

## Article

# In Silico Pleiotropy Analysis in KEGG Signaling Networks Using a Boolean Network Model

Maulida Mazaya <sup>1,†</sup>  and Yung-Keun Kwon <sup>2,\*</sup> 

<sup>1</sup> Research Center for Computing, National Research and Innovation Agency (BRIN), Cibinong Science Center, Jl. Raya Jakarta-Bogor KM 46, Cibinong 16911, West Java, Indonesia

<sup>2</sup> School of IT Convergence, University of Ulsan, 93 Daehak-ro, Nam-gu, Ulsan 44610, Korea

\* Correspondence: kwonyk@ulsan.ac.kr

† This author is the main contributor to this work.

**Abstract:** Pleiotropy, which refers to the ability of different mutations on the same gene to cause different pathological effects in human genetic diseases, is important in understanding system-level biological diseases. Although some biological experiments have been proposed, still little is known about pleiotropy on gene–gene dynamics, since most previous studies have been based on correlation analysis. Therefore, a new perspective is needed to investigate pleiotropy in terms of gene–gene dynamical characteristics. To quantify pleiotropy in terms of network dynamics, we propose a measure called in silico Pleiotropic Scores (*sPS*), which represents how much a gene is affected against a pair of different types of mutations on a Boolean network model. We found that our model can identify more candidate pleiotropic genes that are not known to be pleiotropic than the experimental database. In addition, we found that many types of functionally important genes tend to have higher *sPS* values than other genes; in other words, they are more pleiotropic. We investigated the relations of *sPS* with the structural properties in the signaling network and found that there are highly positive relations to degree, feedback loops, and centrality measures. This implies that the structural characteristics are principles to identify new pleiotropic genes. Finally, we found some biological evidence showing that *sPS* analysis is relevant to the real pleiotropic data and can be considered a novel candidate for pleiotropic gene research. Taken together, our results can be used to understand the dynamics pleiotropic characteristics in complex biological systems in terms of gene–phenotype relations.

**Keywords:** pleiotropy; gene–gene interactions; Boolean network dynamics; signaling networks; feedback loops



**Citation:** Mazaya, M.; Kwon, Y.-K. In Silico Pleiotropy Analysis in KEGG Signaling Networks Using a Boolean Network Model. *Biomolecules* **2022**, *12*, 1139. <https://doi.org/10.3390/biom12081139>

Academic Editors: Vladimir N. Uversky and Matteo Floris

Received: 20 July 2022

Accepted: 15 August 2022

Published: 18 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Pleiotropy is the phenomenon in which one gene can result in multiple phenotypes or traits [1–3]. In human genetic diseases, it means that different mutations within the same gene cause different pathological effects [4,5]. This becomes an important contributor in identifying a novel function of individual genes with respect to gene–gene interactions [6,7] in system-level biological diseases [8,9]. In this regard, many methods have been modeled to understand the pleiotropy. For example, an experimental study [10] performed several laboratory cultures with the *nath-10 polymorphism* and explained its pleiotropic role in the evolution of a cryptic genetic variation in *C. elegans*. In another study, a statistical analysis [11] using canonical correlation analysis identified a novel candidate pleiotropic associations between genetic variants and phenotypes. In addition, a few computational models [12,13] deployed a pairwise combination of genome-wide association data from the complex disease pleiotropy analysis and modular gene expression analysis. Another study using a metabolic model [14] conducted constraint-based simulations for *E. coli* and *S. cerevisiae* and found that pleiotropy is an emergent property of metabolic network.

Finally, protein–protein interaction network analysis was also widely used, and some studies [15–17] confirmed pleiotropic effects in biological molecular function, which lead to complex diseases. Despite the interesting observations in previous studies, most of the previous approaches focused on the pleiotropy analysis induced by undirected molecular correlation networks. Therefore, a new approach is needed to investigate the pleiotropy induced by a directed signaling network because it can explain the pleiotropy caused by a gene–gene dynamical relationship.

To quantify the pleiotropic degree of a gene in terms of network dynamics, we proposed in silico pleiotropic score (*sPS*), which is a measure to represent how differently a gene affects the dynamics of other genes against different mutations, such as knock-out [18,19] and over-expression [20] mutations. In this study, we employed the Boolean network model [21,22] to simulate the network dynamics. A Boolean network model implicitly assumes that all biological components are described by binary values, and their interactions represented by Boolean regulatory functions [23], and it is well-known to capture the silent dynamical properties of biological networks [23,24]. For example, it has been used to analyze oncogene rules in *Non-small cell lung cancer* [25], to model the *C. albicans* yeast for hyphal transition [26], to a matrix cell density sensing to contact inhibition, proliferation, migration, and apoptosis [27], or to illustrate the regulatory effects in cervical cancer [28]. A previous study showed that the dynamics influence of a gene to another genes has some interesting structural characteristics in the signaling network [29]. This study can be extended because the pleiotropy is understood as the difference of the dynamics influence against different mutation types. In this study, through intensive investigations with a signaling network, we observed that most of the dynamically affected genes were related to the experimentally proved pleiotropic genes [30]. Moreover, we found that *sPS* is negatively correlated with the previous standardized method pleiotropy [2]. Further, we investigated the relationships of *sPS* with structural properties and found that they have highly positive correlations with degree/in-degree/out-degree, feedback loops, and centrality measures such as closeness, betweenness, stress, and eigenvector, in the signaling network. This implies that the more central of genes, the more pleiotropic. Finally, we found some biological evidence confirming that *sPS* analysis is relevant to the experimental pleiotropic database, and it can be used for novel candidate pleiotropic gene characteristics. Through a network visualization, we observed that most novel candidate pleiotropic genes are closely located to the known pleiotropic genes. Taken together, these results help to understand the importance of dynamics pleiotropic in complex biological systems in terms of gene–phenotype relations.

## 2. Materials and Methods

### 2.1. Datasets

To examine the in silico dynamics-based pleiotropy, we employed a dataset of a cellular signaling network with 1659 genes and 7964 interactions [31], which was constructed from Kyoto Encyclopedia of Genes and Genomes network (KEGG) database [32]. We then retrieved the list of the pleiotropic genes from human phenotype ontology (HPO) database [30,33], where (a) a gene  $v_i$  is associated with  $n$  phenotypes, and (b) a phenotype  $p$  is associated with  $n$  genes. In our analysis, this definition was used to sort KEGG genes that are associated with any phenotypes in (a) and (b), and to compute their degree of pleiotropy by the number of phenotypes of those genes. Further, we classified all KEGG genes with respect to the functional importance genes by annotating them with cancer genes, drug-target genes, essential genes, tumor-suppressor genes, oncogenes, and disease genes based on TCGA CBioportal [34,35], DrugBank [35,36], DEG [37], TSGene [38,39], ONGene [40], and DisGenet [41,42] databases, respectively (See Tables S1–S3 in Supplementary Materials). We note that this study is not limited to the pleiotropic genes identified from the public databases, because there are other ways to explore the pleiotropy, for example, lab experiments [43] using high-throughput morphometric analysis of hundreds of thousands of

single cells in the budding yeast *Saccharomyces cerevisiae* or experimental design in *Drosophila simulans* with genome sequencing [44] to study pleiotropy.

### 2.2. Boolean Network Model

To examine pleiotropic dynamics of genes induced by different types of mutations in a large-scale network, we applied a synchronous Boolean network model [23], which is one of the simplest computational methods to elucidate the network dynamics [22] and has been used to examine complicated behaviors of biological networks [45–47]. A Boolean network is represented by a directed graph  $G(V, A)$ , where  $V = \{v_1, v_2, \dots, v_N\}$  is a set of nodes, and  $A \subseteq V \times V$  is a set of directed links. Each  $v_i \in V$  has a value of 1 (on) or 0 (off), which means the possible states of the corresponding elements. A directed link  $(v_i, v_j)$  represents a positive (activating) or a negative (inhibiting) relationship from  $v_i$  to  $v_j$  ( $v_i$  and  $v_j$  are called source and target nodes of the link, respectively). Let  $v(t)$  denote the state of node  $v$  at time step  $t$ . When a state of  $v_i$  at time  $t + 1$  is determined by the values of  $k_i$ , with other nodes  $v_{i_1}, v_{i_2}, \dots, v_{i_{k_i}}$  with a link to  $v_i$  at time  $t$ , the update rule of  $v_i$  is represented by a Boolean function  $f_i : \{0, 1\}^{k_i} \rightarrow \{0, 1\}$ . Then, all nodes are synchronously updated, and here, we implemented a nested canalizing functions (NCFs) model [48,49] to describe an update rule  $f_i$  as follows:

$$f_i(v_{i_1}(t), v_{i_2}(t), \dots, v_{i_{k_i}}(t)) = \begin{cases} O_1 & \text{if } v_{i_1}(t) = I_1 \\ O_2 & \text{if } v_{i_1}(t) \neq I_1 \text{ and } v_{i_2}(t) = I_2 \\ O_3 & \text{if } v_{i_1}(t) \neq I_1 \text{ and } v_{i_2}(t) \neq I_2 \text{ and } v_{i_3}(t) = I_3 \\ & \vdots \\ O_{k_i} & \text{if } v_{i_1}(t) \neq I_1 \dots v_{i_{k_i-1}}(t) \neq I_{k_i-1} \text{ and } v_{i_{k_i}}(t) = I_{k_i} \\ O_{def} & \text{otherwise} \end{cases} \quad (1)$$

where  $I_m$  and  $O_m$  ( $m = 1, 2, \dots, k_i$ ) represent the canalizing and canalized Boolean values, respectively, and  $O_{def}$  is generally set to  $1 - O_{k_i}$ . In addition, we specified all  $I_m$  and  $O_m$  values independently and uniformly at random between 0 and 1. We note that many biological networks were successfully represented by NCFs [50–52], and NCFs also properly fit biological experiments' data [49] including pleiotropy analysis [11]. Those support that NCFs can describe the network dynamics considerably similarly to those real biological networks.

In a Boolean network, a network state at time  $t$  can be denoted by a list of state values of all nodes,  $\mathbf{v}(t) = [v_1(t), v_2(t), \dots, v_N(t)] \in \{0, 1\}^N$ . Next, every network state transits to another network state through a set of Boolean update functions  $F = \{f_1, f_2, \dots, f_N\}$  and eventually converges to either a fixed point or a limit-cycle attractor starting from its initial state. This attractor represents the diverse biological network behaviors such as homeostasis or oscillation. The definition of the attractor is defined as follows.

**Definition 1.** Let  $\mathbf{v}(0), \mathbf{v}(1), \dots$ , be a network state trajectory starting at  $\mathbf{v}(0)$ . Then, the attractor denoted by  $G, F, \mathbf{v}(0)$  is represented by an ordered finite list of network states  $\mathbf{v}(\tau), \mathbf{v}(\tau + 1), \dots, \mathbf{v}(\tau + p - 1)$  where  $\tau$  is the smallest time step such that  $\mathbf{v}(t) = \mathbf{v}(t + p)$  for  $\forall t \geq \tau$  with  $\mathbf{v}(i) \neq \mathbf{v}(j)$  for  $\forall i \neq j \in \{\tau, \tau + 1, \dots, \tau + p - 1\}$ . Herein,  $p$  represents the attractor length.

In this study, the examination of attractors is needed to find the affected genes. The affected genes were obtained based on our previous work about gene–gene dynamics influence networks [29]. To implement this, we specified a set of initial states,  $S$ , and computed a state trajectory starting at every  $\mathbf{v}(0) \in S$  until an attractor is found. We note that the network dynamics can depend on the initial network states.

### 2.3. Computation of In-Silico Pleiotropic Scores (sPS)

Given a gene subject to different types of mutations, we propose sPS of the gene to represent how much the other genes are differentially affected in terms of the dynamics that

can be used to deepen the study of existing measures of pleiotropy analysis. Specifically, we considered two mutations, knockout [19,53] and overexpression [54] mutations. The knockout mutation represents the effect of suppressing the expression of a gene or the pharmaceutical inhibition of the secondary messenger production or kinase/phosphate activity [18,55]. On the other hand, the overexpression mutation represents the effect of gene expression change [20,56]. In the Boolean network model, these mutations describe scenarios where the state of mutated gene is frozen to 0 (off) state and 1 (on) state, respectively, during a mutation duration time  $T$ . In this study,  $T$  is a parameter to denote the mutation duration time, and thus a mutation is effective only for  $t \leq T$ . This mutation duration time is considered important since it can affect the mutation process of molecular interaction networks [57,58]. Taken together, these mutations can be implemented by changing  $F$  into  $F'$  for  $t \leq T$  as follows:

$$F' = \{f_1, \dots, f_{i-1}, \alpha, f_{i+1}, \dots, f_N\} \tag{2}$$

where  $\alpha$  is a set to 0 and 1 in the case of the knockout and the overexpression mutations, respectively. Note that the update rule of  $v_i$  is restored to  $f_i$  after the time step  $T$ .

To compute *sPS* of a gene, we employed the notion of the gene–gene dynamics influence used in the previous work [29]. Given a Boolean network  $G(V, A)$  with a set of nodes  $V = \{v_1, v_2, \dots, v_N\}$  specified by a set of corresponding update rules  $F = \{f_1, f_2, \dots, f_N\}$ , we generate a set of random initial states  $S$ . We first define the dynamics influence of gene  $v_i$  on gene  $v_j$ , which represents how much the states sequence of  $v_j$  is changed by a mutation subject to  $v_i$  as follows (see Figure S1 in the Supplementary Information for an illustrative example):

1. For each initial state  $\mathbf{v}(0) \in S$ , we obtain two attractors  $G, F, \mathbf{v}(0)$  and  $G, F', \mathbf{v}(0)$  in the wild-type and the  $v_i$ -mutant networks, respectively. For convenience, let  $G, F, \mathbf{v}(0) = [\mathbf{v}(\tau), \mathbf{v}(\tau + 1), \dots, \mathbf{v}(\tau + p - 1)]$  and  $G, F', \mathbf{v}(0) = \mathbf{v}'(\tau'), \mathbf{v}'(\tau' + 1), \dots, \mathbf{v}'(\tau' + p' - 1)$ .
2. We compute a distance between  $G, F, \mathbf{v}(0)_j$  and  $G, F', \mathbf{v}(0)_j$  defined as follows:

$$d(\mathbf{v}(0), v_i, v_j) = \min_{m \in [0, e-1]} \frac{\sum_{l=0}^{c-1} I(v_j(\tau + l + m) \neq v'_j(\tau' + l))}{c} \tag{3}$$

where  $c$  and  $e$  are the least common multiple and the greatest common divisor, respectively, of  $p$  and  $p'$ , and  $I(\text{condition})$  is an indicator function where outputs 1 if *condition* is true, and 0 otherwise. As a result,  $d(\mathbf{v}(0), v_i, v_j)$  represents the minimum ratio of a bitwise difference between the states sequence of  $v_j$  in the wild-type and the  $v_i$ -mutant attractors over the least common period ( $c$ ) of the two attractors.

3. Lastly, we compute the dynamics influence of  $v_i$  on  $v_j$  denoted by  $\mu(v_i, v_j)$  by averaging out  $d(\mathbf{v}(0), v_i, v_j)$  over a set of initial states in  $S$  as follows:

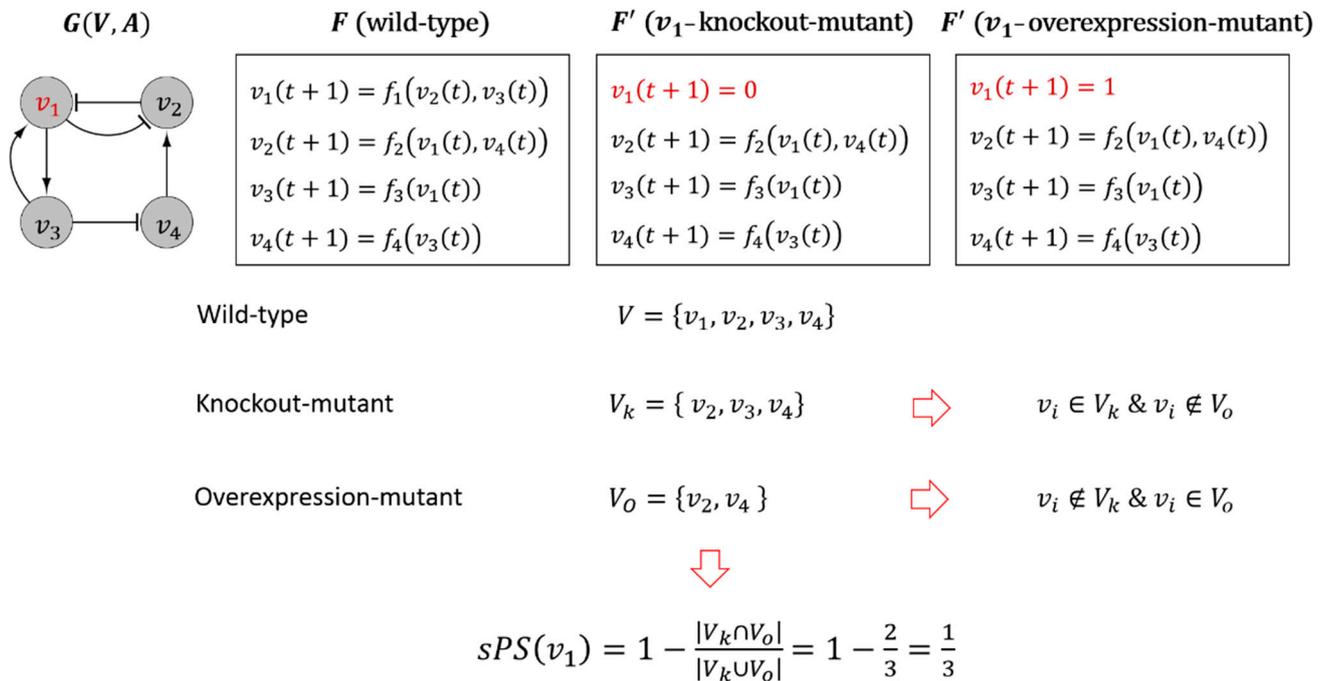
$$\mu(v_i, v_j) = \frac{\sum_{\mathbf{v}(0) \in S} d(\mathbf{v}(0), v_i, v_j)}{|S|} \tag{4}$$

Then, let  $v_i$  an arbitrary source gene in  $V$ . Based on the dynamics influence, we can denote a set of affected genes as  $\{v_j \in V | \mu(v_i, v_j) > 0\}$ . Let  $V_k$  and  $V_o$  the sets of affected genes with respect to the knockout and the overexpression mutations, respectively. We then define *sPS* of a gene  $v_i$  as follows:

$$sPS(v_i) = \frac{1}{|S|} \sum_{\mathbf{v}(0) \in S} \left[ 1 - \frac{|V_k \cap V_o|}{|V_k \cup V_o|} \right]. \tag{5}$$

It represents the proportion of the genes that are included in the symmetric difference of  $V_k$  and  $V_o$  among their union. Figure 1 shows an illustrative example of computing *sPS*. Let  $v_1$  a node subject to the knockout or the overexpression mutation. Through a

computation of the dynamics influence from  $v_1$  to  $v_2, v_3$  and  $v_4$ , we obtained the sets of affected genes regarding the knockout and the overexpression mutations,  $V_k = \{v_2, v_3, v_4\}$  and  $V_o = \{v_2, v_4\}$ , respectively. Thus,  $sPS(v_1) = 1/3$  which implies that the node  $v_1$  is pleiotropic because  $sPS(v_1) > 0$ .



**Figure 1.** An illustrative example of  $sPS$  computation. An example of a signaling network  $G(V, A)$  with a set of update rules  $F$ . Let  $v_1$  a node subject to the knockout or the overexpression mutations. The mutations change  $F$  to  $F'$  where the state value of  $v_1$  is frozen to 0 and 1, respectively, for  $t \leq T$ . The sets of genes whose dynamics are influenced by the mutations are identified as  $V_k$  and  $V_o$ , respectively. Accordingly,  $sPS$  of  $v_1$  is computed as the ratio of the symmetric difference of  $V_k$  and  $V_o$  over the union of them.

2.4. A Standardized Measure of Degree of Pleiotropy

A previous research used a standardized pleiotropic measure [2] to compute phenotypic effects in the baker’s yeast *S. cerevisiae*. Given a gene, they examined the average and the standard deviation of the number of transformed traits from wild-type cells, denoted by  $m_{wt}$  and  $SD$ , respectively, and the number of transformed traits from a cell deficient of the gene, denoted by  $m_d$ . In addition, they defined a standardized measure by the z-transformed pleiotropic score ( $zPS$ ) as follows:

$$zPS = \frac{m_d - m_{wt}}{SD} \tag{6}$$

We note that this notion of pleiotropy is different from that in  $sPS$ . The former focused on the standardized cardinality of the set of affected genes. To examine the notion of  $zPS$  in our work, we set  $m_d$  to the number of affected genes (i.e.,  $|V_k|$  or  $|V_o|$ ). In addition, we specified  $m_{wt}$  and  $SD$  as the average and the standard deviation of the number of associated phenotypes or traits of HPO database for every KEGG gene. In this way, we calculated  $zPS$  of the KEGG associated with HPO database and compared with our  $sPS$  of them.

2.5. Structural Characteristics of Pleiotropic Genes

It is known that the structural characteristics of genes in biological network related to their dynamics stability [59,60]. Here, we considered the following structural properties to investigate the relations to  $sPS$ .

- A feedback loop (FBL) means a sequence chain of nodes where any node is not repeated except the starting and the end nodes [59,61]. In a given network  $G(V, E)$ , an FBL is a closed simple cycle in which all nodes except the starting and ending nodes are not revisited; in other words, a path  $P$  of a length  $L(\geq 1)$  is represented by a sequence of ordered nodes  $u_1 \rightarrow u_2 \rightarrow \dots \rightarrow u_{L+1}$  with no repeated nodes except  $u_1$  and  $u_{L+1}$ . Hence, the  $P$  is called a feedback loop if  $u_1 = u_{L+1}$ . It was known that FBLs play important roles for controlling dynamics behavior of signaling networks [61–64].
- Centrality properties including the closeness defined as the reciprocal of the average distance from a node to every other node [4], the betweenness defined as the ability of a gene to control communication between genes through the shortest paths [65], the stress based on enumeration of the shortest paths [66], and the eigenvector represented by the principal eigenvector of the adjacency matrix of the network, where each node affects all of its neighbors [52].
- Degree of nodes represents the number of edge upon a gene link to another gene [4,64]. In addition, in-degree and out-degree mean the degree of the incoming and the outgoing links only, respectively.

### 2.6. Random Network Generation

To verify that the results of *sPS* in the real molecular interaction networks are relevant with randomly structured networks, we generated random networks using the Barabási Albert (BA) model [67], which is a kind of network growth model with a preferential attachment scheme, or called a probabilistic mechanism, where a new node is free to connect to any node in the network, whether it is a hub or just has a single link.

### 2.7. Parallel Computation

For efficient in silico simulations, we implemented the computational program using PANET [46], which is an analysis tool of the network dynamics using the OpenCL library (The recent version is available at <http://panet-csc.sourceforge.net/>, accessed on 26 November 2019). This allows us to compute large number of attractors in parallel by assigning each initial random state in Equations (3)–(6) to a processing unit of CPUs and/or GPUs.

### 2.8. Statistical Analysis

In this work, we conducted all statistical analysis using the Pearson correlation coefficient and the Student *t*-test using MedCalc Statistical Software version 13.0.6 (MedCalc Software bvba, Ostend, Belgium; <http://www.medcalc.org>; 2014) [68].

## 3. Results

In this work, we simulated *sPS* of all genes in the KEGG signaling networks using a Boolean network model (see Section 2). A total of 1000 initial states were randomly generated to calculate *sPS*, and the mutation duration time  $T$  was varied from 14 to 20 considering the network size of KEGG ( $|V| = 1659$ ).

### 3.1. Comparison of *sPS* with the Observational Pleiotropy

To show that our approach is relevant to the real phenotype data, we plot a contingency table between the degree of pleiotropy between our in silico model and HPO (Table 1). We first listed every KEGG gene that is associated with any phenotypes in the HPO database (see Supplementary Table S2). Next, we specified the degree of pleiotropy with respect to the HPO database as the number of phenotypes or traits of the gene in HPO ('HPO-associated'). Further, we specified the degree of pleiotropy with respect to our in silico model (the mutation duration time  $T$  was set to 20) as the number of KEGG affected genes by the knockout and the over-expression mutations in *sPS* ('*sPS*-associated').

**Table 1.** The list of genes compared between the degree of pleiotropy by HPO and *sPS*. The top 10 genes with the highest degree of pleiotropy are chosen by the number of phenotypes in HPO database ('HPO-associated') and the number of KEGG genes affected by the knockout and the over-expression mutations in *sPS* ('*sPS*-associated').

No.	Gene Name	HPO-Associated	<i>sPS</i> -Associated
1	<i>COL2A1</i>	842	22
2	<i>FGFR1</i>	1045	7
3	<i>FGFR2</i>	1137	3
4	<i>FGFR3</i>	1006	5
5	<i>LIMK1</i>	733	2
6	<i>NRAS</i>	728	1
7	<i>PIK3CA</i>	751	9
8	<i>PRKAR1A</i>	697	1
9	<i>PTEN</i>	838	107
10	<i>TGFBR2</i>	586	78
Total		8363	235

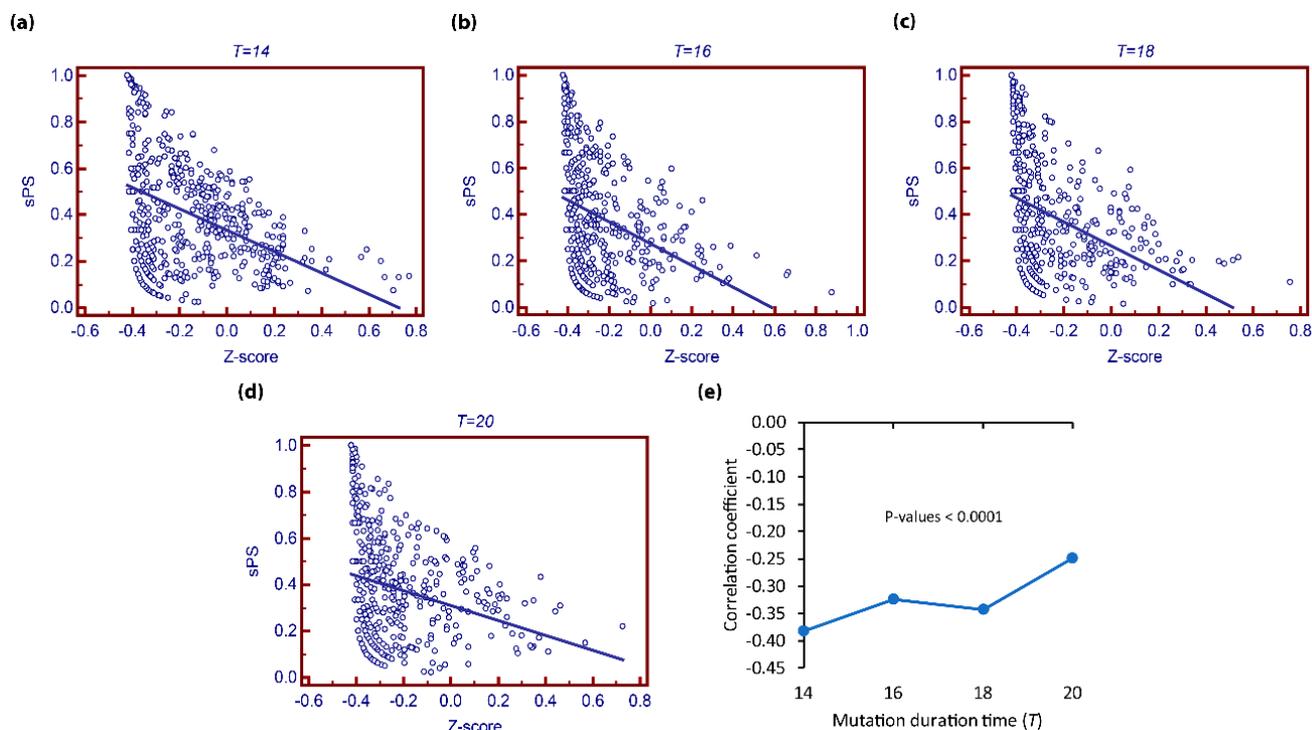
We then selected top 10 HPO-associated genes that are also related to *sPS*. As shown in the table, there are some genes that show a high degree of HPO-associated pleiotropy but a low degree of *sPS*-associated pleiotropy. This shows that our method '*sPS*-associated' supported the HPO-associated genes and can be used to analyze whose pleiotropic phenomenon is known in terms of network dynamics.

Next, we investigated the relationship between *sPS* and *zPS* by varying the mutation duration time (see Figure 2). As shown in Figure 2, there was a negative correlation between them irrespective of the mutation duration time  $T$  (All  $p$ -values  $< 0.0001$  using  $t$ -test). This implies that our measure is different from the previous measure. This is because *zPS* focused on the influential extent of the mutation, whereas *sPS* considers the degree of influential difference caused by a pair of differences. In this regard, *sPS* can convey a novel viewpoint of pleiotropy from the previous approach.

### 3.2. Relation of *sPS* and the Functional Importance Genes

Some pleiotropic genes are relevant to many functionally important genes such as cancer genes, drug targets, essential genes, tumor suppressors, oncogenes, and disease genes. For example, it is known that cancer is one of the lead causes of death in human population [69,70], which comes from the accumulation of sequential mutations resulting from cell abnormalities or genetic instability [71,72]. Accordingly, it is not surprising that pleiotropic analysis has become very common in explore different cancer phenotypes [73]. Another examples is the investigation of drug-target genes through network-based analysis [74,75], which shows a significantly different connectivity, more feedback loops, and more evolutionary than non-drug target genes [76]. Thus, drug targets were considered potential cancer therapeutics [77], and recently, they have been recognized as new therapeutic targets for pleiotropic genes [78]. Furthermore, it has been identified that the deletion of any essential genes can lead to death or infertility [79] and tend to be associated with human disease genes [5,80]. Inspired by those results, we investigated the relationship of *sPS* with functionally important genes. Firstly, we specified every KEGG gene into 'cancer', 'drug-target (DT)', 'essential', 'tumor-suppressors (TSG)', 'oncogene (OCG)', and 'disease gene (DG)' (see Supplementary Table S3 in detail). Secondly, we examined their *sPS* values and classified them into two subset groups, 'non-Zero *sPS*' and 'Zero-*sPS*'. We compared the proportion of the functionally important genes between the two groups. Figure 3a–f shows the result of cancer genes, drug targets, essential genes, tumor suppressors, oncogenes, and disease genes, respectively. As shown in the figure, the ratio in the non-zero *sPS* group is significantly larger than that in the zero-*sPS* group for all types of functionally important genes and all mutation duration time  $T$ . In other words, the functionally important genes tend to be more pleiotropic than the other genes based on our in silico model. It is interesting that this result is relevant with some previous studies. For example, the gene *TERT* was found to be a pleiotropic cancer gene associated with

12 different cancer types [81], while *PTPN2* was confirmed as a pleiotropic gene associated with several autoimmune diseases [82]. This result indicates the promising usefulness of *sPS* to predict the unknown pleiotropic role of functionally important genes.



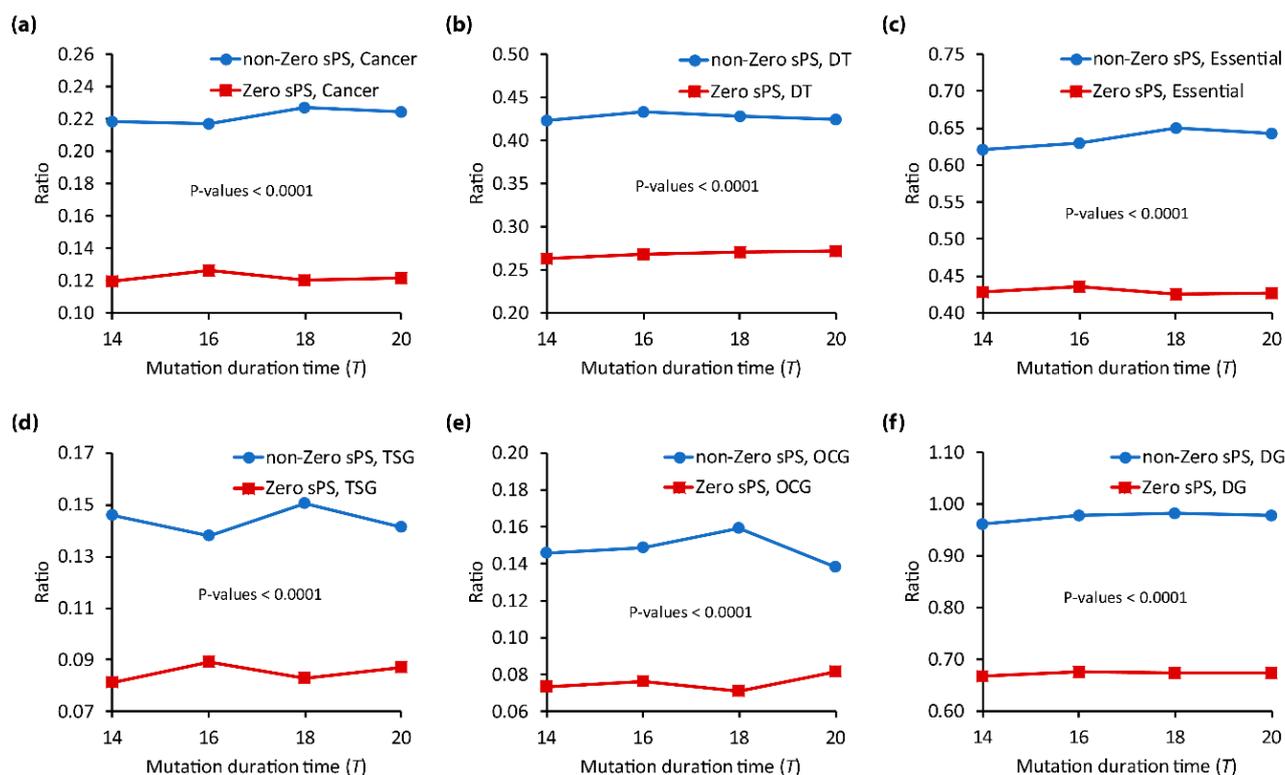
**Figure 2.** Relationship between *sPS* and *zPS* in KEGG network. (a–d) Relations of *sPS* and *zPS* in scatter-plot graph for mutation duration time  $T = 14 - 20$ , respectively. (e) Correlation coefficients of *sPS* and *zPS*. Only genes with positive *sPS* values were examined. All *p*-values are significant (*p*-value < 0.0001).

### 3.3. Relation of *sPS* and the Structural Characteristics

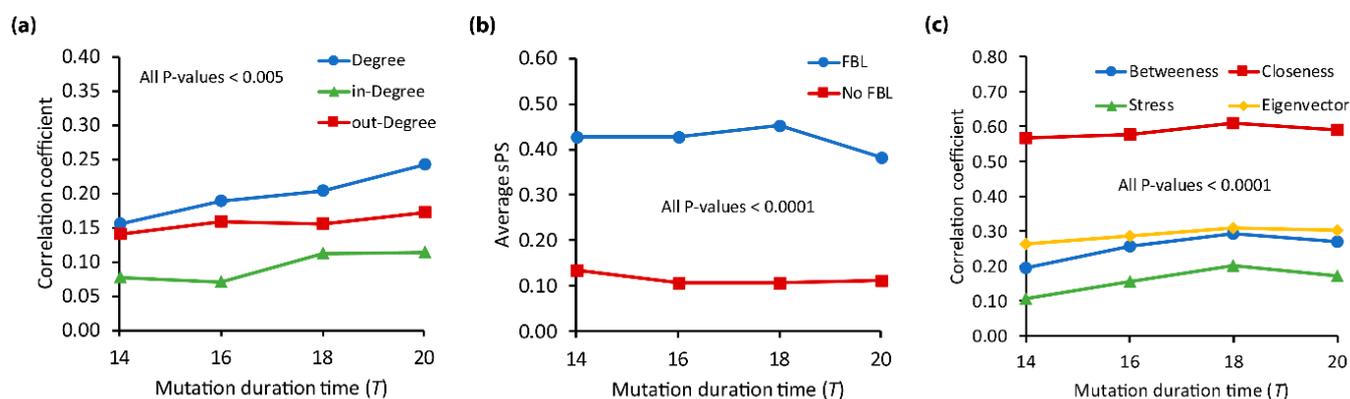
Some previous studies have shown that the structural characteristics of a gene in signaling networks are related to its dynamical behavior [59,60]. In this regard, we examined the relation of *sPS* with the structural characteristics, specifically, the degree, the involvement of feedback loops, and some centrality measures (Figure 4). For every KEGG genes, we first compared the correlation coefficients between degree/in-degree/out-degree of genes with *sPS* values.

As shown in Figure 4a, all degrees showed significantly positive correlations, irrespective of the mutation duration time  $T$ . It means that the *sPS* values tend to be larger as the number of degree/in-degree/out-degree gets larger. In addition, we observed that the correlation coefficient of *sPS* with in-degree was relatively lower than those with degree and out-degree. In fact, the difference of the average *sPS* between two gene groups classified by in-degree values was not as large as that when two gene groups were classified by degree or out-degree values (see Figure S2 in Supplementary Information). It is interesting that degree shows more apparent relation than either specific sub degree type. Next, we examined the relation of *sPS* with the involvement of feedback loops. Thus, we classified each gene into 'FBL' and 'No FBL' groups if the gene was involved with any feedback loops or not, respectively, and then compared the average *sPS* values of the groups. As shown in Figure 4b, the average *sPS* value of 'FBL' group is significantly larger than that of 'No FBL' group. This implies that a gene tends to be more pleiotropic when it is involved with feedback loops. Moreover, we computed the correlation coefficient between the number of feedback loops and *sPS* values and found significant positive relations (see Figure S3 in Supplementary Information; All *p*-values < 0.0001 using *t*-test). This implies that feedback

loops can play an important role in pleiotropy, as indicated in a previous study [83]. This result is intriguing because many previous studies have shown the relation of feedback loops with various dynamical behavior of biological networks [62,84]. For example, the FBL plays role in amplifying (positive feedback loop) or inhibiting (negative feedback loop) of the intracellular signals [62,84], related to disease comorbidity [85], or protein–protein interaction [64]. Thus, this result can add the importance of feedback loop structure in terms of the network dynamics. Finally, we examined the relations of centrality measures such as closeness, betweenness, stress, and eigenvector with *sPS* values and found that all of them have positive correlations (all *p*-values < 0.0001 using *t*-test), irrespective of the mutation duration time (Figure 4c). In other words, it is likely that the more central gene in signaling networks shows a higher *sPS* value. In particular, the correlation coefficient of closeness, which indicates how closely a gene is located to other genes in a network, was the largest. On the other hand, the correlation coefficient of stress was the lowest. It is interesting that our centrality result is consistent with some previous results showing that pleiotropic genes were more central in protein interaction networks [4,17]. In addition, we examined the correlation coefficient of *sPS* values with degree/in-degree/out-degree of nodes, feedback loops, and centrality measures in the BA random networks and could observe consistent results (see Supplementary Figure S4). This implies that our results are principles not only in real networks but also in artificially structured networks.



**Figure 3.** Relation of *sPS* with the functionally important genes in KEGG network. (a–f) Results of cancer genes, drug targets, essential genes, tumor suppressors, oncogenes, and disease genes, respectively. In each subfigure, all genes were classified into ‘non-zero *sPS*’ and ‘zero-*sPS*’ groups. Y-axis is the ratio of the functionally important genes among the total genes in the group. The mutation duration time *T* was varied from 14 to 20.



**Figure 4.** Relation of *sPS* with structural characteristics in KEGG network. (a) Relation to the degree. Y-axis values mean the correlation coefficients between *sPS* and the number of nodes' degree, in-degree, and out-degree. (b) Relation to the involvement of feedback loops. All genes were classified into 'FBL' and 'No FBL' groups, where a gene involves any feedback loops or not, respectively. Y-axis values mean the average of *sPS* values. (c) Relations to the centrality measures such as betweenness, stress, closeness, and eigenvector. Y-axis values mean the correlation coefficients between *sPS* and each centrality measure. Mutation duration time *T* was set to 14–20.

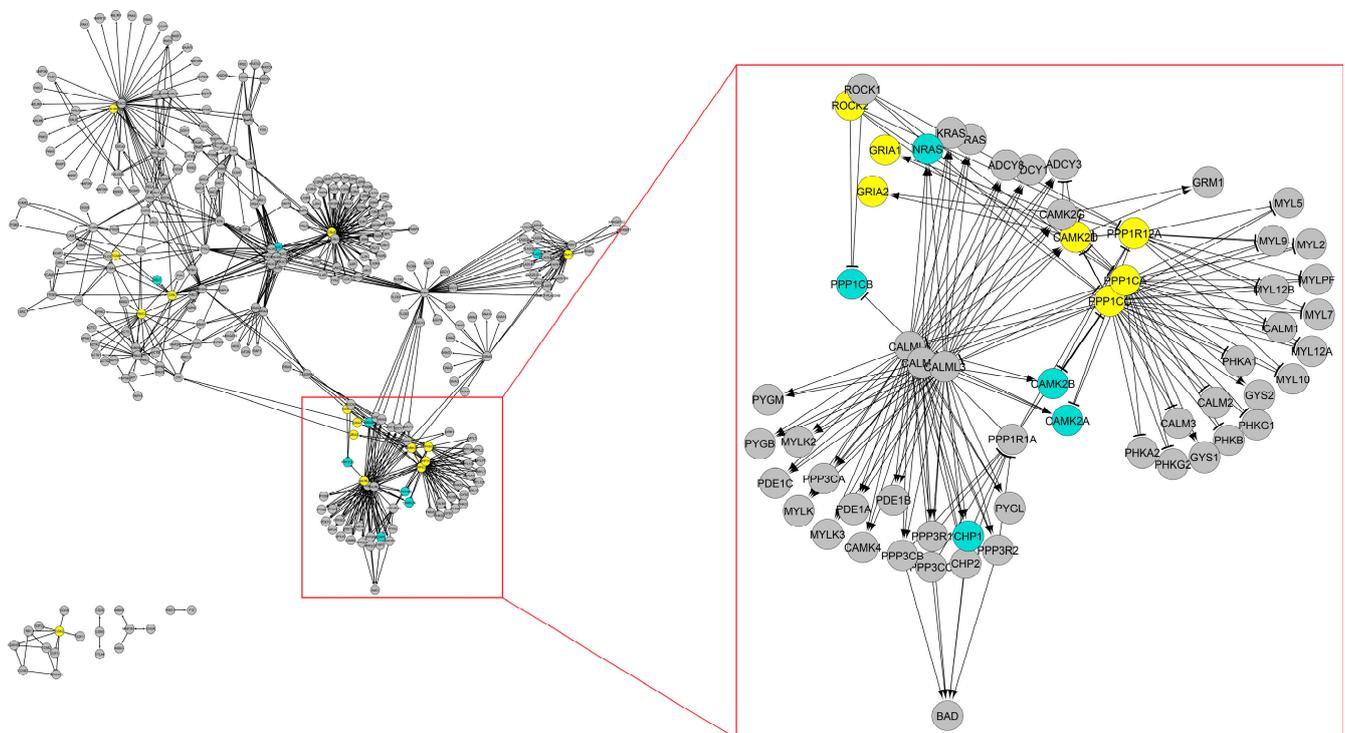
### 3.4. Biological Evidence of Pleiotropic Genes Based on *sPS*

To reveal novel candidate pleiotropic dynamics by the *sPS* measure, we profiled the genes with high *sPS* values from the KEGG network. A total of 29 genes are shown in Table 2. Among them, ten genes were known to be the real pleiotropic genes, for example, the *PIK3CA* (*PIK3*) gene, which is found to be most pleiotropic among targets of drugs abuse in pharmacological experiments [86], and the *ABL1* gene, which is expressed in all tissues of mice pleiotropic phenotype T-cell signaling [87]. In addition, it was reported that *EGFR* gene is a marker of pleiotropic effects in underlying kidney function and cardiovascular disease [88]. It was clear that those genes were involved with many cancer types, associated with drug targets, essentials, and disease genes. On the other hand, we found 19 novel candidate genes in Table 2 that are not known in the experimental pleiotropic database. This suggests that *sPS* can predict the unknown pleiotropic genes. As shown in the table, those genes are very interesting because most of them are associated with various functionally important genes such as cancer, drug target, oncogenes, tumor suppressors, essentials, or disease genes. Specifically, all such genes were associated with disease genes. This implies that disease genes tend to be pleiotropic. Further, we map the listed genes in Table 1 into KEGG sub-network (Figure 5; see Figure S5 in the Supplementary Materials for original network) for visualization. We note that only the listed genes and their neighbors were included in the sub-network. In the figure, the known pleiotropic genes and the novel candidate pleiotropic genes were marked by a blue and a yellow circle, respectively. Interestingly, the novel candidate pleiotropic genes were located closely to the known pleiotropic genes (most of the yellow circle was located from one of blue circle with length 1 in the network). This implies that the novel candidate pleiotropic genes tend to closely interact with the known pleiotropic genes in the signaling network. This was relevant to a study that found that larger effects of pleiotropy can also be caused by correlated effects among traits [89] or the regulatory networks that are so highly interconnected influence neighbor genes to have effects on the core disease genes [90]. In addition, most novel candidate pleiotropic genes were involved in feedback loops in the network. Moreover, their in-degree tends to be smaller than their out-degree. Hence, we can conclude that novel candidate pleiotropic genes are not only associated with certain functional influence genes but also have certain structural characteristics. Taken together, our *sPS* measure can be more efficiently used to reveal novel candidate pleiotropic genes when it is combined with structural characteristics index.

**Table 2.** Significant pleiotropic genes in KEGG network.

No.	Gene Name	HPO	sPS	zPS	NuCancer	DT	ES	TSG	OCG	DG	Deg	In-Deg	Out-Deg	NuFBL
1	ABL1	1	0.6	-0.35	23	1	1	0	1	1	27	15	12	4588
2	PIK3CA	1	0.7	-0.06	19	1	1	0	1	1	51	45	6	40,101
3	EGFR	1	0.45	0.11	21	1	1	0	1	1	73	41	32	225
4	SERPINA1	1	0.21	-0.34	60	1	0	0	0	1	1	0	1	0
5	CAMK2B	1	0.44	-0.04612	0	1	0	0	0	1	15	10	5	19,612
6	PPP1CB	1	0.55	-0.06861	1	0	1	0	0	1	28	3	25	30,526
7	CAMK2A	1	0.63	-0.30841	0	1	1	0	0	1	15	10	5	19,612
8	NRAS	1	0.78	-0.40583	1	0	1	0	1	1	44	28	16	20,154
9	CHP1	1	0.45	-0.08359	0	1	1	0	0	1	10	8	2	2970
10	PLA2G6	1	1	-0.42082	303	1	1	0	0	1	10	10	0	0
11	IGFBP3	0	0.45	0.08	1	1	1	1	0	1	1	0	1	0
12	PRKCA	0	0.48	0.06	0	1	1	0	1	1	24	7	17	10,054
13	ITGAM	0	0.86	-0.4	0	0	0	0	0	1	9	3	6	260
14	ROCK2	0	0.59	-0.12	0	1	1	0	0	1	7	3	4	39,206
15	PPP1CC	0	0.42	0.09	0	1	1	0	0	1	28	3	28	30,526
16	PPP1CA	0	0.52	-0.02	0	1	1	1	0	1	28	3	25	30,526
17	PRKAA1	0	0.40	-0.2035	0	1	1	1	0	1	3	0	3	0
18	PPP1R12A	0	0.73	-0.27844	0	0	1	0	0	1	15	3	12	7624
19	CDK2	0	0.61	-0.36836	0	1	1	1	0	1	10	3	7	4
20	PPP1CC	0	0.451327	0.043803	0	1	1	0	0	1	28	3	25	30,526
21	RALBP1	0	0.428571	-0.00116	0	0	0	0	0	1	6	2	4	648
22	CBLB	0	0.6045	-0.02364	1	0	1	0	1	1	60	4	56	1202
23	WNT11	0	0.5455	-0.34588	0	0	1	1	0	1	17	7	10	0
24	CAMK2D	0	0.75	-0.27844	0	1	0	0	0	1	15	10	5	19,612
25	CRK	0	1	-0.42082	0	0	1	0	1	1	45	31	14	5391
26	CALML5	0	0.4455	0.036309	0	0	0	0	0	1	36	9	27	11,451
27	GRIA1	0	0.90625	-0.39834	0	1	1	0	0	1	10	9	1	13,272
28	GRIA2	0	0.619048	-0.18102	0	1	1	0	0	1	10	9	1	13,272
29	GNAI2	0	0.317647	0.013827	1	0	1	0	1	1	23	6	17	2528

HPO = 1 means the gene was confirmed in the real observational pleiotropic database, otherwise 0. DT = 1, ES = 1, TSG = 1, OCG = 1, DG = 1 means the gene involves with drug-target, essential, tumor-suppressors, oncogenes, or disease genes, respectively, otherwise 0. NuCancer abbreviates the number of associated cancer types. Deg/In-Deg/Out-Deg denote the values of node degree/in-degree/out-degree. NuFBL abbreviates the number of feedback loops involving gene.



**Figure 5.** Pleiotropic genes in KEGG sub-network. Arrow-headed and bar-headed lines indicate the activation (positive) and the inhibition (negative) interactions, respectively. The grey circles belong to the non-observed genes. The blue circles represent confirmed pleiotropic genes from the HPO database. The yellow circles represent the novel candidate pleiotropic genes.

#### 4. Discussion

In this study, we defined pleiotropic scores ( $sPS$ ), which represents how much the dynamics of a gene is affected against a couple of different mutations, herein the knockout and the overexpression mutations, using a Boolean network model. We investigated our approach using the KEGG signaling network. It was interesting to observe that the affected genes with high  $sPS$  values were related to the real pleiotropic data and can explain the dynamical behavior of pleiotropic genes. More interestingly, the various functionally important genes were related to pleiotropic genes. Further, we also found that novel candidate pleiotropic genes tend to closely interact with the known pleiotropic genes in the signaling network. These results will enhance the understanding of dynamical effects on pleiotropic genes, especially in large-scale biological systems. Despite the usefulness of our approach, there are some limitations caused by the Boolean network model used in this study. The first concern is the use of the nested canalizing function to represent a update of a gene status. However, some previous studies have proven its usefulness in gene regulatory interactions. For example, 133 out of 139 rules compiled from a dataset about a transcriptional regulatory network [50], or 39 out of 42 rules inferred from a dataset about signaling pathways [91] can be classified into the nested canalizing functions. Another concern is the synchronous update scheme, which is less realistic than the asynchronous update scheme. In fact, it is likely that the genes in the real signaling networks are regulated in an asynchronous update rule. However, it is required to properly specify some unknown parameters to implement the asynchronous scheme. For example, the asynchronous update assumes that only one node can change state at any given moment, and each node has the same probability of being updated [26]. This implies that the asynchronous update is valid only when a correct strategy to choose an update sequence is known. In this regard, a future study will include an approach to infer the update rule from real biological data instead of generating random update rules. In addition, it will extend to a more generalized analysis considering various mutation types.

#### 5. Conclusions

Pleiotropy refers to the ability of different mutations within the same gene to cause different pathological effects, and many computational methods have been suggested to unravel the dynamics of the pleiotropy. However, little is known in identifying more complicated dynamical relations of gene pleiotropy, since most of them focused on undirected molecular interaction networks. Therefore, a new perspective is needed to investigate the dynamical characteristics induced by gene–gene pleiotropic. In this work, we proposed a measure to compute gene–gene *in silico* pleiotropic scores ( $sPS$ ) representing how much the gene is affected against the different type of mutations on dynamics using a Boolean network model. We considered knockout and overexpression mutations to compute  $sPS$  values. Through intensive investigations, we found that some functional importance genes such as cancer, drug-target genes, tumor suppressors, oncogenes, essential genes, and disease genes tend to have non-zero  $sPS$  values than other genes. Next, we investigated the relationships of  $sPS$  and structural properties and found that there are positive correlations with the number of nodes' degree/in-degree/out-degree, feedback loops, and centrality measures such as closeness, betweenness, stress, and eigenvector. More interestingly, we were able to find some biological evidence confirming that  $sPS$  is relevant to real pleiotropic data and can be used to find novel candidate pleiotropic gene characteristics. Overall, our results suggest the usefulness of  $sPS$  in understanding the dynamics pleiotropic in complex biological systems.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom12081139/s1>, Supporting Text; Figure S1. Dynamics influence computation example; Figure S2. Relations of *sPS* and degree of nodes in KEGG network; Figure S3. Relations of *sPS* values to feedback loops in KEGG network; Figure S4. Relations of *sPS* value in random BA network; Figure S5. Original KEGG network with  $|V| = 1659$ ,  $|A| = 7964$ ; Table S1. KEGG network dataset consists of 1659 nodes and 7964 interactions; Table S2. KEGG-HPO associated [92]; Table S3. KEGG functional importance genes; Table S4. KEGG functional importance genes (continuous).

**Author Contributions:** Conceptualization, M.M. and Y.-K.K.; methodology, M.M. and Y.-K.K.; validation, M.M. and Y.-K.K.; formal analysis, M.M.; investigation, M.M.; resources, M.M. and Y.-K.K.; data curation, M.M.; writing—original draft preparation, M.M.; writing—review and editing, M.M. and Y.-K.K.; visualization, M.M.; supervision, Y.-K.K.; project administration, Y.-K.K.; funding acquisition, Y.-K.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the 2022 Research Fund of University of Ulsan.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article or Supplementary Material.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ittisoponpisan, S.; Alhuzimi, E.; Sternberg, M.; David, A. Landscape of Pleiotropic Proteins Causing Human Disease: Structural and System Biology Insights. *Hum. Mutat.* **2016**, *38*, 289–296. [[CrossRef](#)] [[PubMed](#)]
2. Wang, Z.; Liao, B.-Y.; Zhang, J. Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18034–18039. [[CrossRef](#)]
3. Dudley, A.M.; Janse, D.M.; Tanay, A.; Shamir, R.; Church, G.M. A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol. Syst. Biol.* **2005**, *1*, 2005.0001. [[CrossRef](#)] [[PubMed](#)]
4. Chavali, S.; Barrenas, F.; Kanduri, K.; Benson, M. Network properties of human disease genes with pleiotropic effects. *BMC Syst. Biol.* **2010**, *4*, 78. [[CrossRef](#)]
5. Sivakumaran, S.; Agakov, F.; Theodoratou, E.; Prendergast, J.G.; Zgaga, L.; Manolio, T.; Rudan, I.; McKeigue, P.; Wilson, J.F.; Campbell, H. Abundant Pleiotropy in Human Complex Diseases and Traits. *Am. J. Hum. Genet.* **2011**, *89*, 607–618. [[CrossRef](#)]
6. Liu, Y.; Koyutürk, M.; Barnholtz-Sloan, J.S.; Chance, M.R. Gene interaction enrichment and network analysis to identify dysregulated pathways and their interactions in complex diseases. *BMC Syst. Biol.* **2012**, *6*, 65. [[CrossRef](#)]
7. Polster, R.; Petropoulos, C.J.; Bonhoeffer, S.; Guillaume, F. Epistasis and Pleiotropy Affect the Modularity of the Genotype–Phenotype Map of Cross-Resistance in HIV-1. *Mol. Biol. Evol.* **2016**, *33*, 3213–3225. [[CrossRef](#)]
8. Mi, Z.; Guo, B.; Yin, Z.; Li, J.; Zheng, Z. Disease classification via gene network integrating modules and pathways. *R. Soc. Open Sci.* **2019**, *6*, 190214. [[CrossRef](#)]
9. Yan, J.; Risacher, S.L.; Shen, L.; Saykin, A.J. Network approaches to systems biology analysis of complex disease: Integrative methods for multi-omics data. *Briefings Bioinform.* **2017**, *19*, 1370–1381. [[CrossRef](#)]
10. Duveau, F.; Félix, M.-A. Role of Pleiotropy in the Evolution of a Cryptic Developmental Variation in *Caenorhabditis elegans*. *PLoS Biol.* **2012**, *10*, e1001230. [[CrossRef](#)]
11. Seoane, J.A.; Campbell, C.; Day, I.N.M.; Casas, J.P.; Gaunt, T.R. Canonical Correlation Analysis for Gene-Based Pleiotropy Discovery. *PLoS Comput. Biol.* **2014**, *10*, e1003876. [[CrossRef](#)]
12. Chung, J.; Alzheimer’s Disease Genetics Consortium; Zhang, X.; Allen, M.; Wang, X.; Ma, Y.; Beecham, G.; Montine, T.J.; Younkin, S.G.; Dickson, D.W.; et al. Genome-wide pleiotropy analysis of neuropathological traits related to Alzheimer’s disease. *Alzheimer’s Res. Ther.* **2018**, *10*, 22. [[CrossRef](#)]
13. Collet, J.M.; McGuigan, K.; Allen, S.; Chenoweth, S.F.; Blows, M.W. Mutational Pleiotropy and the Strength of Stabilizing Selection within and between Functional Modules of Gene Expression. *Genetics* **2018**, *208*, 1601–1616. [[CrossRef](#)]
14. Alzoubi, D.; Desouki, A.A.; Lercher, M.J. Alleles of a gene differ in pleiotropy, often mediated through currency metabolite production, in *E. coli* and yeast metabolic simulations. *Sci. Rep.* **2018**, *8*, 17252. [[CrossRef](#)]
15. A Stoney, R.; Ames, R.M.; Nenadic, G.; Robertson, D.L.; Schwartz, J.-M. Disentangling the multigenic and pleiotropic nature of molecular function. *BMC Syst. Biol.* **2015**, *9*, S3. [[CrossRef](#)]
16. Yazdani, A.; Yazdani, A.; Elsea, S.H.; Schaid, D.J.; Kosorok, M.R.; Dangol, G.; Samiei, A. Genome analysis and pleiotropy assessment using causal networks with loss of function mutation and metabolomics. *BMC Genom.* **2019**, *20*, 395. [[CrossRef](#)]
17. Nguyen, T.-P.; Liu, W.-C.; Jordán, F. Inferring pleiotropy by network analysis: Linked diseases in the human PPI network. *BMC Syst. Biol.* **2011**, *5*, 179. [[CrossRef](#)]

18. Li, S.; Assmann, S.M.; Albert, R. Predicting Essential Components of Signal Transduction Networks: A Dynamic Model of Guard Cell Abscisic Acid Signaling. *PLoS Biol.* **2006**, *4*, e312. [[CrossRef](#)]
19. Kwon, Y.-K.; Kim, J.; Cho, K.-H. Dynamical Robustness against Multiple Mutations in Signaling Networks. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2015**, *13*, 996–1002. [[CrossRef](#)]
20. Shastry, B. Overexpression of genes in health and sickness. A bird's eye view. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **1995**, *112*, 1–13. [[CrossRef](#)]
21. Kwon, Y.-K.; Cho, K.-H. Analysis of feedback loops and robustness in network evolution based on Boolean models. *BMC Bioinform.* **2007**, *8*, 430. [[CrossRef](#)]
22. Raeymaekers, L. Dynamics of Boolean Networks Controlled by Biologically Meaningful Functions. *J. Theor. Biol.* **2002**, *218*, 331–341. [[CrossRef](#)]
23. Schwab, J.D.; Kühlwein, S.D.; Ikonomi, N.; Kühl, M.; Kestler, H.A. Concepts in Boolean network modeling: What do they all mean? *Comput. Struct. Biotechnol. J.* **2020**, *18*, 571–582. [[CrossRef](#)]
24. Wang, R.-S.; Saadatpour, A.; Albert, R. Boolean modeling in systems biology: An overview of methodology and applications. *Phys. Biol.* **2012**, *9*, 055001. [[CrossRef](#)]
25. Gupta, S.; Hashimoto, R.F. Dynamical Analysis of a Boolean Network Model of the Oncogene Role of lncRNA ANRIL and lncRNA UFC1 in Non-Small Cell Lung Cancer. *Biomolecules* **2022**, *12*, 420. [[CrossRef](#)]
26. Wooten, D.J.; Zañudo, J.G.T.; Murrugarra, D.; Perry, A.M.; Dongari-Bagtzoglou, A.; Laubenbacher, R.; Nobile, C.J.; Albert, R. Mathematical modeling of the *Candida albicans* yeast to hyphal transition reveals novel control strategies. *PLoS Comput. Biol.* **2021**, *17*, e1008690. [[CrossRef](#)]
27. Guberman, E.; Sherief, H.; Regan, E.R. Boolean model of anchorage dependence and contact inhibition points to coordinated inhibition but semi-independent induction of proliferation and migration. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 2145–2165. [[CrossRef](#)]
28. Gupta, S.; Panda, P.K.; Hashimoto, R.F.; Samal, S.K.; Mishra, S.; Verma, S.K.; Mishra, Y.K.; Ahuja, R. Dynamical modeling of miR-34a, miR-449a, and miR-16 reveals numerous DDR signaling pathways regulating senescence, autophagy, and apoptosis in HeLa cells. *Sci. Rep.* **2022**, *12*, 4911. [[CrossRef](#)]
29. Mazaya, M.; Trinh, H.-C.; Kwon, Y.-K. Construction and analysis of gene-gene dynamics influence networks based on a Boolean model. *BMC Syst. Biol.* **2017**, *11*, 133. [[CrossRef](#)]
30. Robinson, P.N.; Köhler, S.; Bauer, S.; Seelow, D.; Horn, D.; Mundlos, S. The Human Phenotype Ontology: A Tool for Annotating and Analyzing Human Hereditary Disease. *Am. J. Hum. Genet.* **2008**, *83*, 610–615. [[CrossRef](#)]
31. Kim, J.-R.; Kim, J.; Kwon, Y.-K.; Lee, H.-Y.; Heslop-Harrison, P.; Cho, K.-H. Reduction of Complex Signaling Networks to a Representative Kernel. *Sci. Signal.* **2011**, *4*, ra35. [[CrossRef](#)] [[PubMed](#)]
32. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2015**, *44*, D457–D462. [[CrossRef](#)] [[PubMed](#)]
33. Groza, T.; Köhler, S.; Moldenhauer, D.; Vasilevsky, N.; Baynam, G.; Zemojtel, T.; Schriml, L.M.; Kibbe, W.A.; Schofield, P.N.; Beck, T.; et al. The Human Phenotype Ontology: Semantic Unification of Common and Rare Disease. *Am. J. Hum. Genet.* **2015**, *97*, 111–124. [[CrossRef](#)] [[PubMed](#)]
34. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2012**, *2*, 401–404. [[CrossRef](#)]
35. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.E.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **2013**, *6*, p11. [[CrossRef](#)]
36. Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.; et al. DrugBank 3.0: A comprehensive resource for 'Omics' research on drugs. *Nucleic Acids Res.* **2010**, *39*, D1035–D1041. [[CrossRef](#)]
37. Zhang, R.; Lin, Y. DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. *Nucleic Acids Res.* **2008**, *37*, D455–D458. [[CrossRef](#)]
38. Zhao, M.; Kim, P.; Mitra, R.; Zhao, J.; Zhao, Z. TSGene 2.0: An updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res.* **2015**, *44*, D1023–D1031. [[CrossRef](#)]
39. Zhao, M.; Sun, J.; Zhao, Z. TSGene: A web resource for tumor suppressor genes. *Nucleic Acids Res.* **2012**, *41*, D970–D976. [[CrossRef](#)]
40. Liu, Y.; Sun, J.; Zhao, M. ONGene: A literature-based database for human oncogenes. *J. Genet. Genom.* **2017**, *44*, 119–121. [[CrossRef](#)]
41. Piñero, J.; Ramírez-Anguaita, J.M.; Saüch-Pitarch, J.; Ronzano, F.; Centeno, E.; Sanz, F.; I Furlong, L. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* **2019**, *48*, D845–D855. [[CrossRef](#)] [[PubMed](#)]
42. Piñero, J.; Bravo, À.; Queralt-Rosinach, N.; Gutiérrez-Sacristán, A.; Deu-Pons, J.; Centeno, E.; García-García, J.; Sanz, F.; Furlong, L.I. DisGeNET: A comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.* **2016**, *45*, D833–D839. [[CrossRef](#)] [[PubMed](#)]
43. Geiler-Samerotte, K.A.; Li, S.; Lazaris, C.; Taylor, A.; Ziv, N.; Ramjeawan, C.; Paaby, A.B.; Siegal, M.L. Extent and context dependence of pleiotropy revealed by high-throughput single-cell phenotyping. *PLoS Biol.* **2020**, *18*, e3000836. [[CrossRef](#)] [[PubMed](#)]

44. Christodoulaki, E.; Nolte, V.; Lai, W.-Y.; Schlötterer, C. Natural variation in *Drosophila* shows weak pleiotropic effects. *Genome Biol.* **2022**, *23*, 116. [[CrossRef](#)]
45. Campbell, C.; Albert, R. Stabilization of perturbed Boolean network attractors through compensatory interactions. *BMC Syst. Biol.* **2014**, *8*, 53. [[CrossRef](#)]
46. Trinh, H.-C.; Le, D.-H.; Kwon, Y.-K. PANET: A GPU-Based Tool for Fast Parallel Analysis of Robustness Dynamics and Feed-Forward/Feedback Loop Structures in Large-Scale Biological Networks. *PLoS ONE* **2014**, *9*, e103010. [[CrossRef](#)]
47. Mendes, N.D.; Lang, F.; Le Cornec, Y.-S.; Mateescu, R.; Batt, G.; Chaouiya, C. Composition and abstraction of logical regulatory modules: Application to multicellular systems. *Bioinformatics* **2013**, *29*, 749–757. [[CrossRef](#)]
48. Kauffman, S.; Peterson, C.; Samuelsson, B.; Troein, C. Genetic networks with canalizing Boolean rules are always stable. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17102–17107. [[CrossRef](#)]
49. Kauffman, S.; Peterson, C.; Samuelsson, B.; Troein, C. Random Boolean network models and the yeast transcriptional network. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 14796–14799. [[CrossRef](#)]
50. Harris, S.E.; Sawhill, B.K.; Wuensche, A.; Kauffman, S. A model of transcriptional regulatory networks based on biases in the observed regulation rules. *Complexity* **2002**, *7*, 23–40. [[CrossRef](#)]
51. Samal, A.; Jain, S. The regulatory network of *E. coli* metabolism as a Boolean dynamical system exhibits both homeostasis and flexibility of response. *BMC Syst. Biol.* **2008**, *2*, 21. [[CrossRef](#)] [[PubMed](#)]
52. Trinh, H.-C.; Kwon, Y.-K. Effective Boolean dynamics analysis to identify functionally important genes in large-scale signaling networks. *Biosystems* **2015**, *137*, 64–72. [[CrossRef](#)]
53. Davidich, M.I.; Bornholdt, S. Boolean Network Model Predicts Knockout Mutant Phenotypes of Fission Yeast. *PLoS ONE* **2013**, *8*, e71786. [[CrossRef](#)]
54. Shao, J.; Fujiwara, T.; Kadowaki, Y.; Fukazawa, T.; Waku, T.; Itoshima, T.; Yamatsuji, T.; Nishizaki, M.; A Roth, J.; Tanaka, N. Overexpression of the wild-type p53 gene inhibits NF- $\kappa$ B activity and synergizes with aspirin to induce apoptosis in human colon cancer cells. *Oncogene* **2000**, *19*, 726–736. [[CrossRef](#)] [[PubMed](#)]
55. Dudgeon, C.; Chan, C.; Kang, W.; Sun, Y.; Emerson, R.; Robins, H.; Levine, A.J. The evolution of thymic lymphomas in p53 knockout mice. *Genes Dev.* **2014**, *28*, 2613–2620. [[CrossRef](#)]
56. Prelich, G. Gene Overexpression: Uses, Mechanisms, and Interpretation. *Genetics* **2012**, *190*, 841–854. [[CrossRef](#)]
57. Taylor, M.A.; Wilczek, A.M.; Roe, J.L.; Welch, S.M.; Runcie, D.E.; Cooper, M.D.; Schmitt, J. Large-effect flowering time mutations reveal conditionally adaptive paths through fitness landscapes in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 17890–17899. [[CrossRef](#)]
58. Bozic, I.; Nowak, M.A. Timing and heterogeneity of mutations associated with drug resistance in metastatic cancers. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15964–15968. [[CrossRef](#)]
59. Prill, R.J.; Iglesias, P.; Levchenko, A. Dynamic Properties of Network Motifs Contribute to Biological Network Organization. *PLoS Biol.* **2005**, *3*, e343. [[CrossRef](#)]
60. Klein, C.; Marino, A.; Sagot, M.-F.; Milreu, P.V.; Brilli, M. Structural and dynamical analysis of biological networks. *Briefings Funct. Genom.* **2012**, *11*, 420–433. [[CrossRef](#)]
61. Kwon, Y.-K. Properties of Boolean dynamics by node classification using feedback loops in a network. *BMC Syst. Biol.* **2016**, *10*, 83. [[CrossRef](#)] [[PubMed](#)]
62. Thomas, R.; Thieffry, D.; Kaufman, M. Dynamical behaviour of biological regulatory networks—I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull. Math. Biol.* **1995**, *57*, 247–276. [[CrossRef](#)]
63. Hetmanski, J.H.R.; Zindy, E.; Schwartz, J.M.; Caswell, P.T. A MAPK-Driven Feedback Loop Suppresses Rac Activity to Promote RhoA-Driven Cancer Cell Invasion. *PLoS Comput. Biol.* **2016**, *12*, e1004909. [[CrossRef](#)] [[PubMed](#)]
64. Yeger-Lotem, E.; Sattath, S.; Kashtan, N.; Itzkovitz, S.; Milo, R.; Pinter, R.Y.; Alon, U.; Margalit, H. Network motifs in integrated cellular networks of transcription–regulation and protein–protein interaction. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5934–5939. [[CrossRef](#)]
65. Freeman, L.C. A Set of Measures of Centrality Based on Betweenness. *Sociometry* **1977**, *40*, 35–41 [[CrossRef](#)]
66. Shimmel, A. Structural parameters of communication networks. *Bull. Math. Biol.* **1953**, *15*, 501–507. [[CrossRef](#)]
67. Barabási, A.-L.; Albert, R. Emergence of Scaling in Random Networks. *Science* **1999**, *286*, 509–512. [[CrossRef](#)]
68. Schoonjans, F.; Zalata, A.; Depuydt, C.; Comhaire, F. MedCalc: A new computer program for medical statistics. *Comput. Methods Programs Biomed.* **1995**, *48*, 257–262. [[CrossRef](#)]
69. Cui, Q.; Ma, Y.; Jaramillo, M.; Bari, H.; Awan, A.; Yang, S.; Zhang, S.; Liu, L.; Lu, M.; O’Connor-McCourt, M.; et al. A map of human cancer signaling. *Mol. Syst. Biol.* **2007**, *3*, 152. [[CrossRef](#)]
70. Wang, X.; Zhang, Y.; Han, Z.-G.; He, K.-Y. Malignancy of Cancers and Synthetic Lethal Interactions Associated with Mutations of Cancer Driver Genes. *Medicine* **2016**, *95*, e2697. [[CrossRef](#)]
71. Loeb, K.R.; Loeb, L.A. Significance of multiple mutations in cancer. *Carcinogenesis* **2000**, *21*, 379–385. [[CrossRef](#)]
72. Kent, D.G.; Green, A.R. Order Matters: The Order of Somatic Mutations Influences Cancer Evolution. *Cold Spring Harb. Perspect. Med.* **2017**, *7*, a027060. [[CrossRef](#)] [[PubMed](#)]
73. Wang, L.; Zhu, M.; Wang, Y.; Fan, J.; Sun, Q.; Ji, M.; Fan, X.; Xie, J.; Dai, J.; Jin, G.; et al. Cross-Cancer Pleiotropic Analysis Reveals Novel Susceptibility Loci for Lung Cancer. *Front. Oncol.* **2020**, *9*, 1492. [[CrossRef](#)] [[PubMed](#)]

74. Yildirim, M.A.; Goh, K.-I.; Cusick, M.E.; Barabasi, A.L.; Vidal, M. Drug—Target network. *Nat. Biotechnol.* **2007**, *25*, 1119–1126. [[CrossRef](#)] [[PubMed](#)]
75. Kotlyar, M.; Fortney, K.; Jurisica, I. Network-based characterization of drug-regulated genes, drug targets, and toxicity. *Methods* **2012**, *57*, 499–507. [[CrossRef](#)] [[PubMed](#)]
76. Lv, W.; Xu, Y.; Guo, Y.; Yu, Z.; Feng, G.; Liu, P.; Luan, M.; Zhu, H.; Liu, G.; Zhang, M.; et al. The drug target genes show higher evolutionary conservation than non-target genes. *Oncotarget* **2015**, *7*, 4961–4971. [[CrossRef](#)]
77. Zhu, P.; Aliabadi, H.M.; Uludağ, H.; Han, J. Identification of Potential Drug Targets in Cancer Signaling Pathways using Stochastic Logical Models. *Sci. Rep.* **2016**, *6*, 23078. [[CrossRef](#)]
78. Bedi, O.; Dhawan, V.; Sharma, P.L.; Kumar, P. Pleiotropic effects of statins: New therapeutic targets in drug design. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2016**, *389*, 695–712. [[CrossRef](#)]
79. Winzeler, E.A.; Shoemaker, D.D.; Astromoff, A.; Liang, H.; Anderson, K.; Andre, B.; Bangham, R.; Benito, R.; Boeke, J.D.; Bussey, H.; et al. Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis. *Science* **1999**, *285*, 901–906. [[CrossRef](#)]
80. Goh, K.-I.; Cusick, M.E.; Valle, D.; Childs, B.; Vidal, M.; Barabási, A.-L. The human disease network. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8685–8690. [[CrossRef](#)]
81. Bien, S.A.; Peters, U. Moving from one to many: Insights from the growing list of pleiotropic cancer risk genes. *Br. J. Cancer* **2019**, *120*, 1087–1089. [[CrossRef](#)] [[PubMed](#)]
82. Jia, X.; Shi, N.; Feng, Y.; Li, Y.; Tan, J.; Xu, F.; Wang, W.; Sun, C.; Deng, H.; Yang, Y.; et al. Identification of 67 Pleiotropic Genes Associated With Seven Autoimmune/Autoinflammatory Diseases Using Multivariate Statistical Analysis. *Front. Immunol.* **2020**, *11*, 30. [[CrossRef](#)] [[PubMed](#)]
83. Large, E.E.; Padmanabhan, R.; Watkins, K.L.; Campbell, R.F.; Xu, W.; McGrath, P.T. Modeling of a negative feedback mechanism explains antagonistic pleiotropy in reproduction in domesticated *Caenorhabditis elegans* strains. *PLoS Genet.* **2017**, *13*, e1006769. [[CrossRef](#)] [[PubMed](#)]
84. Ananthasubramaniam, B.; Herzel, H. Positive Feedback Promotes Oscillations in Negative Feedback Loops. *PLoS ONE* **2014**, *9*, e104761. [[CrossRef](#)]
85. Le, D.-H.; Kwon, Y.-K. The effects of feedback loops on disease comorbidity in human signaling networks. *Bioinformatics* **2011**, *27*, 1113–1120. [[CrossRef](#)] [[PubMed](#)]
86. Pei, F.; Li, H.; Liu, B.; Bahar, I. Quantitative Systems Pharmacological Analysis of Drugs of Abuse Reveals the Pleiotropy of Their Targets and the Effector Role of mTORC1. *Front. Pharmacol.* **2019**, *10*, 191. [[CrossRef](#)]
87. Gu, J.J.; Ryu, J.R.; Pendergast, A.M. Abl tyrosine kinases in T-cell signaling. *Immunol. Rev.* **2009**, *228*, 170–183. [[CrossRef](#)]
88. Graham, S.E.; Nielsen, J.B.; Zawistowski, M.; Zhou, W.; Fritsche, L.G.; Gabrielsen, M.E.; Skogholt, A.H.; Surakka, I.; Hornsby, W.E.; Fermin, D.; et al. Sex-specific and pleiotropic effects underlying kidney function identified from GWAS meta-analysis. *Nat. Commun.* **2019**, *10*, 1847. [[CrossRef](#)]
89. Hill, W.G.; Zhang, X.-S. Assessing pleiotropy and its evolutionary consequences: Pleiotropy is not necessarily limited, nor need it hinder the evolution of complexity. *Nat. Rev. Genet.* **2012**, *13*, 296. [[CrossRef](#)]
90. Brown, S.D.M.; Lad, H.V. The dark genome and pleiotropy: Challenges for precision medicine. *Mamm. Genome* **2019**, *30*, 212–216. [[CrossRef](#)]
91. Naldi, A.; Carneiro, J.; Chaouiya, C.; Thieffry, D. Diversity and Plasticity of Th Cell Types Predicted from Regulatory Network Modelling. *PLoS Comput. Biol.* **2010**, *6*, e1000912. [[CrossRef](#)] [[PubMed](#)]
92. Trinh, H.C.; Kwon, Y.K. Edge-based sensitivity analysis of signaling networks by using Boolean dynamics. *Bioinformatics* **2016**, *32*, i763–i771. [[CrossRef](#)]