

A systems approach to traditional oriental medicine

Hyun Uk Kim^{1,2,4}, Jae Yong Ryu¹, Jong Ok Lee^{1,2} & Sang Yup Lee^{1–4}

Analyzing structural similarities between compounds derived from traditional oriental medicine and human metabolites is a systems-based approach that can help identify mechanisms of action and suggest approaches to reduce toxicity.

The majority of modern drugs contain a single active ingredient or component developed around a proof-of-concept test of their activity against a single biological target. That any drug works by acting on a single target is of course rather simplistic, given that the human body is a complex system of intertwined and redundant metabolic, regulatory and interacting networks. Here, we analyze the molecular and systems-level characteristics of traditional oriental medicines (TOMs), which, unlike conventional drugs, are based on a multicomponent, multitarget approach. In the following Commentary, we propose how structural similarities of TOM-derived compounds to human metabolites can facilitate the identification of their mechanisms of action and potentially reduce toxicities. Finally, we propose that such strategies can facilitate further integration of TOM into modern drug discovery and development.

Metabolism, mechanism and medicine

TOM has a long history of widespread clinical

use in Asian countries in distinctive formats that share some common principles^{1,2}. Knowledge concerning the use of TOM over the millennia has grown empirically through experience and folklore, but products are often indicated to treat a great variety of seemingly unrelated symptoms, without reference to a mechanism of action or effect on an underlying disease-causing mechanism. At the same time, billions of people around the world use TOM each day, and these natural products have been the starting point for several successful Western drug development programs, including lovastatin (Mevacor) for cholesterol, artemisinin for malaria and carfilzomib (Kyprolis) for multiple myeloma.

The ultimate goal of TOM is to restore and maintain the energy and the overall balance of the human body¹. This is fundamentally different from Western drug discovery and development, which is based on the concept of targeting a specific receptor, ion channel, enzyme or regulatory protein involved in causing the disease and symptom. Considering that the human body is a complex interacting system of many networks at the level of cells, tissues, organs and whole body, Western medicine is increasingly embracing polypharmacology as a complement to the traditional approach of using a single active ingredient or component to address a single target^{3,4}. Thus, multicomponent, multitarget approaches are of interest as potential means of treating complex diseases caused by abnormal interactions among multiple molecules and pathways in a biological system^{5–7}. A notable example from TOM is the Realgar-Indigo naturalis formula, in which the major components tetraarsenic tetrasulfide, indirubin and tanshinone IIA generate synergies by intervening in distinct

biological processes to treat acute promyelocytic leukemia⁸.

One of the problems preventing TOM from more widespread adoption by modern healthcare is our poor understanding of the mechanisms of action of many of these natural products. This is of great clinical importance because many TOM active chemical ingredients have the potential to cause liver and kidney toxicity. An improved understanding of their mechanisms of action could provide insight into approaches for formulating TOM chemical compounds that have not only good synergistic medicinal effects but also fewer toxicities or side effects. It would also enable the systematic combination of different agents to address biological pathways containing multiple targets and provide a scientific rationale for relevant diseases where the underlying pathology would be addressed.

Thus far, conventional drug discovery efforts have tended to focus on signaling components, pathways and networks when discussing pathological conditions, rather than focusing on human metabolism and its components (e.g., enzymes, transport proteins and regulatory proteins), despite the importance of the latter in physiological phenomena⁹. Among the different chemical, pharmacological and systems biological approaches taken to characterize TOM¹⁰, the structural similarities of the TOM-derived compounds to human metabolites can facilitate identifying their mechanisms of action. The concept of metabolite-likeness has been important in synthetic drug discovery as represented by metabolite analogs treating various diseases, for instance, cancers^{11,12}. Indeed, several TOM-derived active compounds that are structurally similar to metabolites have been shown to interact with the target proteins.

¹Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 Plus Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea. ²BioInformatics Research Center, KAIST, Daejeon, Republic of Korea.

³BioProcess Engineering Research Center, KAIST, Daejeon, Republic of Korea. ⁴The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.

e-mail: leesy@kaist.ac.kr

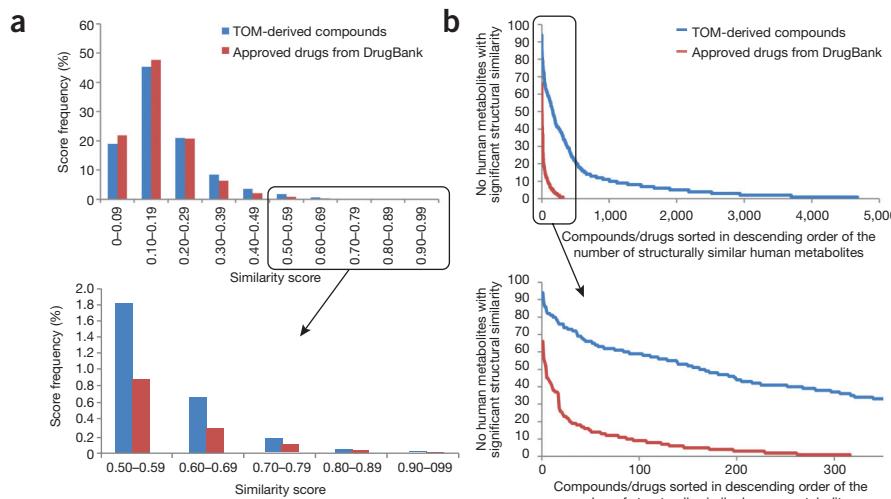


Figure 1 Overall trends in pairwise structural comparisons of the compounds in TOM and the approved drugs from DrugBank against human metabolites. **(a)** Score frequencies for each similarity score interval of 0.1. **(b)** The number of human metabolites structurally similar to 4,679 compounds in TOM and 316 approved drugs from DrugBank having similarity scores ≥ 0.77 and < 1.0 .

For example, genistein, which is structurally similar to estrogen, functions as an agonist for the estrogen receptor beta¹³, whereas emodin, which is structurally similar to cortisone, inhibits 11 β -hydroxysteroid dehydrogenase type 1 (ref. 14). However, the functions of many other TOM-derived compounds remain elusive.

In the following article, we provide a systematic study of the structural similarity of TOM-derived compounds and approved small-molecule drugs to human metabolites. Although natural products have been subjected to this type of large-scale comprehensive chemical analyses before^{15–18}, more detailed analyses are needed at the systems level. We also review design principles of

synergistic combinations of recently reported TOM-derived compounds and conclude by proposing a central role for systems biology in further advancing the TOM field.

Structural similarities

We carried out large-scale structural comparisons of TOM-derived compounds and approved drugs against human metabolites. Data files containing structural information of the active compounds observed in TOMs and human metabolites were downloaded from the TCM Database@Taiwan¹⁹ and the Kyoto Encyclopedia of Genes and Genomes (KEGG)²⁰, respectively. As a control, the same chemical data files for approved drugs were downloaded from the DrugBank 3.0

(ref. 21). The structural comparisons were conducted using the small molecule subgraph detector (SMSD²²; see **Supplementary Note** for more detail). A comparison of the chemical structures of TOM-derived compounds with human metabolites is shown in **Figure 1** (see also **Supplementary Note** and **Supplementary Fig. 1**).

Results of the extensive pairwise structural comparisons reveal that the compounds in TOM are structurally more similar to human metabolites than approved drugs from DrugBank (i.e., they have greater similarity scores; **Fig. 1a** and **Supplementary Note**). When the similarity score is lower than 0.2, more compounds in DrugBank are similar to human metabolites than those in TOM. Conversely, for similarity scores greater than 0.2, more TOM compounds are similar to human metabolites than those in DrugBank. Differences in the number of structurally similar compounds among TOMs and approved drugs become greater than twofold for similarity scores between 0.5 and 0.7 (**Fig. 1a**).

Many more TOM compounds than approved drugs are structurally similar to human metabolites (**Fig. 1b**; see **Supplementary Note** for definition of structurally similar compounds). Although TOM compounds and approved drugs have similar numbers of biological pathways with structurally similar metabolites (150 and 152 pathways, respectively; **Supplementary Table 1**), the number of TOM compounds with at least one structurally similar metabolite is much higher. Specifically, the fractions of compounds from TOM and DrugBank that have at least one structurally similar metabolite corresponded to 32.7% and 22.7%, respectively.

Box 1 Categorizing combinations

We categorized the effects of combinations of TOM compounds into four types of synergistic combinations^{25,36} (**Fig. 2**): first, complementary action; second, neutralizing action; third, facilitating action; and fourth, pharmacokinetic potentiation. The first three combinations can be considered to be pharmacodynamic synergism³⁶, whereas the last refers to pharmacokinetic synergism.

Complementary action usually involves multiple compounds positively regulating the target or targeting process together; neutralizing action involves one TOM compound downregulating the side effects of another TOM compound; facilitating action refers to the action of one compound enhancing the activity of another compound (e.g., one compound opens up a transporter for the influx of another compound into the cell); and finally, pharmacokinetic potentiation represents enhanced bioavailability of the compound (e.g., by blocking metabolic degradation of the active ingredient compound by another compound).

Categorization of such synergistic combinations is also applicable to the use of non-herbal synthetic drug combinations³⁶. (It is interesting to note that these four categories of drug synergy have parallels in the traditional Kun-Shin-Chao-Sa design principle of TOM³⁷; see **Supplementary Note**).

Using these categories for the action of specific components of a TOM, one can make predictions about how they will interact. For instance, a synergistic neutralizing combination would involve a major component of a TOM that has a mechanism of action associated with both efficacy and side effects and a second complementary component that alleviates the side effects (but not the efficacy; **Fig. 2** and **Supplementary Table 2**). It should be noted, however, that some combinations may confound these simple predictions; for example, a synergistic complementary combination where the major and complementary components show equal potencies at different target sites (**Fig. 2**).

Measurements of metabolite-likeness may inform knowledge on mechanisms of action of TOM-derived compounds (**Supplementary Table 1**). Mapping metabolites structurally similar to TOM compounds should be useful in narrowing down the candidate pathways and/or reactions for elucidating the mechanisms of action of the TOM-derived synergistic effects (see below). For instance, metabolites (e.g., carbamazepine, 2-hydroxycarbamazepine and 3-hydroxycarbamazepine) similar to indirubin found in Realgar-Indigo naturalis formula appeared to be involved in drug metabolism (**Supplementary Table 2**), and, in fact, indirubin has been known to induce drug metabolism by activating cytochromes²³. Interestingly, in the case of danshensu, whose similar metabolites are largely associated with tyrosine metabolism (e.g., 3-(4-hydroxyphenyl)lactate, 3,4-dihydroxy-L-phenylalanine, 2-carboxy-2,3-dihydro-5,6-dihydroxyindole, 3-(4-hydroxyphenyl)pyruvate, and L-tyrosine; **Supplementary Table 2**), it was shown to inhibit tyrosine phosphorylation, thereby contributing to maintaining cell-cell adhesion²⁴. It should be noted that when searching for such metabolites, the score thresholds for the structural similarity might need to be systematically adjusted to more thoroughly investigate different sets of metabolites predicted. More detailed information on structurally similar metabolites for the newly added 14 TOM-derived synergistic combinations is available in **Supplementary Table 2**.

Synergistic combinations

One potential advantage of TOM is that the presence of multiple active ingredients or compounds in any one product presents the opportunity to affect multiple targets⁵. Ma *et al.*²⁵ previously reported 24 verified combinations of TOM compounds, for which mechanisms of action are known. We have also analyzed the synergistic combinations of TOM compounds that have been reported to date through a thorough search of the literature since 2000; we regard combination studies performed before this period as non-systematic (**Supplementary Table 3**).

Our literature search identified a total of 38 (including 14 new) combinations of TOM compounds, all of which have information on compound structures and their interacting molecule targets (**Box 1** and **Fig. 2**). We then analyzed components on the basis of four categories of activity—complementary action, neutralizing action, facilitating action and pharmacokinetic potentiation (**Box 1**). On the basis of this analysis, complementary actions appear to be the most frequent mechanism of action, represent-

ing 19 combinations (with ten newly found combinations). The next most common mechanism of action of a combination was neutralizing action (11 combinations with 2 newly found combinations), followed by

pharmacokinetic potentiation (6 combinations with 1 newly found combination) and facilitating actions (2 combinations with 1 newly found combination; (**Figs. 2** and **3** and **Supplementary Table 2**).

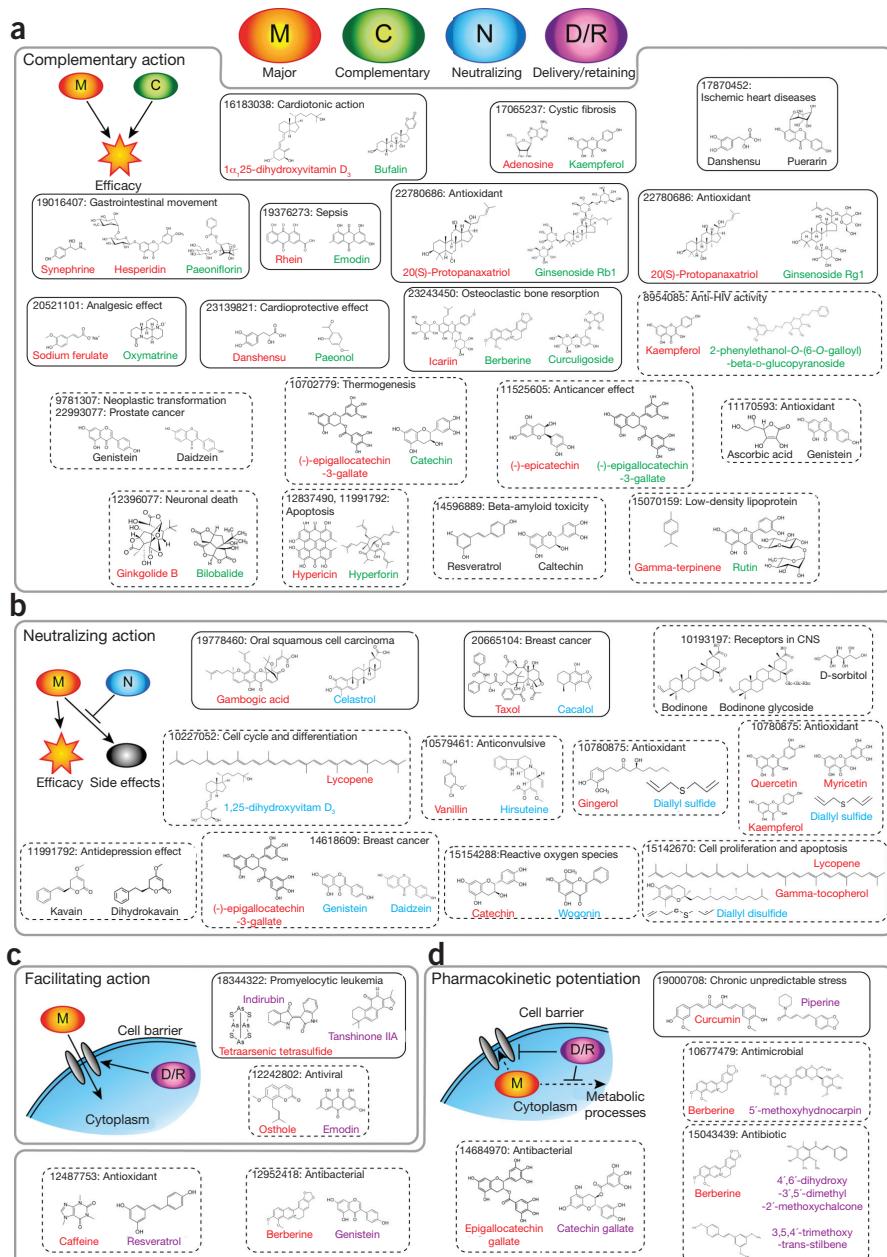


Figure 2 Mechanisms of action of 38 synergistic combinations of TOM-derived compounds. Each combination is presented in a small box; boxes with solid lines indicate combinations reviewed in this study, boxes with dashed lines, combinations from Ma *et al.*²⁵. **(a-d)** Numbers shown in each small box represent PubMed identifier (PMID) of the relevant literature. Next to them are relevant diseases or treatment effects associated with the presented synergistic combinations, including complementary action (**a**), neutralizing action (**b**), facilitating action (**c**) and pharmacokinetic potentiation (**d**). In each combination, the major component is shown in red, whereas supportive components are shown in green (complementary component; **a**, light blue (neutralizing component; **b**) or purple (delivery/retaining component; **c** and **d**). If the roles of each component in synergistic combination are not clear, their component names are shown in black. In **b**, the combination of Taxol and cacalol (PMID: 20665104) can also be said to belong to the complementary action because cacalol itself also induces apoptosis of breast cancer cells.

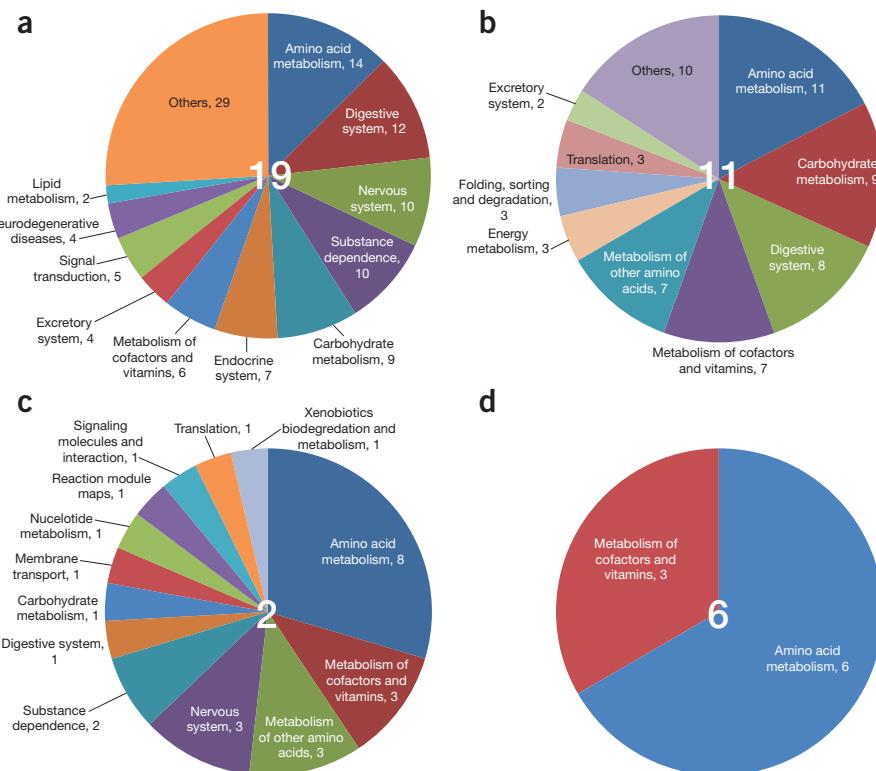


Figure 3 Distribution of submetabolisms associated with the metabolites having high structural similarities (similarity scores ≥ 0.77 and < 1.0) with the TOM-derived compounds. **(a–d)** Synergistic combinations include complementary action (**a**), neutralizing action (**b**), facilitating action (**c**) and pharmacokinetic potentiation (**d**). Large white numbers in the middle of each pie chart represent the number of elucidated synergistic combinations. Numbers next to each submetabolism category indicate the frequency of submetabolisms targeted by synergistic combinations of the compounds in TOM. Submetabolism categories are based on the KEGG. ‘Others’ in **a** and **b** indicate submetabolisms having one or two structurally similar metabolite(s).

In terms of metabolic pathways most commonly targeted by TOM compounds, it appears that amino acid metabolism has the greatest number of metabolites with similar structures to synergistic combinations of active ingredients (Fig. 3). In addition, metabolic pathways associated with cofactors and vitamins were shared by all four types of synergistic combinations. Each synergistic combination showed distinct ranks of biological pathways with respect to the number of their structurally similar metabolites, indicating their potentially differentiated mechanisms of action.

Perspectives

The degree of structural similarity between TOM compounds and human metabolites is much higher than that for human metabolites and conventional small-molecule drugs. The greater number of such compounds in TOM may provide clues as to their potential for addressing multiple targets *in vivo*. Furthermore, mapping of the structurally similar metabolites onto their associated biological pathways may suggest potential sites of interaction between TOM-derived compounds and

human biology. Analyses of the mechanisms of action of TOM compounds offer opportunities for discovering new targets and drugs that can be complementary to conventional approaches used for developing approved drugs and their combinations to date²⁶.

A previous study on the structural similarities of natural products from the TCM Database@Taiwan¹⁹ supports our data¹⁸. We note that one report has shown that natural products have greater similarities to chemical drugs than human metabolites in terms of Tanimoto similarity (0.85 versus 0.58)¹⁷. A reason for this difference may be the use of compounds from the ZINC NP database²⁷, which has information on only commercially available compounds. This suggests that the search for TOM-derived compounds with the potential for multitargeting should reach beyond the commercially available natural products.

Natural evolution and the greater pool of compounds present in TOMs might have contributed to their similarities with human metabolites being greater than those observed for approved drugs^{5,18}. Another explanation is that approved drugs usually go through

chemical optimization for target selectivity, resulting in a smaller number of structurally similar metabolites. Consequently, each individual compound present in the vast array of TOMs might influence a broader range of human metabolic pathways by simply interacting with a greater number of biological molecules.

TOM-derived components with high structural similarities to human metabolites may suggest likely mechanisms of action. Enzymes consuming metabolites structurally similar to TOM-derived compounds would be expected to be potential targets of those compounds^{13,14,28}. This concept was previously validated by several theoretical and experimental studies, including two that identified metabolite-like chemicals that effectively inhibited the growth of microbial pathogens²⁹ and treat liver cancer¹¹, respectively, and one that predicted the unintended off-target sites of commercial drugs³⁰. Thus, we expect detailed studies based on metabolite likeness to reveal mechanisms of action of the compounds in TOM. The strategy is particularly suited for identifying compounds that target metabolism, such as emodin¹⁴, mentioned above. Compounds that interact and perturb regulatory or signaling proteins might or might not necessarily show structural similarities to metabolites. Thus, the combined use of the compounds in TOM targeting the metabolic enzymes, regulatory and transport proteins, and other compounds, either natural or synthetic, targeting regulatory or signaling proteins would cover a wide range of cellular drug targets, potentially opening up synergies between TOM, its active ingredients or components and conventional small-molecule drugs.

Our analysis of the types of interactions between different components of TOMs suggests that complementary actions between components are most common. This does not necessarily mean that a TOM shows efficacy mostly through complementary actions; categorization as a complementary action might simply be more amenable to biochemical analyses for validation than other types of synergistic combinations. More importantly, the mechanisms of action for each combination should be interpreted carefully because they can be more complicated than described in the literature as a result of influences from genetic and environmental (e.g., lifestyle) conditions of an individual person.

For TOM to be fully integrated into modern medicine, rigorous evaluation of the mechanisms of action of TOM active ingredients is key. Numerous TOM-derived compounds, either alone or in combination, have yet to be characterized with respect to their effects on specific

diseases. It may be that archived documents on TOM that were written many centuries ago can still be valuable in this respect as they often have clear guidelines on the clinically validated use of specific plants and animals for treating diseases or improving disease conditions. We anticipate that the identification of active compounds of TOM and their effective targets will remain a challenge for some years, but systems biology offers one promising approach for addressing the problem. Various systems biological approaches need to be employed for deciphering them at the systems level.

The multicomponent/multitargeting principle of TOM, which aims to sustain the healthy status of the human body in its entirety, is very compatible with the holistic view of systems biology, which considers interactions among components (e.g., genes, RNAs, proteins, metabolites and metabolic fluxes) at various levels and networks in a given biological system^{6,31}. For example, network-based approaches of systems biology will provide additional insights on the compounds in TOM by linking them with their targets at the whole metabolic and regulatory network level using high-throughput techniques, which will consequently allow us to predict more effective synergistic compounds³².

Furthermore, the newly updated genome-scale human metabolic network models, such as Recon³³ and HMR³⁴ series, can be ideal platforms for data integration and analysis associated with TOM. At the moment, a wide range of omics data are available for various types of human cells. With the development of advanced methods for integrating omics data with the human metabolic network model (e.g., Recon 2.0 and HMR 2.0)³⁵, it will

be possible to advance our knowledge on the mechanisms of action of the specific TOM in the context of specific diseases, for example, by linking the expression levels of the genes and their respective proteins with the administered TOM compounds. Such systems biology studies will no doubt contribute to the modernization of TOM and the establishment of the multicomponent, multitargeting strategy as a new paradigm in medicine. Once identified, active compounds in TOM can be more efficiently produced by metabolic engineering and/or chemical synthesis. Instead of relying on impure plant extracts or tiny amounts of purified compounds for a limited spectrum of studies, the pure active compounds produced in large amounts will allow us to perform more systematic studies in various combinations.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper (doi:10.1038/nbt.3167).

ACKNOWLEDGMENTS

We are grateful to our laboratory members for sharing their computer resources and Y. Kim for his helpful advice in using supercomputers at Korea Institute of Science and Technology Information (KISTI). This work was supported by the Bio-Synergy Research Project (2012M3A9C4048759) of the Ministry of Science, ICT and Future Planning through the National Research Foundation. This work was also supported by the Novo Nordisk Foundation.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

- Cheung, F. *Nature* **480**, S82–S83 (2011).
- Jung, J.Y., Lee, J.H. & Chung, S.H. *Uisahak* **20**, 1–28 (2011).
- Dar, A.C., Das, T.K., Shokat, K.M. & Cagan, R.L. *Nature* **486**, 80–84 (2012).
- Gujral, T.S., Peshkin, L. & Kirschner, M.W. *Proc. Natl. Acad. Sci. USA* **111**, 5048–5053 (2014).
- Schmidt, B.M., Ribicky, D.M., Lipsky, P.E. & Raskin, I. *Nat. Chem. Biol.* **3**, 360–366 (2007).
- van der Greef, J. *Nature* **480**, S87 (2011).
- Xu, Z. *Nature* **480**, S90–S92 (2011).
- Wang, L. *et al.* *Proc. Natl. Acad. Sci. USA* **105**, 4826–4831 (2008).
- McKnight, S.L. *Science* **330**, 1338–1339 (2010).
- Liu, X., Wu, W.Y., Jiang, B.H., Yang, M. & Guo, D.A. *Trends Pharmacol. Sci.* **34**, 620–628 (2013).
- Agren, R. *et al.* *Mol. Syst. Biol.* **10**, 721 (2014).
- Hebar, A., Valent, P. & Selzer, E. *Expert Rev. Clin. Pharmacol.* **6**, 23–34 (2013).
- McCarty, M.F. *Med. Hypotheses* **66**, 1093–1114 (2006).
- Feng, Y. *et al.* *Br. J. Pharmacol.* **161**, 113–126 (2010).
- Dobson, P.D., Patel, Y. & Kell, D.B. *Drug Discov. Today* **14**, 31–40 (2009).
- Ganesan, A. *Curr. Opin. Chem. Biol.* **12**, 306–317 (2008).
- Khanna, V. & Ranganathan, S. *J. Cheminform.* **3**, 30 (2011).
- Jayaseelan, K.V., Moreno, P., Truszkowski, A., Ertl, P. & Steinbeck, C. *BMC Bioinformatics* **13**, 106 (2012).
- Chen, C.Y. *PLoS ONE* **6**, e15939 (2011).
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. *Nucleic Acids Res.* **40**, D109–D114 (2012).
- Knox, C. *et al.* *Nucleic Acids Res.* **39**, D1035–D1041 (2011).
- Rahman, S.A., Bashton, M., Holliday, G.L., Schrader, R. & Thornton, J.M. *J. Cheminform.* **1**, 12 (2009).
- Spink, B.C., Hussain, M.M., Katz, B.H., Eisele, L. & Spink, D.C. *Biochem. Pharmacol.* **66**, 2313–2321 (2003).
- Ding, M. & Yuan, Y.J. *J. Pharm. Pharmacol.* **59**, 1027–1033 (2007).
- Ma, X.H. *et al.* *Drug Discov. Today* **14**, 579–588 (2009).
- Hawkes, D. & Benhamu, J. *Trends Pharmacol. Sci.* **35**, 111–112 (2014).
- Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S. & Coleman, R.G. *J. Chem. Inf. Model.* **52**, 1757–1768 (2012).
- Kim, T.Y., Kim, H.U. & Lee, S.Y. *Metab. Eng.* **12**, 105–111 (2010).
- Kim, H.U. *et al.* *Mol. Syst. Biol.* **7**, 460 (2011).
- Keiser, M.J. *et al.* *Nature* **462**, 175–181 (2009).
- Kim, H.U., Sohn, S.B. & Lee, S.Y. *Biotechnol. J.* **7**, 330–342 (2012).
- Li, S., Zhang, B. & Zhang, N. *BMC Syst. Biol.* **5 Suppl 1**, S10 (2011).
- Thiele, I. *et al.* *Nat. Biotechnol.* **31**, 419–425 (2013).
- Mardinoglu, A. *et al.* *Nat. Commun.* **5**, 3083 (2014).
- Blazier, A.S. & Papin, J.A. *Front. Physiol.* **3**, 299 (2012).
- Jia, J. *et al.* *Nat. Rev. Drug Discov.* **8**, 111–128 (2009).
- Oh, C. *Uisahak* **22**, 1–40 (2013).