The essentiality of drug targets: an analysis of current literature and genomic databases

Xiao Ji, Deepak K. Rajpal and Johannes M. Freudenberg

Computational Biology, GSK R&D, Collegeville, PA, USA

Essential genes encode proteins thought to be crucial for the survival of an organism. However, the role of essential genes in human disease and their suitability as drug targets is less clear. Here, we use a recent catalog of nearly 9000 known essential and nonessential genes to evaluate their involvement as therapeutic targets in human diseases. We find that essential genes are more likely than nonessential genes to play a part in specific therapeutic areas such as cardiovascular diseases and neoplasms. We also find significant differences between essential and nonessential genes among protein classes relevant to drug discovery. Taken together, our analyses suggest that the essentiality status of a potential new target is an important consideration in drug discovery.

Introduction

Essential genes (EGs) are a subset of genes in the genome that have central functional roles and are required for the survival of an organism. As such, they are involved in crucial developmental, metabolic and signaling pathways, tend to be located more centrally in protein–protein interaction networks, are more highly expressed, and are more conserved throughout evolution [1]. However, less is known about the essentiality of drug targets. It has been suggested that drugs targeting nonessential proteins have higher approval rates [2], probably because these proteins are safer to target [1]. By contrast, EGs have been suggested as potential new targets for oncology drugs [3]. To further elucidate whether the essentiality of a gene is a factor for determining its suitability as a therapeutic drug target, we set out to more systematically assess currently known EGs and nonessential genes (NEGs).

To determine the essentiality of a gene, a number of approaches have been applied such as induced mutagenesis, gene knockdown and gene knockout experiments in model organisms including *Saccharomyces cerevisiae* (budding yeast) [4,5], *Caenorhabditis elegans* [6–8], *Drosophila melanogaster* (fruit fly) [9–11] and *Mus musculus* (house mouse) [12–14]. In mouse, which is often used for translational research, ~30% of the genes are estimated to be essential for viability [13,14]. In human, RNAi libraries targeting the human genome enabled earlier studies to identify EGs in human cell lines [15–17]. More recently, systematic efforts to uncover EGs in human cell lines through genome-wide screens involving the CRISPR/Cas9 gene editing system have been described [18–20]. A comprehensive list of human EGs has been reported [21] by combining the legacy mouse phenotyping data from mouse genome informatics (MGI) [12], the newly uncovered lethal and sub-viable mouse genes from the International Mouse Phenotyping Consortium [14], and EGs from three human cell line CRISPR/Cas9-based studies [18–20]. This list comprises 3915 EGs that cause pre-weaning lethality in knockout mice or impede cell growth in human cell lines with CRISPR knockouts, as well as 4919 NEGs that produce viable phenotypes in knockout mice [21]. Interestingly, there is considerable overlap and remarkable concordance between the results from mouse-model- and human-cell-line-based approaches [14].

The role of gene essentiality in human disease is somewhat contradictory. Earlier studies of human orthologs of EGs in the mouse suggested that the majority of human disease genes are nonessential, because genetic mutations in EGs prevent viability and thus do not contribute to acquired human disease [22–25]. By contrast, some of these studies also showed evidence suggesting essentiality of human disease genes, such as high connectivity in gene networks [24] and evolutionary conservation [22]. More-recent studies of human orthologs of EGs in mouse [26,27] and

Corresponding author: Freudenberg, J.M. (johannes.m.freudenberg@gsk.com)

1359-6446/
https://doi.org/10.1016/j.drudis.2018.11.002

the analysis of EGs in published genome-wide association studies [14] began to redefine the understanding of gene essentiality in human disease implying a larger overlap between EGs and known human disease genes than previously thought.

Given the involvement of gene essentiality in human disease, the question arises whether EGs are suitable targets for drug discovery. It has been shown that in protein–protein interaction (PPI) networks essential proteins and drug targets have more interactors than other proteins in the PPI network [28]. That same study also found that drug targets have less co-expression with other genes and higher tissue specificity compared with EGs, suggesting that drug targets do not necessarily show a trend toward greater essentiality [28]. Furthermore, it has been argued that drugs targeting EGs are associated with stronger unintended side effects [29], although further research is needed to confirm this conjecture. What complicates matters further is the observation that currently known EGs and NEGs are over-represented among human drug targets at a similar level compared with genes where essentiality is not currently known [30].

Here, we further explore the following questions: (i) should gene essentiality be considered as a feature to select potential new successful drug targets and (ii) if so, are there any differences between specific therapeutic areas or target classes? We use an experimentally derived list of 3915 EGs and 4919 NEGs [14,21] compiled from experiments in knockout mice [12,14] and human cell lines with CRISPR knockouts [18–20]. To assess the properties of known and predicted drug targets, we also obtained data from: the Pharmaprojects database [31], which captures marketed drugs and drugs in development; the DGIdb database, which compiles gene categories associated with druggability [32]; and the Open Targets (OT) platform [33], which contains any available evidence for gene–disease associations. We mapped the experimental factor ontology (EFO) terms that were originally used in OT to MeSH terms and we used top-level MeSH disease terms to define therapeutic areas.

EGs and NEGs have more human disease associations than genes of undetermined essentiality status

The OT database reflects the depth and breadth of currently available evidence supporting a given target–disease relationship and also provides scores summarizing the accumulated evidence in a single number ranging from 0 (no evidence) to 1 (strong evidence) [33]. Using the OT database, we compared the strength of evidence for the association of EGs and NEGs with human disease terms. First, we extracted 2 698 330 gene–disease pairs among 31 149 gene symbols and 8901 disease labels and corresponding OT scores, among which 513 916 were EG–disease pairs and 623 672 were NEG–disease pairs. The median number of disease associations with any evidence recorded in OT was 158 for EGs and 162 for NEGs (see Fig. S1 in supplementary material online). No significant difference in the number of disease associations was observed between EGs and NEGs ($P = 0.28$; two-sided Wilcoxon rank-sum test). Interestingly, the median number of disease associations for genes with unavailable essentiality status was 49, which was significantly less than that of EGs or NEGs ($P < 2.2 \times 10^{-16}$; two-sided Wilcoxon rank-sum test). Moreover, the OT overall scores of EGs and those of NEGs were significantly higher than those of genes with unavailable essentiality status ($P < 2.2 \times 10^{-16}$; two-sided Wilcoxon rank-sum test; see Fig. S2 in supplementary material online). These results strongly suggest that the set of genes with currently known essentiality status (i.e., EGs or NEGs) is biased toward genes that have been more extensively studied for their associations with diseases. This could be caused by biases in the selection of genes for experimental follow-up studies with CRISPR screens, for example. Conversely, genes for which the essentiality status is known have readily accessible documented phenotype associations and therefore are more likely to be included in subsequent studies reporting their associations with disease. A recent effort has defined gene essentiality in terms of evolutionary conservation or, more specifically, intolerance to loss-of-function variants in terms of viability or fitness of the organism [1], which will enable further exploration of genes with currently unknown essentiality.

EGs have higher disease association scores than NEGs

Given that that there appeared to be more known disease associations with EGs and NEGs, we next examined the strength of the evidence in OT supporting such associations. We found that the OT overall scores of EG–disease pairs are significantly higher than those of NEG–disease pairs ($P < 2.2 \times 10^{-16}$; two-sided Wilcoxon rank-sum test). In particular, there was an enrichment of top OT overall scores (defined as scores in the upper 75th percentile) among EG–disease pairs compared with NEG–disease pairs (see Fig. S2 in supplementary material online; odds ratio = 1.23, $P < 2.2 \times 10^{-16}$; two-sided Fisher's exact test). We then asked which data types included in the OT platform contribute to the increased OT overall scores in EG–disease pairs. The OT overall score combines scores for seven different evidence types and we found that EG–disease pairs were associated with higher OT scores in six out of the seven evidence types in OT including 'genetic association', 'somatic mutation', 'animal model', 'known drug', 'affected pathway' and 'literature' (Fig. 1a). Only the 'RNA expression' OT scores, which measure the changes of gene expression levels between normal and corresponding disease samples, tended to be lower for EG–disease pairs than those of NEG–disease pairs.

EGs have higher disease-association scores in certain therapeutic areas

Overall, because EGs are neither enriched nor depleted among disease-associated genes (when compared with NEGs), as well as in the subset of genes that are targeted by currently approved drugs, we next systematically assessed the distribution of evidence on EG–disease associations across different therapeutic areas. It had previously been shown that human disease genes with essential mouse orthologs are associated with carcinomas [26]. We therefore compared the OT overall scores of EG–disease pairs and NEG–disease pairs categorized by specific therapeutic areas (Fig. 1b) as defined by the top-level disease terms from the Medical Subject Headings (MeSH) Browser [34]. We identified 12 therapeutic areas with significantly higher OT overall scores in EG–disease pairs compared with NEG–disease pairs, as well as eight therapeutic areas with significantly lower OT overall scores in EG–disease pairs (false discovery rate, FDR <0.1; two-sided Wilcoxon rank-sum test). The disease term 'congenital, hereditary and neonatal diseases and abnormalities' is the therapeutic area with the greatest

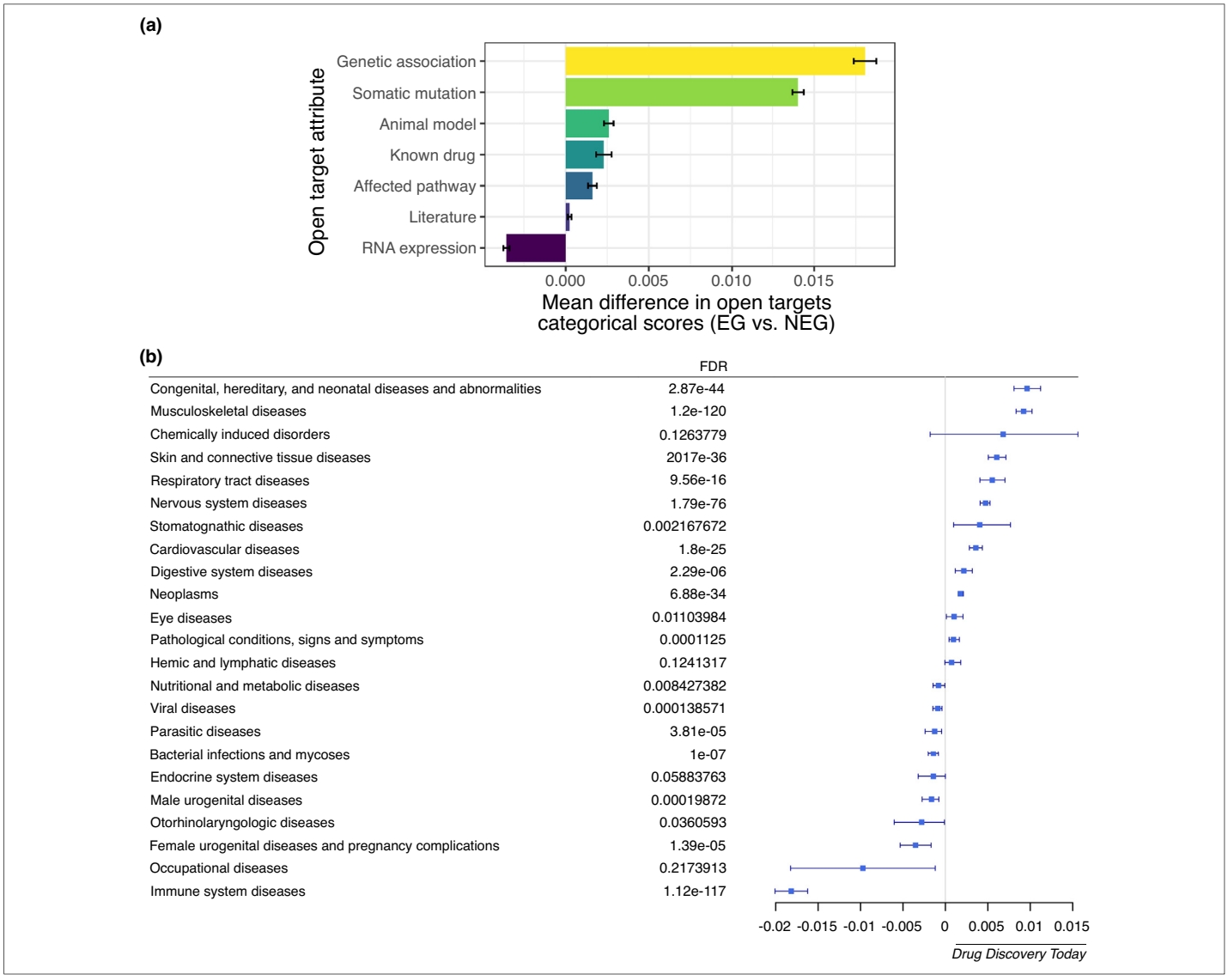

FIGURE 1

Comparison between the Open Targets (OT) scores of essential genes (EGs) and nonessential genes (NEGs). **(a)** Differences in OT scores between EG–disease pairs and NEG–disease pairs by evidence type. The OT platform provides scores summarizing the accumulated evidence for a target–disease pair in a single number ranging from 0 (no evidence) to 1 (strong evidence). The OT overall score combines scores for seven different evidence types. We found that EG–disease pairs were associated with higher OT scores in six out of the seven evidence types in OT including 'genetic association', 'somatic mutation', 'animal model', 'known drug', 'affected pathway' and 'literature'. The mean difference (and their 95% confidence intervals) in OT evidence scores between EG–disease pairs and NEG–disease pairs is shown. The mean difference between EG–disease pairs and NEG–disease pairs in all seven categorical scores was statistically significant (Bonferroni-adjusted P value <0.05; two-sided Student's t-test). **(b)** Differences in OT overall scores between EG–disease pairs and NEG–disease pairs stratified by therapeutic areas. Twenty-three therapeutic areas represented by top-level MeSH disease terms are plotted. The effect sizes and their 95% confident intervals is based on the two-sided Wilcoxon rank-sum test. Gene–disease pairs in therapeutic areas such as musculoskeletal diseases show enrichment of essential genes whereas other areas such as immune system diseases have an over-representation of NEGs.

difference in the OT overall scores between EG–disease pairs and NEG–disease pairs, which correlates the definition of EGs (i.e., genes that cause pre- or neo-natal organismal lethality or cell death when knocked out) with their importance during the prenatal development of an organism [14,21]. Other therapeutic areas where EG–disease pairs were associated with higher OT overall scores (e.g., musculoskeletal diseases, skin and connective tissue diseases, respiratory tract diseases, nervous system diseases, stomatognathic diseases, cardiovascular diseases, digestive system diseases, hemic and lymphatic diseases) involve organs or tissues that play crucial functional parts in the human body. In addition, there was strong evidence supporting the greater association of EG targets to neoplasm (FDR = 6.88×10^{-34}), which suggests that cancer cells tend to be susceptible to genetic or expression changes of EGs. By contrast, the therapeutic areas where EG–disease pairs had lower OT overall scores than NEG–disease pairs include infectious diseases (including viral diseases, parasitic diseases, bacterial infections and mycoses), male and female urogenital diseases, nutritional and metabolic diseases, as well as immune system diseases. It is expected that EGs were less likely to be associated with immune system diseases owing to the possible redundancy of genes involved in maintaining an effective human immune sys-

tem [35]. Other reasons could stem from biases resulting from the experimental methods used to determine gene essentiality. That is, a developing prenatal mammalian organism or an *in vitro* system does not require an immune system.

Approved drugs targeting EGs tend to have more indications

Next, we focused our analyses more specifically on known drug targets using a pharmaceutical industry pipeline database [36] from which we extracted records for 25 090 target–indication pairs between 2140 unique gene targets and 859 unique diseases [37]. Among them, 2473 target–indication pairs were associated with successfully launched drugs (we refer to them as 'successful target–disease pairs'); 1672 pairs failed in the clinic at Phase III, 6599 pairs failed at Phase II, 3803 pairs failed at Phase I, 10 478 failed at the preclinical phase and 65 pairs were withdrawn because of other reasons. Among the successful target–disease pairs, there were 1049 EG–disease pairs and 1292 NEG–disease pairs. Among the 22 522 target–indication pairs that failed in preclinical or clinical phases (we refer to them as 'failed target–disease pairs') there were 9351 EG–disease pairs and 11 448 NEG–disease pairs. From the 2473 successful target–indication pairs from the Pharmaprojects data we identified 170 unique EG targets, 240 unique NEG targets and 31 targets for which the essentiality status was unavailable. Although the median number of associated indications for 170 EG targets and that for 240 NEG targets were equal (median = 3), we observed a moderate enrichment of successful targets associated with more than three diseases among the 170 EGs compared with the 240 NEGs (79 EGs vs 89 NEGs; $P = 0.036$, odds ratio = 1.47; one-sided Fisher's exact test; see Fig. S3 in supplementary material online).

EG targets are more likely to be successful in certain therapeutic areas

Using 22 522 failed target–indication pairs as background, we did not observe an enrichment or depletion of EG–disease pairs among 2473 successful target–disease pairs ($P = 0.895$, odds ratio = 0.994; two-sided Fisher's exact test) suggesting that essentiality by itself does not predict clinical success or failure overall. However, given the differences in therapeutic areas described above, we investigated whether EG–disease pairs were enriched among successful target–disease pairs in specific therapeutic areas represented by top-level MeSH disease terms. Among the 26 therapeutic areas, we selected 18 areas that contained at least 50 target–disease pairs in the Pharmaprojects data for a meta-analysis (Fig. 2a). We identified five therapeutic areas with significant enrichment of successful EG–disease pairs compared with NEG–disease pairs (FDR <0.1; two-sided Fisher's exact test), including cardiovascular diseases (FDR = 1.15×10^{-4}), neoplasms (FDR = 0.04), hemic and lymphatic diseases (FDR = 0.09), pathological conditions, signs and symptoms (FDR = 0.092), and nervous system diseases (FDR = 0.09). In addition, musculoskeletal diseases showed marginal enrichment of successful EG–disease pairs (FDR = 0.103). These results are consistent with our EG–disease association analysis using the OT data (Fig. 1b). Similarly, in the therapeutic areas of male urogenital diseases and viral diseases we also observed a decreased amount of evidence for EG–disease association and depletion of successful EG–indication pairs. These results indicated that our findings are generally consistent across OT data and Pharmaprojects data.

In the top two therapeutic areas with significant enrichment of successful EG–indication pairs from the Pharmaprojects data (i.e., cardiovascular diseases and neoplasms), we compared the success rates between EG–disease pairs and NEG–disease pairs and the ratio between EG–disease pairs and NEG–disease pairs across different phases in the drug development process. In cardiovascular diseases, the success rates (defined as the proportion of target–disease pairs that successfully moved to the next phase of development) of EG–disease pairs were consistently higher than those of NEG–disease pairs across all drug developmental phases (Fig. 2b). The ratio between EG–disease pairs and NEG–disease pairs (represented by the proportion of EG–disease pairs among the total number of EG–disease pairs and NEG–disease pairs) increased along with the progression in the drug development pipeline. Remarkably, the proportion of EG–disease pairs increased from 39.3% in the preclinical phase to 57.4% in the final approval stage (Fig. 2c). In neoplasms, higher success rates of EG–disease pairs were observed only in the preclinical phase and clinical Phase I and II, whereas the success rate of EG–disease pairs is comparable with that of NEG–disease pairs in clinical Phase III (Fig. 2d). Similarly, the ratios between EG–disease pairs and NEG–disease pairs increased as the drug development pipeline progressed from the preclinical phase to clinical Phase III. The proportions of EG–disease pairs in Phase III and the final approval stage were as high as 65% (Fig. 2e). These results suggest that gene essentiality is predictive of successful drug targets along the process of the drug development pipeline in specific therapeutic areas including cardiovascular diseases and neoplasms. In summary, when considering different therapeutic areas separately, we observed a consistent trend between the disease association data and the pharma pipeline data, where we observed EGs to be more likely to play a part in the disease and as drug targets in cardiovascular diseases, neoplasms, nervous system diseases and musculoskeletal diseases, whereas NEGs are more likely to play a part in infectious diseases and urogenital diseases (Figs. 1 b, 2 a).

EGs are under-represented among certain druggable protein classes

Although, conceptually, the essentiality of a gene is not directly related to the druggability of the gene or gene product, we were interested in investigating the essentiality of genes from potentially druggable protein families, which were initially defined by Hopkins and Groom [36]. We identified 30 druggable protein classes with at least 50 genes from DGIdb (version 3.0) [32] and evaluated the enrichment of EGs or NEGs among these protein classes (Fig. 3). We observed that EGs were depleted among many known druggable protein families, including G-protein-coupled receptors (GPCRs), hormone receptors, cytochrome P450, ion channels, phospholipases, protease inhibitors, neutral zinc metallopeptidases, transporters and proteases (FDR <0.1; two-sided Fisher's exact test). In addition, druggable proteins located at the cell surface or external side of the plasma membrane were less likely to be encoded by EGs. Notably, we found that EGs were enriched among several druggable gene categories related to cancer treatment including transcription factor complexes, histone modification proteins, transcription factor binding proteins, DNA

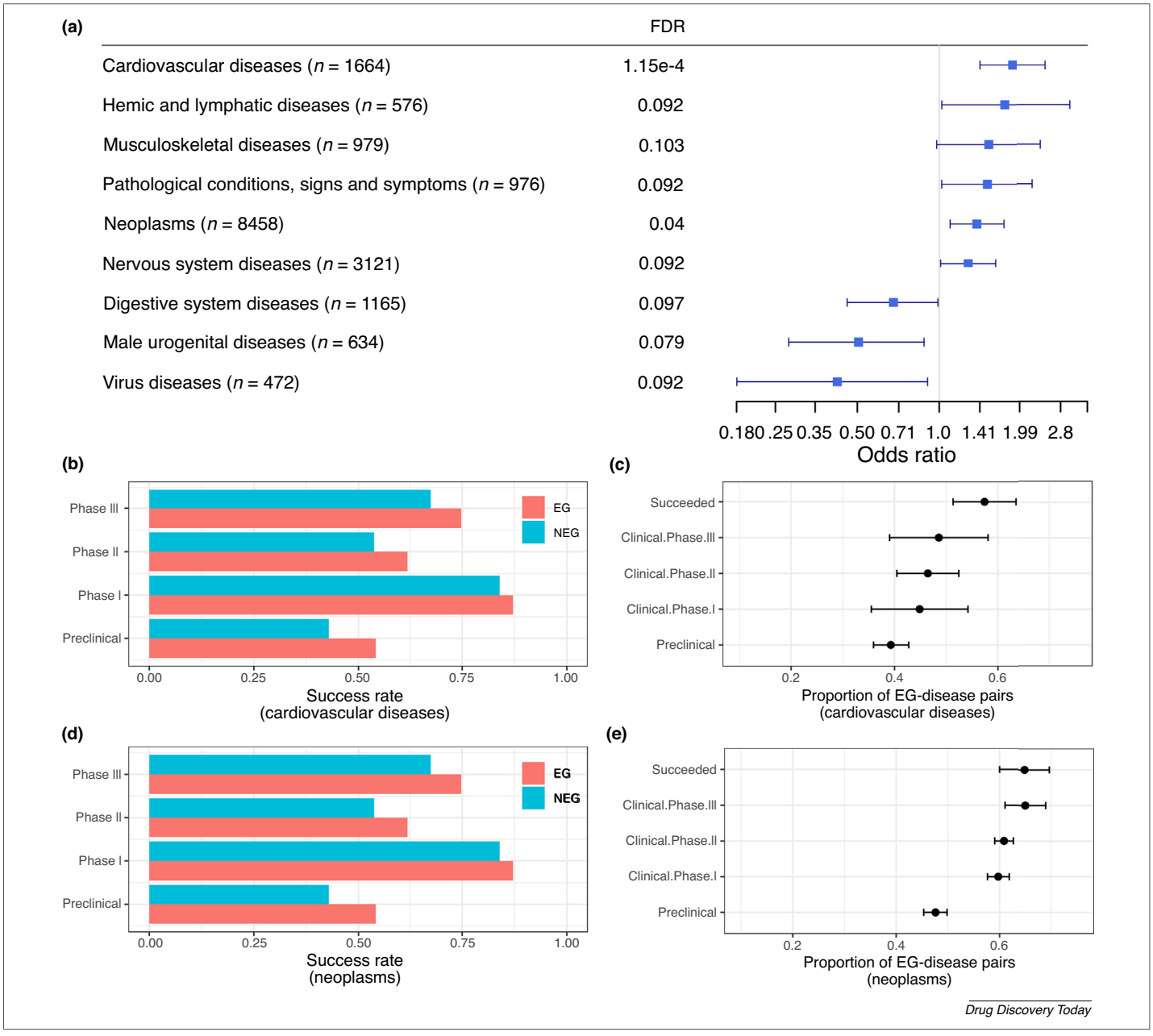

FIGURE 2

Comparison between the pharmaceutical drug development records of essential genes (EGs) and those of nonessential genes (NEGs). **(a)** Enrichment or depletion of EG–disease pairs among successful target pairs in Pharmaprojects data, which capture all available information on marketed drugs and drugs in development including protein targets and indications that were mapped to MeSH terms. For our analysis, we selected 18 therapeutic areas that had at least 50 target–disease pairs in the Pharmaprojects database. The therapeutic areas with enrichment or depletion of EG–disease pairs (FDR = 0.103 for musculoskeletal diseases and FDR <0.1 for other eight therapeutic areas, two-sided Fisher's exact test) are shown. The effect sizes and their 95% confident intervals were derived from the two-sided Fisher's exact test. The total number (*n*) of target–disease pairs in each therapeutic area are indicated. **(b–e)** Success rates and proportion of EG–disease pairs along the process of drug development pipeline. In two specific therapeutic areas (i.e., cardiovascular diseases and neoplasm) the success rates between EG–disease pairs (in red) and NEG–disease pairs (in turquoise) (b,d) and the proportions of EG–disease pairs among the sum of EG–disease pairs and NEG–disease pairs across different phases in the drug development process (along with their 95% confidence intervals) (c,e) are shown.

polymerases, proteins in DNA repair apparatus and tumor suppressors (FDR <0.1; two-sided Fisher's exact test). Moreover, we found that EGs were enriched among 399 'clinically actionable' genes (FDR = 2.35×10^{-23}; two-sided Fisher's exact test) that were known to undergo somatic genomic mutations in cancer and were actively being used in targeted clinical sequencing panels for individualized cancer treatment [32,38,39].

A common strategy for cancer treatment involves targeting cell growth, cell division and cell differentiation processes including DNA and RNA synthesis, transcriptional regulation and epigenetic regulation [40]. We observed over-representation of EGs among several druggable gene categories listed in DGIdb including transcription factor complexes, histone modification proteins, transcription factor binding proteins, DNA polymerases, proteins in

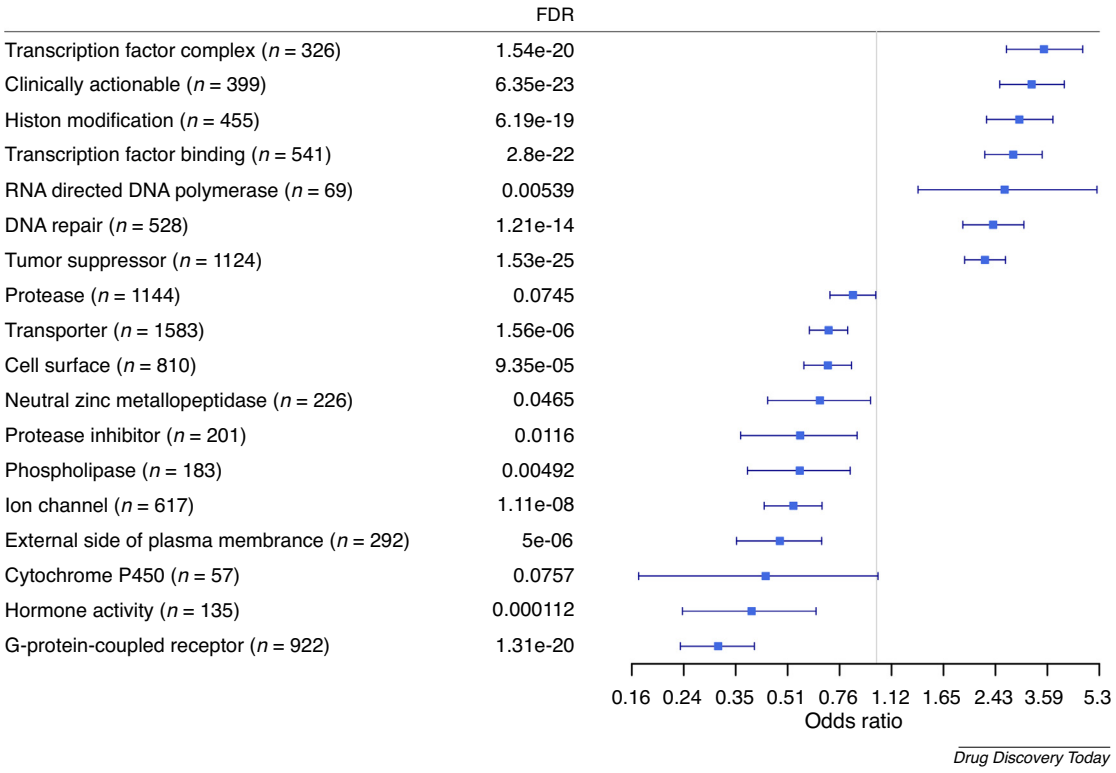

FIGURE 3

Enrichment or depletion of essential genes (EGs) among druggable genes categorized by DGIdb. The DGIdb database compiles protein classes associated with druggability [such as G-protein-coupled receptor (GPCR), ion channel, etc.]. EGs and nonessential genes (NEGs) were mapped to these annotations. Eighteen categories with enrichment or depletion of EGs (FDR <0.1, two-sided Fisher's exact test) are shown. The effect sizes and their 95% confident intervals came from two-sided Fisher's exact test. The total number (*n*) of genes in each druggable gene category was indicated. Categories such as transcription factor complex and histone modification show strong enrichment of essential genes while categories such as GPCRs show a strong over-representation of NEGs.

DNA repair apparatus and tumor suppressors, which could explain why EGs are more likely to be targets of approved cancer drugs. A potential extension of our analysis could be to integrate additional data sources to determine whether genes found to be essential in human cancer cell lines, but not normal cell types, are suitable targets of anticancer drugs. A closely related concept in this space is synthetic lethality, which has recently received renewed interest potentially leading to new drugs for patients with cancerous loss-of-function mutations [41]. Interestingly, cardiovascular diseases, nervous system diseases and musculoskeletal diseases are also therapeutic areas where EGs are more likely than NEGs to be associated as well as to be targets of approved drugs.

Concluding remarks and future directions

Based on our analysis, gene essentiality is an important factor that, by itself however, is not predictive of the success or failure of a potential new drug target, suggesting that other factors need to be considered as well. These include the protein families, such as GPCRs and ion channels, which are overrepresented among NEGs. By contrast, transcription factors and histone modifiers are more likely to be essential but they are only recently being considered to be more amenable to pharmacological intervention [42]. Another important consideration is the potential safety implication of EG targets that have the tendency to play a more central part in protein–protein interaction networks, potentially leading to more-pleiotropic effects and an increased number of side effects. Accordingly, drugs that target more-highly-connected genes (such as EGs) tend to be used as therapeutics for more-severe diseases such as cancer and autoimmune disorders [29]. Lastly, although EGs fall across the continuum of gene expression between being tissue-specific and ubiquitously expressed, they are enriched for the housekeeping genes that are ubiquitously expressed across tissues [27].

In conclusion, with a current list of EGs and NEGs, we find that gene essentiality is an important consideration when evaluating potential new drug targets, which affects different therapeutic areas in different ways. In cardiovascular diseases, neoplasms, nervous system diseases and musculoskeletal disease EGs are more likely to be successful drug targets; whereas in other areas, such as infectious diseases and urogenital diseases, NEGs are more likely than EGs to be targets of approved drugs. Recent efforts have aimed to systematically quantify the probability of success of a target–indication pair using model features such as genetic [43] and transcriptomic [44] data. Our analysis shows that gene essentiality is an additional, similarly important target characteristic that should be incorporated in these quantitative models.

Acknowledgments

We would like to thank Pankaj Agarwal and Philippe Sanseau for their comments and insight and critical review of the manuscript. We would also like Mark Hurle for providing the Pharmaprojects dataset. Finally, we would like to thank the anonymous reviewers for their thoughtful comments and valuable feedback.

Conflicts of interest

All authors are employees and shareholders of GlaxoSmithKline.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.drudis.2018.11.002.

References

1 Bartha, I. *et al.* (2018) Human gene essentiality. *Nat. Rev. Genet.* 19, 51–62
2 Narasimhan, V.M. *et al.* (2016) Health and population effects of rare gene knockouts in adult humans with related parents. *Science* 352, 474–477
3 Rancati, G. *et al.* (2018) Emerging and evolving concepts in gene essentiality. *Nat. Rev. Genet.* 19, 34–49
4 Giaever, G. *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391
5 Winzeler, E.A. *et al.* (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285, 901–906
6 Clark, D.V. *et al.* (1988) The unc-22(IV) region of *Caenorhabditis elegans*: genetic analysis of lethal mutations. *Genetics* 119, 345–353
7 Johnsen, R.C. and Baillie, D.L. (1991) Genetic analysis of a major segment [LGV (left)] of the genome of *Caenorhabditis elegans*. *Genetics* 129, 735–752
8 Kamath, R.S. and Ahringer, J. (2003) Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* 30, 313–321
9 Bellen, H.J. *et al.* (2004) The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. *Genetics* 167, 761–781
10 Boutros, M. *et al.* (2004) Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* 303, 832–835
11 Dietzl, G. *et al.* (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448, 151–156
12 Eppig, J.T. *et al.* (2005) The Mouse Genome Database (MGD): from genes to mice—a community resource for mouse biology. *Nucleic Acids Res.* 33, D471–D475
13 White, J.K. *et al.* (2013) Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell* 154, 452–464
14 Dickinson, M.E. *et al.* (2016) High-throughput discovery of novel developmental phenotypes. *Nature* 537, 508–514
15 Harborth, J. *et al.* (2001) Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J. Cell Sci.* 114, 4557–4565
16 Luo, B. *et al.* (2008) Highly parallel identification of essential genes in cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* 105, 20380–20385
17 Silva, J.M. *et al.* (2008) Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* 319, 617–620
18 Blomen, V.A. *et al.* (2015) Gene essentiality and synthetic lethality in haploid human cells. *Science* 350, 1092–1096
19 Hart, T. *et al.* (2015) High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* 163, 1515–1526
20 Wang, T. *et al.* (2015) Identification and characterization of essential genes in the human genome. *Science* 350, 1096–1101
21 Ji, X. *et al.* (2016) Increased burden of deleterious variants in essential genes in autism spectrum disorder. *Proc. Natl. Acad. Sci. U. S. A.* 113, 15054–15059
22 Domazet-Loso, T. and Tautz, D. (2008) An ancient evolutionary origin of genes associated with human genetic diseases. *Mol. Biol. Evol.* 25, 2699–2707
23 Feldman, I. *et al.* (2008) Network properties of genes harboring inherited disease mutations. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4323–4328
24 Goh, K.I. *et al.* (2007) The human disease network. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8685–8690
25 Park, D. *et al.* (2008) Analysis of human disease genes in the context of gene essentiality. *Genomics* 92, 414–418
26 Dickerson, J.E. *et al.* (2011) Defining the role of essential genes in human disease. *PLoS One* 6, e27368
27 Georgi, B. *et al.* (2013) From mouse to human: evolutionary genomics analysis of human orthologs of essential genes. *PLoS Genet.* 9, e1003484
28 Yildirim, M.A. *et al.* (2007) Drug–target network. *Nat. Biotechnol.* 25, 1119–1126
29 Yao, L. and Rzhetsky, A. (2008) Quantitative systems-level determinants of human genes targeted by successful drugs. *Genome Res.* 18, 206–213
30 Kim, B. *et al.* (2017) *In silico* re-identification of properties of drug target proteins. *BMC Bioinformatics* 18, 248
31 Informa (2018) Pharmaprojects Database. Available at: https://citeline.com/products/pharmaprojects
32 Griffith, M. *et al.* (2013) DGIdb: mining the druggable genome. *Nat. Methods* 10, 1209–1210
33 Koscielny, G. *et al.* (2017) Open Targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res.* 45, D985–D994
34 Sorden, N. (2016) New MeSH browser available. *NLM Tech. Bull.* 413, e2
35 Fischer, A. and Rausell, A. (2016) Primary immunodeficiencies suggest redundancy within the human immune system. *Sci. Immunol* 1, 5861
36 Hopkins, A.L. and Groom, C.R. (2002) The druggable genome. *Nat. Rev. Drug Discov.* 1, 727–730
37 Yao, J. *et al.* (2018) Predicting clinically promising therapeutic hypotheses using tensor factorization. *Bioinformatics* http://dx.doi.org/10.1101/272740
38 Wagle, N. *et al.* (2012) High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov.* 2, 82–93
39 Cheng, D.T. *et al.* (2015) Memorial Sloan Kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J. Mol. Diagn.* 17, 251–264
40 Kumar, S. *et al.* (2015) Drug targets for cancer treatment: an overview. *Med. Chem.* 5, 115–123
41 Mullard, A. (2017) Synthetic lethality screens point the way to new cancer drug targets. *Nat. Rev. Drug Discov.* 16, 589–591
42 Papavassiliou, K.A. and Papavassiliou, A.G. (2016) Transcription factor drug targets. *J. Cell Biochem.* 117, 2693–2696
43 Nelson, M.R. *et al.* (2015) The support of human genetic evidence for approved drug indications. *Nat. Genet.* 47, 856–860
44 Rouillard, A.D. *et al.* (2018) Systematic interrogation of diverse omic data reveals interpretable, robust, and generalizable transcriptomic features of clinically successful therapeutic targets. *PLoS Comput. Biol.* 14, e1006142