Current Perspectives on Multi-Omics Data Integration With Application on Toxicity Biomarkers Discovery

Jane Synnergren[1]* and Pierre Dönnes[2]

[1]*Systems Biology Research Centre, University of Skövde, Sweden*

[2]*SciCross AB, Skövde, Sweden*

Submission: March 12, 2018; **Published:** March 14, 2018

*Corresponding author: Jane Synnergren, Systems Biology Research Centre, School of Bioscience, University of Skövde, Sweden, Tel: +46 (0)708 806495; Email: jane.synnergren@his.se

Opinion

Biomarkers are becoming increasingly important in pharmaceutical drug discovery for early detection of therapeutic and adverse responses, and have promising utility for patient stratification. There is currently a huge interest and high expectations around integrated analyses of multiple high-throughput omics datasets, with applications in for example personalized medicine and biomarker discovery. The recent advances in development of high-throughput platforms for omics data generation have greatly contributed to the huge amounts of datasets that now are available through publicly accessible biological databases. These datasets constitute valuable resources for the identification of novel biomarkers for various purposes including toxicity assessment.

The latest revolution in technologies for generation of large-scale omics data has turned biomarker research towards more data-driven approaches using high-throughput techniques that produce large-scale datasets. Genomics, transcriptomics, proteomics and metabolomics datasets are now generated at a high pace, using well established technologies. In addition, microbiome sequencing, lipidomics and multiplex immunoassays provide valuable information. These methods provide new possibilities for advanced biomarker research of particular relevance for toxicity and drug safety assessment. However, many challenges related to large-scale data remain, requiring novel approaches for analysis, visualization, interpretation and integration of data. For example, often these high-throughput datasets are incomplete and needs manual curation before they can be used for automated data mining. Furthermore, the increasing diversity of experimental techniques, the high dimensionality of the resulting data, the noise in high-throughput measurements and the nature of the underlying biology result in substantial challenges in omics data analyses that need innovative solutions [1].

In response to this, more data science oriented techniques are now being explored and evaluated for predictions of toxic response with promising results. of special interest are various deep learning methods that successfully have been applied to diverse fields such as image analysis, speech recognition, natural language processing and various bioinformatics problems. Interestingly, reports are also emerging on how various deep learning and other artificial intelligence based techniques have been applied in computational methods for toxicity predictions [2-4]. It is indeed encouraging to see how the combination of emerging computational approaches, high-throughput screening assays and model systems now are beginning to provide comprehensive and clinically relevant data on toxicity to be used together with advanced data science oriented algorithms. These multidisciplinary approaches will most likely pave the way for more advanced toxicity predictions with improved accuracy.

The interest in omics data integration for applications in toxicity testing, drug discovery and personalized medicine has dramatically increased the last few years and several multi-national initiatives and consortiums have been established aiming for development of methods for automated or semi-automated integration of multi-omics data [5,6]. These initiatives are contributing to building new knowledge on omics data integration that will strengthen the development of methods for prediction of specific toxic responses in defined populations. Proposed approaches aim towards patient stratification for optimal risk-benefit including the establishment of safe levels of exposure to new compounds and the identification of subgroups of patients that may have increased risk for adverse drug responses [7]. Most of the currently used methods for risk assessment does not account for differences between individuals when it comes to response of drug exposure, which is a severe limitation. Moreover, pre-clinical safety testing is

often performed on a limited number of compounds and use mainly animal models, which are costly, time-consuming and have low translatability to humans [7]. In vitro models based on human cells and large-scale omics technologies in early drug discovery will most likely help to generate data that give better insights to the development of clinically relevant biomarkers. Thus, algorithms for performing in silico predictions of safety risks in humans based on comprehensive datasets could provide accurate and cost-effective tools to identify potential health risks to specific populations [7]. However, previous prediction algorithms have been limited by lack of data about population variability and difficulties in extrapolating from model organisms [8] and these limitations are now starting to find solutions. Ongoing efforts in data sharing using well-described ontologies and standard formats are also an important driver in the field. Further refinement and integration of these approaches will be of critical importance for taking future steps towards accurate prediction of the toxicity of drugs using multiple omics data.

References

1. B. Berger, J Peng, M Singh (2013) Computational solutions for omics data. Nat Rev Genet 14(5): 333-346.

2. S Ekins (2016) The Next Era: Deep Learning in Pharmaceutical Research. Pharm Res 33(11): 2594-2603.

3. A Mayr, G Klambauer, T Unterthiner, S Hochreiter (2016) DeepTox: Toxicity Prediction using Deep Learning. Front Environ Sci 3: 80.

4. L Zhang, CM McHale, N Greene, RD Snyder, IN Rich, et al. (2014) Emerging approaches in predictive toxicology. Environ Mol Mutagen 55(9): 679-688.

5. G Koscielny, Peter An, D Carvalho Silva, J A Cham, L Fumis, et al. (2017) Open Targets: a platform for therapeutic target identification and validation. Nucleic Acids Res 45: D985-D994.

6. R N Smith, J Aleksic, D Butano, A Carr, S Contrino, et al. (2012) InterMine: a flexible data warehouse system for the integration and analysis of heterogeneous biological data. Bioinformatics 28(23): 3163-3165.

7. F Eduati, LM Mangravite, T Wang, H Tang, JC Bare, et al. (2015) Prediction of human population responses to toxic compounds by a collaborative competition. Nat Biotechnol 33(9): 933-940.

8. L Zeise, FY Bois, WA Chiu, D Hattis, I Rusyn, et al. (2013) Addressing human variability in next-generation human health risk assessments of environmental chemicals. Environ Health Perspect 121(1): 23-31.

Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
  (Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission
https://juniperpublishers.com/online-submission.php