Bioorganic Chemistry 86 (2019) 273–276

Contents lists available at ScienceDirect

Bioorganic Chemistry

journal homepage: www.elsevier.com/locate/bioorg

Essay: Avoiding unfounded health claims on small molecules in scientific literature

Bart I. Roman*

Department of Green Chemistry and Technology, Coupure Links 653, 9000 Gent, Belgium
Cancer Research Institute Gent (CRIG), Corneel Heymanslaan 10, 9000 Gent, Belgium

ARTICLE INFO

Keywords:
Health claims
Drug discovery
False positives
Hit validation
Biological filters

ABSTRACT

The publication of unfounded health claims on small molecules in peer-reviewed scientific literature is a problem that requires attention. It undermines the evidence-based decision making processes of modern-day society, weakens the credibility of the scientific enterprise, and diverts resources to futile research efforts. In the present essay we discuss some human and scientific causes behind the issue. We propose a number of actions to be taken up by scientists, referees and publishers. One particularly important factor is the issue of enigmatic compound behavior in biological assays. We therefore also introduce the idea of biological filters, a pattern recognition method to triage enigmatic compounds into valuable hits and false positives, based on the entirety of their biological effects in cell-based systems.

1. Introduction

A cornerstone of scientific reporting is providing the necessary empirical data that supports or allows to reject a claim or hypothesis. Synthetic organic chemistry papers generally do well in this respect, since validation of a hypothesis is largely inherent to the work, e.g. the isolation and/or spectroscopic characterization of crucial chemical species to substantiate a proposed reaction mechanism or structure.

Unfortunately, chemists sometimes appear to abandon this evidence-based path of reasoning when it comes to making statements about the utility of their compounds as physiologically active agents. Too often, and also in reputable journals, one is confronted with statements such as “the activity of [our compounds] in [any single *in vitro* test] underlines their promising potential as therapeutics for [any disease] or as agents for maintaining [any aspect of health and well-being].” Such claims should be substantiated by at least a minimal data set on activity validation, selectivity and physicochemical properties.

The publication of these unfounded health claims on small molecules in peer-reviewed literature poses a hazard to the scientific enterprise, society and the economy. Literature containing such claims is fueling large amounts of futile research that consume valuable resources. It moreover poisons the development of bona fide health products, as it constitutes prior art that may prevent patentability. Finally, legislative bodies [1] and consumers [2] rely on science for an evidence-based valuation of nutritional and physiological claims regarding substances and health foods, and for the rejection of quackery. In the following paragraphs we want to highlight some of the human and scientific factors at the origin of the issue, and suggest relevant corrective and preventive actions.

2. Unfounded health claims: a multifactorial issue

A good share of unsubstantiated health claims seems to concentrate on particular classes of compounds, which often have a highly enigmatic biological activity profile. Many of these compounds have been flagged and fall under the definition of invalid metabolic panacea (IMPs) [3] or pan assay interference inhibitors (PAINS) [4]. Despite numerous warnings in high-profile journals, new papers on such compounds appear on a daily basis. Case in point, the problem of unfounded health claims started to vex us while repeatedly encountering certain polyphenols, e.g. epigallocatechin gallate **1**, in different projects and biological contexts (Chart 1) [5–7]. Other usual suspects are for example curcumin **2** and many rhodanines **3** [8]. These molecules are known covalent modifiers, membrane disruptors, and/or metal complexers. They perturb normal protein function in nonspecific and non-druglike ways. Polyphenols, curcumin and rhodanines, amongst other enigmatic molecules, are moreover plagued by poor solubility and bioavailability. This all has resulted in the publication of numerous incorrect health claims by non-critical observers. The ready synthetic accessibility of many of these compound families has moreover led to

* Address: Department of Green Chemistry and Technology, Coupure Links 653, 9000 Gent, Belgium.
E-mail address: bart1.roman@ugent.be.

https://doi.org/10.1016/j.bioorg.2019.02.002
Received 12 September 2018; Received in revised form 11 December 2018; Accepted 1 February 2019
Available online 02 February 2019
0045-2068/

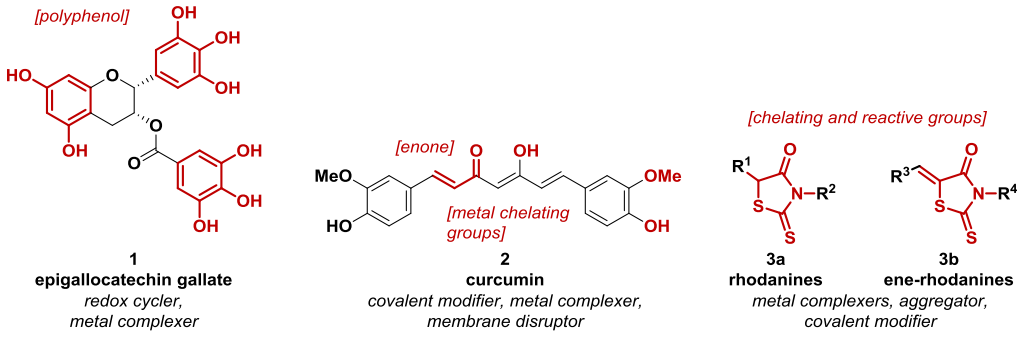


Chart 1. Examples of enigmatic compounds and classes with problematic activity profiles, giving rise to unsubstantiated health claims.

an overwhelming body of– in the vast majority of cases – futile structure-activity explorations and (pre)clinical work [9,10]. In spite of all the warnings, why is a good percentage of the literature alerts in our mail box still dealing with flagged molecules? Scientists are shaped by their culture and society as much as any other human being, which may bring about irrational expectations or beliefs in the virtues of certain chemicals. Many of the literature compounds with an unsubstantiated activity profile are indeed natural products which have been used in traditional healing practices, or as herbs and spices, for over centuries. It may be hard for researchers to maintain a critical distance to topics of such intense culturally-driven interest. The magnitude of the impact of this phenomenon in chemistry deserves further investigation [11–13].

Unfounded health claims are not limited to enigmatic compounds alone. The present system of scientific recognition, in bolder terms the maxim 'publish or perish', also is an important driver of precarious activity claims [14]. If after a hard and lengthy study researchers have to conclude that their new compounds or methodology lack sufficient impact to be publishable on their own right, a rescue mechanism often – and regrettably – kicks in where any hint of biological activity is used to safeguard publication potential. It becomes especially dangerous when chemists themselves – in search of rapid positive biological data – cross into the area of biology and reproduce literature protocols of 'simple' experiments (e.g. antioxidant activity) with little or no context provided by specialists [15–18].

A related issue is insufficient insight into the multitude of aspects that make up medicinal chemistry and preclinical compound development [19]. Proper compound characterization and validation include counter screens, physicochemical and pharmacokinetic profiling and a view on selectivity, SAR and room for chemical optimization. Lack of appreciation of these aspects leads to reports claiming, for example, promising anticancer effects for a compound based on mere cytotoxicity data in a cancer cell line, biological effects at physiologically irrelevant concentrations at which compounds are invariably promiscuous, or blind focus on only the most active compound in a screen.

Target validation constitutes an additional order of complexity that is often underappreciated [20]. Too often compounds are attributed health claims based on affinity for improperly validated targets. A common issue is a lack of evidence that target engagement produces a meaningful therapeutic effect. It is, for example, striking how popular antioxidant activity still is in certain layers of the literature, while clinical support for antioxidant intake beyond normal dietary habits as preventative therapy is lacking, and may even be detrimental in some settings [21–24].

3. Improving the state of the literature

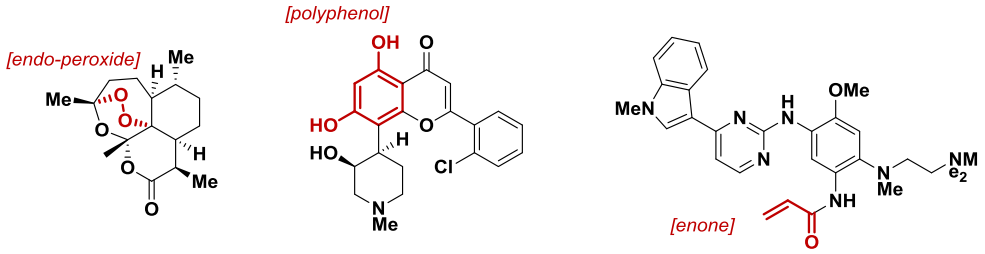
Installing much-needed additional scientific rigor regarding health claims will require considerable attention of all parties involved: researchers, publishers and referees. Many unfounded claims can be traced back to hiatuses in the training of chemists or lack of experience in the field. Additional insight in the black art of medicinal chemistry and superior appreciation of the philosophy of science and scientific reporting are needed. Chemists must always seek aid from relevant specialists to generate and interpret data on matters outside their area of expertise. They should not give in to the pressure of publishing premature or only positive results and must keep pushing for a greater appreciation of quality rather than quantity from their host institutions, granting bodies and the community as a whole [14].

Journals should implement stricter criteria for the acceptance of biological activity data, just as they now nearly all do for confirmation of purity and structure. The availability of orthogonal activity data should be standard. Some journals have already adopted a rigorous policy regarding the examination of actives against known classes of problematic compounds, in particular for the classical PAINS [25]. Such analyses should become universal. We also call for criteria regarding the choice of wording when it comes to describing compounds (e.g. bioactives, tools, hits, leads, active pharmaceutical ingredient versus drug) and effects (e.g. biochemical, biological, functional, therapeutic effects). Publishers should provide checklists of required data before certain terms can be used. Multidisciplinary findings should only be accepted when backed by a multidisciplinary authorship and after peer review by a multidisciplinary panel of referees.

The above only addresses a number of human factors in the proliferation of unfounded health claims. There is, however, also an urgent need for better a fundamental understanding of the activity profile of particular chemical families, which cannot be considered PAINS or IMPs but have an apparent enigmatic behavior that leaves medicinal chemists in doubt and discord [26,27]. Clearer insights would enable the construction of better guidelines and tools to prevent scientists, publishers and reviewers from getting lured into reporting unfounded activity claims on false positives, and from rejecting '*false* false positives'.

Current filtering tools are mostly focused on eliminating such compounds by use of chemical descriptors. However, these approaches remain too much of a black box when used stand-alone. Apart from a limited number of structural motifs such as quinones, most medicinal chemists will agree that there is a long list of substructure alerts that are not more effective than random chance in identifying a problematic structure [28]. The relevance of structural flags is also context-dependent. Many filters might for example not have let pass the peroxide bridge in the artemisinins **4** (see Chart 2), the polyphenolic scaffold of CDK9 kinase inhibitor flavopiridol **5** or the enone function in several third-generation epidermal growth factor receptor tyrosine kinase inhibitors such as osimertinib **6** [29,30]. Also, certain toxic substructures are tolerable in e.g. acute malignancies but not in a milder, chronic setting. Others have adapted screening protocols, such as running biochemical screens in the presence and absence of detergent to detect irrelevant inhibition caused by compound aggregation [26]. Here too, however, no definitive list of permanent aggregators could be compiled for curating libraries or triaging HTS hits because aggregation phenomena proved highly assay- and condition-dependent.

How can filter methods then be improved to avoid 'false false



4
artemisinin
part of standard P. falciparium malaria treatment

5
flavopiridol
nanomolar inhibitor of cyclin-dependant kinases, FDA orphan drug status for AML

6
osimertinib
FDA-approved third-generation epidermal growth factor receptor tyrosine kinase inhibitor

Chart 2. Bona fide inhibitors and drugs containing flagged substructures.

positives'? The most direct way of assessing bona fide activity of a compound is by running a significant number of orthogonal assays using different readouts and systems. This is impractical for most academic labs and too costly for large screening libraries, if enough orthogonal set-ups for assessing a particular phenomenon can be developed in the first place. Beall et al. have recently [4] called for a similar but more collective effort concerning the screening of all relevant PAINS structures against multiple targets in multiple assays. In essence, however, use of the data coming forth from such exercises will still require extrapolation to other assays and from biochemical set-ups to cell-based systems or *in vivo* models.

4. Introducing a new concept: biology-based holistic filters

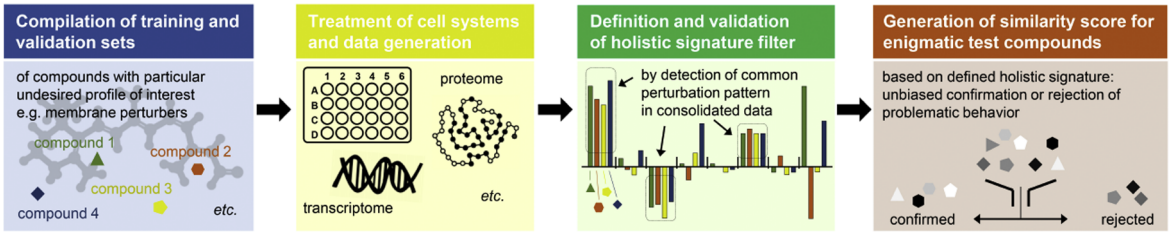
We call for the development of an additional, conceptually different approach in which biological filters are used to triage compounds based on the entirety of their biological effects in complex (i.e. cell-based) systems rather than in isolated models. Such holistic filters will recognize patterns of direct interaction targets and biological consequences that are indicative of a particular problematic behavior (Scheme 1). Pattern recognition can be developed for any undesired profile, ranging from non-specific PAINS behaviors to true polypharmacology or non-tolerable activities. The proposed approach relates to the aforementioned strategies much like phenotypic screening relates to target-based methods: compounds are evaluated in a cell-based (or higher-complexity) model that simultaneously probes for a multitude of biological and physicochemical effects in a model with relevance to the *in vivo* setting, rather than in isolated biochemical systems. Compared to chemical filters, which are in essence the outcome of a structure-property/activity relationship, the suggested biology-based filters are solely based on the observed properties/activities themselves and thus eliminate the need for correlating biology with chemistry.

To generate the necessary data for pattern recognition, we will profile sets of compounds with undesirable behavior (e.g. sets of chelators, redox active compounds, membrane perturbers, frequent hitters, hitters of hERG and other non-tolerable activities) and controls in relevant cell systems using a range of proteomics and transcriptomics techniques. Bioinformatics tools will then be used to deduce signature patterns for each set. Essential in these exercises will be the careful compilation of a manageable range of cell systems that ensures a comprehensive coverage of the relevant biological context that is to be the domain of applicability of the filter (e.g. cancer, inflammation, etc.). We are gathering preliminary data for a proof-of-concept of the proposed biological filters, with current focus on secondary metabolites in cancer. We hope to disclose our initial findings in the near future.

The suggested strategy is not without its challenges. Generating the data necessary to develop the proposed filters is a considerable task, but the same is true for structure-based approaches proposed by others. Similar efforts have moreover been successful in other domains – such as the Connectivity Map (CMap) of the NIH LINCS consortium [31] – thus proving the feasibility of the concept. Data of the CMap might even be suitable for the training or external validation of the recognition maps of the biological filters. Cost and accessibility of RNA sequencing and proteomics analyses may be regarded as an issue. However, other approaches require extensive biochemical profiling in sporadically used assays (e.g. individual aggregation, perturbation and binding assays). Most universities and research institutions nowadays have access to transcriptomics and proteomics core facilities which run the necessary analyses on a routine basis. Prices for sequencing and MS-based proteomics have also dropped sharply over the last years, which makes the strategy economically competitive [32]. It is moreover reasonable to expect that groups working on therapeutics for a particular disease area will have access to an elementary panel of relevant cell systems for that field of application.

5. Conclusions

Our proposed biology-based filters would enable an unbiased and unambiguous recognition of 'ugly' parts of chemical space for a particular application context. Complementing this strategy with structure-based triage would result in an even more powerful consensus-based approach. In combination with the proposed actions to tackle the human factors in the story, we hope to have provided a number of effective strategies to help clear the literature from poor 'hit' molecules and unfounded activity claims, or at least to have stimulated the scientific debate on the need to do so.



Scheme 1. Development and application of a biology-based holistic filter for a particular undesired profile of interest (e.g. membrane perturbation).

Conflicts of interest

The author declares no conflicts of interest.

References

[1] European Parliament & Council, Regulation (EC) 1924/2006 on nutrition and health claims made on foods, Off. J. Eur. Communities. L 404 (2006) 1–15.
[2] D.M. Marcus, Dietary supplements: What's in a name? What's in the bottle? Drug Test. Anal. 8 (2016) 410–412, https://doi.org/10.1002/dta.1855.
[3] J. Bisson, J.B. McAlpine, J.B. Friesen, S.-N. Chen, J. Graham, G.F. Pauli, Can invalid bioactives undermine natural product-based drug discovery? J. Med. Chem. 59 (2016) 1671–1690, https://doi.org/10.1021/acs.jmedchem.5b01009.
[4] J.B. Baell, J.W.M. Nissink, Seven year itch: pan-assay interference compounds (PAINS) in 2017—utility and limitations, ACS Chem. Biol. 13 (2018) 36–44, https://doi.org/10.1021/acschembio.7b00903.
[5] B.I. Roman, T.S.A. Heugebaert, M.E. Bracke, C.V. Stevens, Assessment of the anti-neoplastic potential of chalcones in animal models, Curr. Med. Chem. 20 (2013), https://doi.org/10.2174/092986713804806630.
[6] B.I. Roman, T. De Ryck, A. Patronov, S.H. Slavov, B.W.A. Vanhoecke, A.R. Katritzky, M.E. Bracke, C.V. Stevens, 4-Fluoro-3′,4′,5′-trimethoxychalcone as a new anti-invasive agent. From discovery to initial validation in an in vivo metastasis model, Eur. J. Med. Chem. 101 (2015), https://doi.org/10.1016/j.ejmech.2015.06.029.
[7] S. Verhasselt, B.I. Roman, O. De Wever, K. Van Hecke, R. Van Deun, M.E. Bracke, C.V. Stevens, Discovery of (S)-3′-hydroxyblebbistatin and (S)-3′-aminoblebbistatin: polar myosin II inhibitors with superior research tool properties, Org. Biomol. Chem. 15 (2017), https://doi.org/10.1039/c7ob00006e.
[8] J. Baell, M.A. Walters, Chemistry: chemical con artists foil drug discovery, Nature. 513 (2014) 481–483, https://doi.org/10.1038/513481a.
[9] K.M. Nelson, J.L. Dahlin, J. Bisson, J. Graham, G.F. Pauli, M.A. Walters, The essential medicinal chemistry of curcumin, J. Med. Chem. 60 (2017) 1620–1637, https://doi.org/10.1021/acs.jmedchem.6b00975.
[10] T. Tomašić, L. Peterlin Mašič, Rhodanine as a scaffold in drug discovery: a critical review of its biological activities and mechanisms of target modulation, Expert Opin. Drug Discov. 7 (2012) 549–560, https://doi.org/10.1517/17460441.2012.688743.
[11] O.J. Jegede, P.A.O. Okebukola, The effect of instruction on socio-cultural beliefs hindering the learning of science, J. Res. Sci. Teach. 28 (1991) 275–285, https://doi.org/10.1002/tea.3660280308.
[12] D. Baker, P.C.S. Taylor, The effect of culture on the learning of science in non-western countries: the results of an integrated research review, Int. J. Sci. Educ. 17 (1995) 695–704, https://doi.org/10.1080/0950069950170602.
[13] H. Zhang, Z. Zhang, The functional-analogical explanation in Chinese science and technology - a case study of the theory of Yin-Yang and five elements, in: Magnani, L, Li, P. (Ed.), Model. Reason. Sci. Technol. Med., Springer-Verlag, Berlin, Heidelberger Platz 3, D-14197 BERLIN, GERMANY, 2007, pp. 245+.
[14] D.R. Grimes, C.T. Bauch, J.P.A. Ioannidis, Modelling science trustworthiness under publish or perish pressure, R. Soc. Open Sci. 5 (2018) 171511, , https://doi.org/10.1098/rsos.171511.
[15] H.A. Moharram, M. Youssef, Methods for Determining the Antioxidant Activity: A Review, 2014.
[16] S.G. Klochkov, M.E. Neganova, S.V. Afanas'eva, E.F. Shevtsova, Synthesis and antioxidant activity of securinine derivatives, Pharm. Chem. J. 48 (2014) 15–17, https://doi.org/10.1007/s11094-014-1035-5.
[17] I. Fernandez-Pastor, A. Fernandez-Hernandez, F. Rivas, A. Martinez, A. Garcia-Granados, A. Parra, Synthesis and antioxidant activity of hydroxytyrosol alkyl-carbonate derivatives, J. Nat. Prod. 79 (2016) 1737–1745, https://doi.org/10.1021/acs.jnatprod.6b00124.
[18] A. Gupta, S. Kalpana, J. Malik, Synthesis and In vitro antioxidant activity of new 3-substituted-2-oxindole derivatives, Indian J. Pharm. Sci. 74 (2012) 481, https://doi.org/10.4103/0250-474X.108445.
[19] J. Hughes, S. Rees, S. Kalindjian, K. Philpott, Principles of early drug discovery, Br. J. Pharmacol. 162 (2011) 1239–1249, https://doi.org/10.1111/j.1476-5381.2010.01127.x.
[20] J. Blagg, P. Workman, Chemical biology approaches to target validation in cancer, Curr. Opin. Pharmacol. 17 (2014) 87–100, https://doi.org/10.1016/j.coph.2014.07.007.
[21] D. Wilson, K. Goggin, K. Williams, M.M. Gerkovich, N. Gqaleni, J. Syce, P. Bartman, Q. Johnson, W.R. Folk, Consumption of Sutherlandia frutescens by HIV-Seropositive South African adults: an adaptive double-blind randomized placebo controlled trial, PLoS One 10 (2015) e0128522, https://doi.org/10.1371/journal.pone.0128522.
[22] J. Watson, Oxidants, antioxidants and the current incurability of metastatic cancers, 120144–120144, Open Biol. 3 (2013), https://doi.org/10.1098/rsob.120144.
[23] J. Bouayed, T. Bohn, Exogenous antioxidants—double-edged swords in cellular redox state: health beneficial effects at physiologic doses versus deleterious effects at high doses, Oxid. Med. Cell. Longev. 3 (2010) 228–237, https://doi.org/10.4161/oxim.3.4.12858.
[24] V.I. Sayin, M.X. Ibrahim, E. Larsson, J.A. Nilsson, P. Lindahl, M.O. Bergo, Antioxidants accelerate lung cancer progression in mice, 221ra15-221ra15, Sci. Transl. Med. 6 (2014), https://doi.org/10.1126/scitranslmed.3007653.
[25] Guidelines for Authors of the Journal of Medicinal Chemistry, 2018. http://pubs.acs.org/paragonplus/submission/jmcmar/jmcmar_authguide.pdf (accessed November 27, 2018).
[26] A. Jadhav, R.S. Ferreira, C. Klumpp, B.T. Mott, C.P. Austin, J. Inglese, C.J. Thomas, D.J. Maloney, B.K. Shoichet, A. Simeonov, Quantitative analyses of aggregation, autofluorescence, and reactivity artifacts in a screen for inhibitors of a thiol protease, J. Med. Chem. 53 (2010) 37–51, https://doi.org/10.1021/jm901070c.
[27] S. Jasial, Y. Hu, J. Bajorath, How frequently are pan-assay interference compounds active? Large-scale analysis of screening data reveals diverse activity profiles, low global hit frequency, and many consistently inactive compounds, J. Med. Chem. 60 (2017) 3879–3886, https://doi.org/10.1021/acs.jmedchem.7b00154.
[28] S.J. Capuzzi, E.N. Muratov, A. Tropsha, Phantom PAINS: problems with the utility of alerts for P an- A ssay In terference compound S, J. Chem. Inf. Model. 57 (2017) 417–427, https://doi.org/10.1021/acs.jcim.6b00465.
[29] J.B. Baell, Feeling Nature's PAINS: natural products, natural product drugs, and pan assay interference compounds (PAINS), J. Nat. Prod. 79 (2016) 616–628, https://doi.org/10.1021/acs.jnatprod.5b00947.
[30] L. Chen, W. Fu, L. Zheng, Z. Liu, G. Liang, Recent progress of small-molecule epidermal growth factor receptor (EGFR) Inhibitors against C797S resistance in non-small-cell lung cancer, J. Med. Chem. 61 (2018) 4290–4300, https://doi.org/10.1021/acs.jmedchem.7b01310.
[31] A. Subramanian, R. Narayan, S.M. Corsello, D.D. Peck, T.E. Natoli, X. Lu, J. Gould, J.F. Davis, A.A. Tubelli, J.K. Asiedu, D.L. Lahr, J.E. Hirschman, Z. Liu, M. Donahue, B. Julian, M. Khan, D. Wadden, I.C. Smith, D. Lam, A. Liberzon, C. Toder, M. Bagul, M. Orzechowski, O.M. Enache, F. Piccioni, S.A. Johnson, N.J. Lyons, A.H. Berger, A.F. Shamji, A.N. Brooks, A. Vrcic, C. Flynn, J. Rosains, D.Y. Takeda, R. Hu, D. Davison, J. Lamb, K. Ardlie, L. Hogstrom, P. Greenside, N.S. Gray, P.A. Clemons, S. Silver, X. Wu, W.-N. Zhao, W. Read-Button, X. Wu, S.J. Haggarty, L.V. Ronco, J.S. Boehm, S.L. Schreiber, J.G. Doench, J.A. Bittker, D.E. Root, B. Wong, T.R. Golub, A next generation connectivity map: L1000 Platform and the First 1,000,000 Profiles, Cell 171 (2017) 1437–1452.e17, https://doi.org/10.1016/j.cell.2017.10.049.
[32] J.A. Reuter, D.V. Spacek, M.P. Snyder, High-throughput sequencing technologies, Mol. Cell. 58 (2015) 586–597, https://doi.org/10.1016/j.molcel.2015.05.004.