Enliven Archive
Intensifying Thoughts

Review Article | www.enlivenarchive.org | Enliven: Challenges in Cancer Detection and Therapy

ISSN: 2376-046X

Precision Oncology beyond Targeted Therapy: Combining Omics Data with Machine Learning Matches the Majority of Cancer Cells to Effective Therapeutics

Nancy William

Department of Oncology, University of Canberra, Building 1/11 Kirinari St, Bruce ACT 2617, Australia

*Corresponding author: Nancy William, Department of Oncology, University of Canberra, Building 1/11 Kirinari St, Bruce ACT 2617, Australia, E-mail: nancywilliam90@yandex.com

Received Date: 16th January 2018
Accepted Date: 08th March 2018
Published Date: 9th March 2018

Citation: William N (2018) Precision Oncology beyond Targeted Therapy: Combining Omics Data with Machine Learning Matches the Majority of Cancer Cells to Effective Therapeutics Enliven: Enliven: Challenges Cancer Detect Therv 3(1): 003.

Precision medicine has gained prominence is contemporary medical science. Besides, automated and standardized assessment of patient data is supported by open source machine-learning [1]. The aim of precision medicine includes making correct prediction of ideal pharmacotherapies from the genomic profiles of tumors. In an ideal situation, the predictions mostly depend on robust cause and effect relationships [2]. The researchers propose an open source platform to predict personalized drug responses from gene expression profiles. The platform uses a standard recursive feature elimination (RFE) technique and an extremely adaptable support vector machine (VSM) algorithm [3]. They used the National Cancer Institute panel to collect drug response and gene expression data obtained from sixty human cancer cell lines (NCI-60) [4]. The data was used to build drug specific models which are highly precise in making prediction of the drug responsiveness of diverse cancer cell lines inclusive of those which comprise the latest NCI-DREAM Challenge [5]. Furthermore, the researchers prove that the predictive precision is maximized when the learning dataset fails to pre-filter genes which are regarded as activating factors of cancer onset or progression [6]. After applying their model to ovarian cancer (OC) patient gene expression datasets which publically available, they were able to make predictions which were similar with the responses contained in the literature [7]. Importantly, by developing an open source algorithm they intended to facilitate its testing in diverse cancer contexts and types which can promote modifications and improvements which are largely community-driven in the successive applications [8,9].

Quantitative trait loci (QTL) mapping, genome-wide association studies (GWAS), and human genome sequencing have immensely assisted in creating greater understanding about the molecular pathways linked to human diseases [10,11]. The use of open-source scripts and liberal data sharing have benefitted such efforts [12]. Lately, the field of personalized cancer drug prediction has integrated machine learning (ML), each linked with varying levels of success for instance, the pRRocphetic and Bioconductor [13]. As indicated, the researchers proposed an open source software platform to predict response of cancer drug whereby during the pilot phase they explored the NCBI Gene Expression Omnibus (GEO) to access publicly available datasets which were then they were subjected to formatting and before partitioning the array files into experimental and learning sets [14,15].

It is evident that when microarray probe level expression data were used in the process of model building, the predictive accuracy was greatly enhanced as compared to instances when average gene expression values utilized [16]. Similarly, the precision was increased when a variety of cancer types were used to build upon the models. Lastly, the predictive accuracy was considerably lowered after learning datasets which are built using preconceived biological models were pre-filtered [17]. The researcher proved that the predictive accuracy can be increased by developing SVM-based models across diverse types of cancer.

Choosing a suitable learning datasets to build the predictive models of cancer drug response is difficult since scientists or researchers have not fully described the molecular processes which cause cancer progression or onset [18]. Therefore, it is possible that a gene expression pattern that is linked to a certain cancer type might be causative factor of cancer development [19]. The researchers made comparison of two SVM-derived models developed to make prediction of the cancer drug carboplatin which is usually extensively prescribed in the market to determine their relative accuracies [20]. Drug response profiles and gene expression profiles were used to build the respective models. There has been limited use of open source algorithms in performing cancer drug prediction due to the lack of Bio conductor, Source forge, and Git Hub among other online repositories of prediction software as compared to the resources which are accessible for alternative machine learning applications, for instance, the Large Online Image (LOI) repository competitions which facilitate depositing of other computational solutions [21,22]. The researchers are confident that the open source support vector machine (SVM)-based algorithm can be valuable towards improving the personalized cancer medicine and

cancer drug prediction [21]. The algorithm combines a "one-by one" data normalization pipeline with a standard SVM methodology which is then implemented in seven commonly used chemotherapeutic drugs [24].

The drug response had high levels of predictive accuracy when the model was constructed using data from diverse cancer types [25]. The study outcome corroborates the current available evidence that argue that optimum cancer drug response's molecular signatures are not essentially defined by the tissue of origin of the cancer type [26]. Furthermore, the results prove that considerable enhancements can possibly be done in machine learning-based algorithms to improve their predictive accuracy especially through modulating the learning dataset's type or format as utilized in the model building process [27].

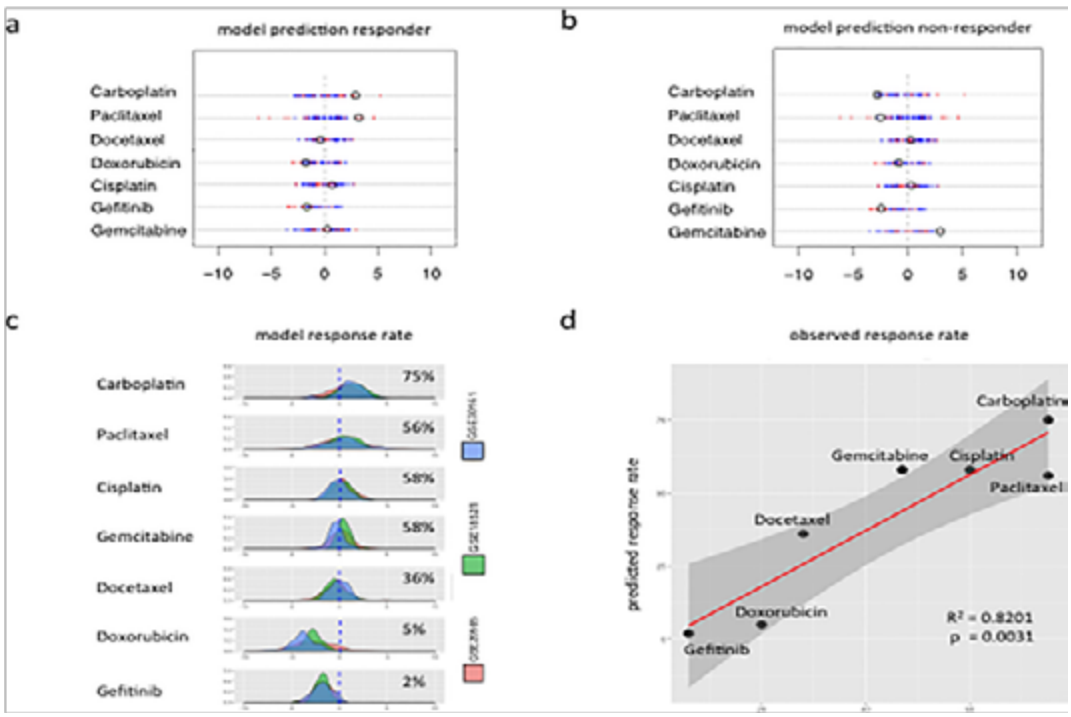

Figure 1: Aggregate and Individual prediction of response to chemotherapeutic drugs (Source: Huang, Cai, Roman Mezencev, John F. McDonald, Fredrik Vannberg (2017) Open source machine-learning algorithms for the prediction of optimal cancer drug therapies. PloS one 12: 1-18).

References

1. Huang C, Mezencev R, McDonald JF, Vannberg F (2017) Open Source Machine-Learning Algorithms for The Prediction of Optimal Cancer Drug Therapies. PloS one 12.
2. Costello JC, Heiser LM, Georgii E, Gonen M, Menden MP, et al. (2014) A Community Effort to Assess and Improve Drug Sensitivity Prediction Algorithms. Nat Biotechnol 32: 1202-1212.
3. Haverty PM, Lin E, Tan J, Yu Y, Lam B et al. (2016) Reproducible Pharmacogenomic Profiling of Cancer Cell Line Panels. Nature 533: 333-337.
4. Haibe-Kains B, El-Hachem N, Birkbak NJ, Jin AC, Beck AH, et al. (2013) Inconsistency in Large Pharmacogenomic Studies. Nature 504: 389-393.
5. Geeleher P, Cox N, Huang RS (2014) pRRophetic: An R Package for Prediction of Clinical Chemotherapeutic Response from Tumor Gene Expression Levels. PLoS One 9: e107468.
6. Piccolo SR, Withers MR, Francis OE, Bild AH, Johnson EW (2013) Multiplatform Single-Sample Estimates of Transcriptional Activation. Proceedings of the National Academy of Sciences 110: 17778-17783.
7. Andre E, Brett K, Novoa RA, Ko J, Swetter SM, et al. (2017) Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. Nature 542: 115-118.
8. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, et al. (2015) Imagenet Large Scale Visual Recognition Challenge. International Journal of Computer Vision 3: 211-252.
9. Clough Emily, Tanya Barrett (2016) The Gene Expression Omnibus Database. Statistical Genomics 1418: 93-110.
10. Liu B, Zhang D, Xu R, Xu J, Wang X, et al. (2013) Combining Evolutionary Information Extracted from Frequency Profiles with Sequence-Based Kernels for Protein Remote Homology Detection. Bioinformatics 30: 472-479.
11. Chih-Wei H, Chih-Chung C, Chih-Jen L (2003) A Practical Guide to Support Vector Classification. Bioinformatics 7: 1-16.
12. Yvan S, Inza I, Larranaga P (2007) A Review of Feature Selection Techniques in Bioinformatics. Bioinformatics 19: 2507-2517.
13. Radovic M, Ghalwash M, Filipovic N, Obradovic Z (2017) Minimum Redundancy Maximum Relevance Feature Selection Approach for Temporal Gene Expression Data. BMC Bioinformatics 18: 9.

14. Gaul DA, Mezencev R, Long TQ, Jones CM, Benedict BB, et al. (2015) Highly-Accurate Metabolomic Detection of Early-Stage Ovarian Cancer. Scientific Reports 5: 16351.
15. Guan W, Zhou M, Hampton CY, Benigno BB, Walker LD, et al. (2009) Ovarian Cancer Detection from Metabolomic Liquid Chromatography/ Mass Spectrometry Data by Support Vector Machines. BMC Bioinformatics 10: 259.
16. Wen Y, Gorsic LK, Wheeler HE, Ziliak DM, Huang RS, et al. (2011) Chemotherapeutic-Induced Apoptosis – A Phenotype for Pharmacogenomics Studies. Pharmacogenet Genomic 21: 476-488.
17. International Cancer Genome Consortium, Hudson TJ, Anderson W, Artez A, Barker AD (2010) International Network of Cancer Genome Projects. Nature 464: 993-998.
18. Swathi M (2018) Enhancement of K-Mean Clustering for Genomics of Drugs. Enliven: J Genet Mol Cell Biol 5: 001.
19. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, et al. (2014) Multiplatform Analysis of 12 Cancer Types Reveals Molecular Classification Within and Across Tissues of Origin. Cell 158: 929-944.
20. Azuaje F (2017) Computational Models for Predicting Drug Responses in Cancer Research. Brief bioinform 18: 820-829.
21. Frampton GM, Fichtenholtz A, Otto GA, Wang K, Downing SR, He J, et al. (2013) Development and Validation of a Clinical Cancer Genomic Profiling Test Based On Massively Parallel DNA Sequencing. Nat Biotechnol 31: 1023-1031.
22. Ozols RF (2006) Challenges for Chemotherapy in Ovarian Cancer. Annals of Oncology 17: 181-187.
23. Hansen SW (2001) Gemcitabine in the Treatment of Ovarian Cancer. Int J Gynecol Cancer 2: 39-41.
24. Bode AM, Dong Z (2017) Precision Oncology – The Future of Personalized Cancer Medicine?. NPJ Precis Oncol 1: 2.
25. Stephanie S, Verfaillie CM (2002) BCR/ABL: From Molecular Mechanisms of Leukemia Induction to Treatment of Chronic Myelogenous Leukemia. Oncogene 21: 8547-8549.
26. Swathi M (2017) Clustering Enhancement Using Similarity Indexing to Reduce Entropy. Enliven: Bioinform 4: 001.
27. Saez-Rodriguez J, Costello JC, Friend SH, Kellen MR, Mangravite L, et al. (2016) Crowdsourcing Biomedical Research: Leveraging Communities as Innovation Engines. Nature Reviews Genetics 17: 470-486.
28. Gysels E, Renevey P, Celka P (2005) SVM-Based Recursive Feature Elimination to Compare Phase Synchronization Computed from Broadband and Narrowband EEG Signals in Brain–Computer Interfaces. Signal Processing 85: 2178-2189.
29. Liu T, Tao P, Li X, Qin Y, Wang C (2015) Prediction of Subcellular Location of Apoptosis Proteins Combining Tri-Gram Encoding Based On PSSM and Recursive Feature Elimination. J Theor Biol 366: 8-12.
30. Swathi M (2017) Drug Prediction of Cancer Genes Using SVM. Enliven: Pharmacovigilance and Drug Safety 4: 001.