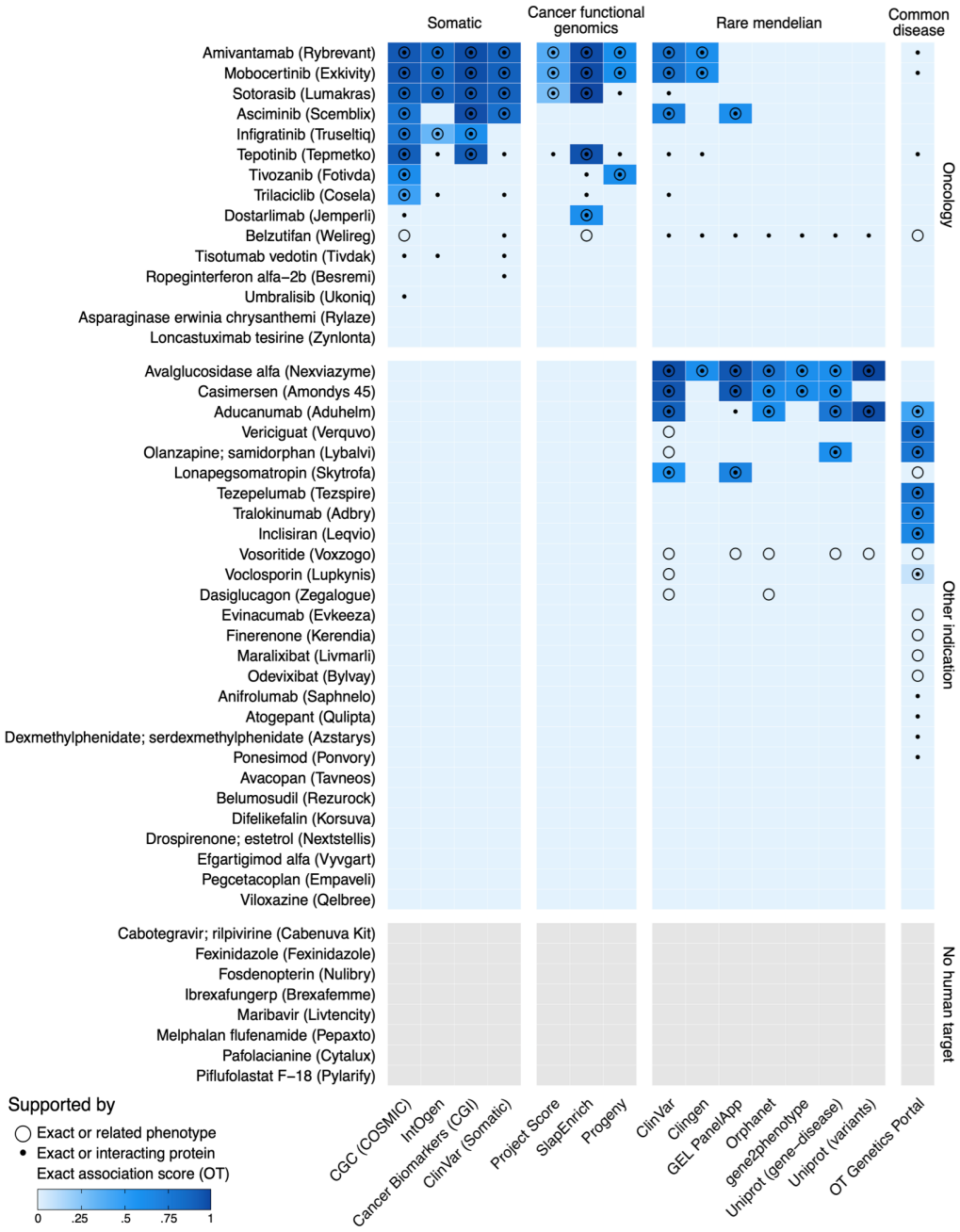
On some occasions, the genetic evidence is not linked to the specific indication but to a strongly related trait. We manually included related traits to the drug indication with availability of genetic evidence for the drug target. The full list of related traits includes microalbuminuria (biomarker of chronic kidney disease), glycodeoxycholate sulfate (one of the bile acids that cause pruritus), height (growth delay) and coronary artery disease (myocardial infarction). These cases cover an additional list of 7 drugs otherwise not captured by direct association between the target and the disease.

Finally, we also illustrated cases in which there is genetic evidence supporting a physically interacting protein. This type of approach has been systematically leveraged in previous research[2]. Here we established a stringent criterion in which we required a relatively strong support (MI>0.42) for physical (and not functional) interactions as reported by the Intact database. 7 more drugs were exclusively supported by this type of evidence.

The analysis was derived from the data aggregated and harmonised by the Open Targets Platform[3]. The code and data necessary to reproduce the figures can be found here https://gist.github.com/d0choa/499c98bd205b39c98304ee603d034546

References

1. Mullard, A. 2021 FDA approvals. *Nat. Rev. Drug Discov.* **21**, 83–88 (2022).
2. MacNamara, A. *et al.* Network and pathway expansion of genetic disease associations identifies successful drug targets. *Sci. Rep.* **10**, 20970 (2020).
3. Ochoa, D. *et al.* Open Targets Platform: supporting systematic drug-target identification and prioritisation. *Nucleic Acids Res.* **49**, D1302–D1310 (2021).



Supplementary Figure 1 | Supporting genetic evidence for 50 drugs approved by the FDA in 2021. Evidence supporting drug target (T)–disease indication (D) association from the Open Targets Platform[3] (https://platform.opentargets.org, November 2021). Evidence for each drug is divided by oncology drugs (top panel), drugs for other indications with human targets (middle panel) and drugs without a human target (predominantly infectious disease drugs or imaging agents; bottom panel). The x-axis is grouped by the predominant genetic data in the respective resources. Support by data source is displayed as: a) evidence for T and D or any other descendant in the EFO ontology aggregated using the Open Targets Platform association score (blue heatmap colours); b) evidence for T and D or closely related phenotypes to D added by manual curation (open circles); and/or c) evidence for D and T or any protein physically interacting with T (IntAct database MI score > 0.42; black dots). All drugs with at least one of the 3 types of support (a,b,c) in any data source are counted as supported by genetic evidence (n = 33).