Krajowa Konferencja Zastosowań Matematyki w Biologii i Medycynie

Jugowice, 11th–15th September 2017

RESPONSE PROFILES FOR HIGH AND THERAPEUTIC RADIATION DOSES IN BREAST CANCER PATIENTS

Anna Papież[1], Christophe Badie[2], Joanna Polańska[1]

[1]Data Mining Group, Silesian University of Technology
ul. Akademicka 16, 44-100 Gliwice, Poland
[2]Public Health England, Centre for Radiation, Chemical & Environmental Hazards
Chilton, Didcot, Oxfordshire OX11 ORQ United Kingdom
[1]apapiez@polsl.pl, jpolanska@polsl.pl, [2]christophe.badie@phe.gov.uk

ABSTRACT

Nowadays, there still remains a great amount of knowledge to discover regarding the molecular mechanisms underlying response to ionising radiation. One of the most common and important applications of ionising radiation is radiotherapy in cancer treatment and the need to improve appropriate dosimetry for constantly developing personalised medicine is pressing. In this work we conduct a combined analysis of two microarray experiments on dose response in breast cancer patients in terms of data mining for potential biomarkers and the selection of processes occurring under particular doses. Initially, a standard statistical inference approach is adopted, however the most important results are obtained in a custom analysis of dose response profiles using Jonckheere-Terpstra test for trends coupled with order restricted clustering techniques. Validation is performed *in silico* using the Gene Ontology and Reactome database. The results provide candidate gene biomarker lists for differences in high and therapeutic dose samples, as well as lists of processes identified as overrepresented GO terms, also for the distinction between radiosensitive and radioresitant patients. Such preliminary analysis may serve as a tool for efficient reduction of time and costs required to perform large screening biological experiments.

INTRODUCTION

Ionising radiation is an omnipresent factor, which has a significant impact on many aspects of human life. Small doses are absorbed on an everyday basis while using utensils such as radios or microwave ovens, whereas higher doses occurring during accidents may have very detrimental effects [1]. However, high doses used under controllable conditions have beneficial effects, i.e. they are used widely for therapeutic purposes. In fact, medical procedures such as X-ray imaging or radiotherapy constitute the main source of man-made radiation exposure [2]. For instance, 2 Gy of ionising radiation, which is classified as a high dose, is a commonly used fraction of a total dose given to the patient during radiation therapy in various types of cancer [3]. This is a standard, however, it is known that radiosensitivity is a trait specific for each individual and depending on its level the reaction to radiotherapy may be extremely different. Radiosensitive patient obtaining too high doses too frequently have a high chance of developing late adverse effects, while for radioresitant individuals the standard procedure may be insufficient for the healing effect to progress. Therefore, there is a pressing need for a thorough understanding of the processes underlying radiation response in the field of dosimetry, for the objective of therapy personalisation.

In this work we examine two datasets from breast cancer patient samples. In one experiment these blood samples were treated with a therapeutic dose of 2 Gy and in the other, with a high dose of 4Gy. By means of dose response profiles and overrepresentation *in silico* analysis the study provides guidance for some insight into the processes activated and inhibited along with a raise in high doses of radiation used for medical purposes.

MATERIALS AND METHODS

The data sets consisted of two independent expression sets obtained in the course of microarray experiments on radiosensitivity. The experiments were designed with the objective of identifying genes differentiating radioresistant and radiosensitive women in a group of breast cancer patients undergoing radiotherapy using extracted lymphocytes from donated blood samples.

The first experiment was carried out on the HuGene Affymetrix 1.0 ST oligonucleotide microarray platform. There were 60 samples, of which 30 labelled as radiosensitive and 30 as radioresistant. These samples were divided into two conditions: controls and irradiated with a therapeutic level dose of 2 Gy. The second was performed using a custom Breakthrough 20K cDNA microarray chip. The study group consisted of 31 radiosensitive and 28 radioresistant patients and the treatment samples were subjected to a high dose of 4 Gy.

Data sets were normalised separately according to their platform type as in [4]. Next, in order to enable joint analysis of the two sets, they were adjusted for batch effects using empirical Bayes methods [5]. Afterwards, statistical inference was performed using appropriate gene differentiation tests with regard to normality of the distribution and variance homogeneity.

For downstream analysis, up- and down-regulation testing was performed between the three dose groups and six types of dose response were considered (depicted in Figure 1). Moreover, the Jonckheere-Terpstra test for trend [6, 7] was utilised for the identification of genes with similar expression patterns. Genes were grouped as strictly up-regulated when the outcome of the trend test was significant for an increasing trend hypothesis and insignificant for the monotonic trend hypothesis and contrariwise for down-regulated features. Afterwards, the genes were further investigated using order restricted clustering [8], to determine which of them belong in similar dose response groups. The genes with common expression profiles were examined in terms of functional analysis to determine their Gene Ontology [9] interactions and Reactome pathways [12] with overrepresented terms. Furthermore, the features identified as differentiating specifically within radiosensitive and radioresistant patient groups were also studied for GO Biological Process overrepresentation with Storey's FDR procedure chosen for multiple testing correction [10]. As for the ontology algorithm, the *parent/child* method [11] was used, which takes into account the structure of the GO tree graph.

RESULTS

Differentiation testing

After extracting the common genes for both experimental platforms, the joint analysis was carried out on a total of 9852 genes. Firstly, these genes were investigated with regard to the control samples in order to ensure the same base level for both experimental datasets. For each gene, the appropriate statistical differentiation test was chosen taking into account the normality and variance homogeneity assumption fulfilment. Among the control samples 7429 genes did not produce a significant difference between the normalised oligonucleotide and cDNA array experiments.

Therefore, this genelist was further examined for features discriminating between high and therapeutic dose response. 1455 genes emerged as differentially expressed between the two doses, regardless of the radiosensitivity status. When taking into account radiosensitive (RS) and radioresistant (RR) samples separately, the overlap of differentially expressed genes is presented in Figure 2.

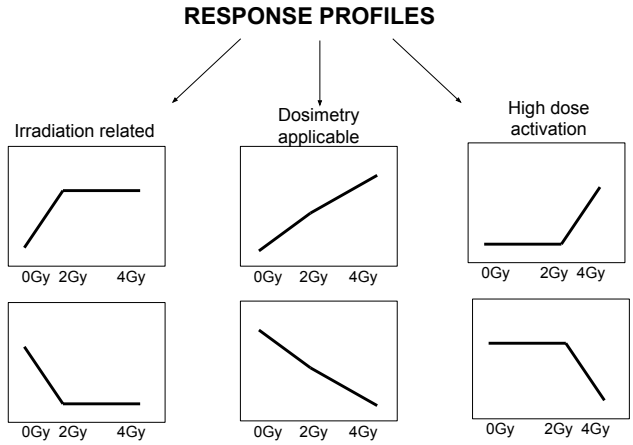

Figure 1. Diagram illustrating the types of response profiles under consideration in downstream analysis.

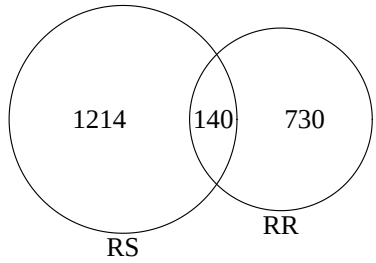

Figure 2. Venn diagram illustrating the overlap of differentially expressed genes for high and therapeutic doses separately in radiosensitive (RS) and radioresistant (RR) patients.

On performing Gene Ontology enrichment analysis using Fisher's exact test with regard to all of the differentially expressed genes, after multiple testing Benjamini-Hochberg correction procedure, we obtained 1 significantly overrepresented term, i.e.: cellular amino acid metabolic process. This GO term is of a general nature, therefore, ontology overrepresentation analysis was carried out using genes differentially expressed between doses in radiosensitive and radioresistant groups of samples. Within the radiosensitive group no significantly overrepresented terms were discovered, however, in the radioresistant patients 31 terms were statistically significant and among those radiation response processes may be found such as:

- response to stress;
- oxidative phosphorylation;
- immune response-regulating signalling pathway.

Trend testing

The analysed genes, which did not present significant differences between controls in both experiments, were classified into one of the six types of response profiles (Figure 1). The number of genes allocated in each of the examined groups is presented in Table 1. When analysing genes

falling into the response profile types, the particular interest lays in the dosimetry applicable group, thus these genes were examined for related pathways using the Reactome Pathway Browser. A total of 34 pathways were identified as deregulated, among those DNA repair related processes, such as the Gap-filling DNA repair synthesis and ligation in GG-NER 3 and regulation of RAS by GAPs [14], as well as a number of other terms concordant with signal transduction, immune repsonse, developmental biology, DNA replication and repair.

Table 1. Numbers of genes presenting the same types of dose response profiles.

Number of genes in response profiles					
Irradiation related		Dosimetry applicable		High dose activation	
Up-No change	610	Up-Up	117	No change-Up	48
Down-No change	1067	Down-Down	969	No change-Down	319

Then, the genes were additionally assessed for the presence of trend with the Jonckheere-Terpstra test. The number of features with increasing, decreasing and monotonic trends is presented in Table 2.

Table 2. Numbers of genes showing significant dose trend. The strictly increasing and decreasing genes are those which do not appear in the monotonic trend group.

	Increasing	Monotonic	Decreasing
N^o of genes	717	377	53
	Strictly increasing		Strictly decreasing
N^o of genes	363		30

The genes with strictly increasing and decreasing dose response served as a basis for GO term enrichment analysis. The strictly up-trending genes yielded 99 significantly overrepresented terms and the down-trending: 38 GO terms. Among the terms linked to decreasing features, there were processes related to hemopoiesis and homoeostasis, as well as GPI anchor metabolism and biosynthesis. Whereas, among the terms enriched with dose-increasing genes, cellular response to ionizing radiation could be found, along with processes such as tube morphogenesis and Wnt signalling.

DISCUSSION

Starting from standard statistical inference analysis between high and therapeutic doses of ionising radiation, the results show that differentiating genes in radioresistant and radiosensitive groups participate in key processes related to stress response under irradiation. However, it is insufficient to recognise differences between these two discrete dose points, and therefore, dose response profiles were put under investigation, for a more profound knowledge of dose response mechanisms.

Thus, the response profiles were submitted under investigation, and especially the up- and down-regulation across the measured doses, as these present an opportunity for dosimetric applications. The functional validation through pathway analysis using Reactome indicates links to DNA repair and oncogenic mechanisms. A great majority of these genes are down-regulated, which points out toward the inhibition of the underlying processes.

The results of Jonckheere-Terpstra testing show that a majority of the genes presenting significant trends are in the increasing direction, nevertheless, in both cases significantly overrepresented GO terms are revealed. They account, however, for different types of processes related to radiation response. The down-trending genes tend to be more engaged in hemopoiesis and homoeostasis, which have been previously shown to play a role in stem cell injury from ionizing radiation [13].

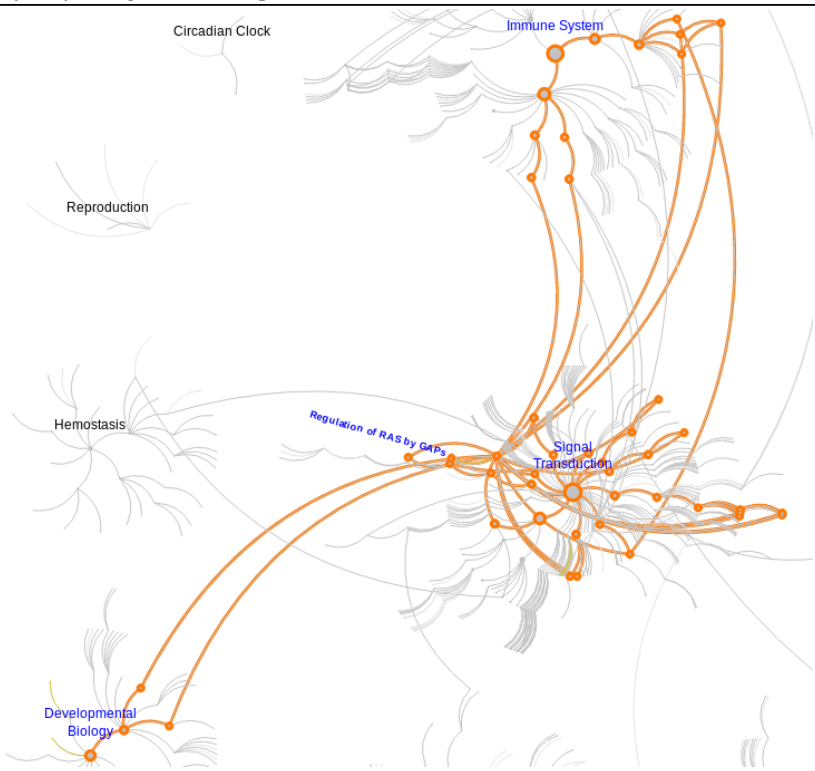

Figure 3. Reactome graph presenting pathways deregulated by dosimetry applicable genes linked to Regulation of RAS by GAPs. The deregulated pathways are highlighted in bold.

Also, GPI anchors being important apoptosis regulators when deregulated by ionising radiation may have a significant impact on cellular resistance [15]. On the other hand, up-trending genes have proved to be explicitly involved in cellular response to ionizing radiation, but also indirectly such indications have been found e.g. in the Wnt signaling pathway or tube morphogenesis. The former is and the latter has been reported to be linked to breast cancer mechanisms in a study comprising a large dataset analysed in a non-customary manner going beyond differential expression [16].

CONCLUSIONS

In order to enable personalised individual therapy, the underlying mechanisms of disease need to be profoundly studied and understood. In case of treating breast cancer using radiotherapy the uniqueness of response to exposition to an ionising radiation dose in terms of its power and frequency emphasises the urge to complete this task in the nearest future. In this work, it has been shown that limiting the efforts necessary to execute by specialists in molecular biology is possible by adapting statistical approaches to a certain biological question. In this case analysis of dose profiles using Jonckheere-Terpstra trend testing and order restricted clustering points out to particular potential biomarkers and signalling pathways and processes involved in cellular response to high and therapeutic doses of ionising radiation. Such *in silico* analysis with functional validation narrows down the search space for expert, potentially saving time and effort and allowing for improvement in planing the design of future biological experiments established in order to study the impact of specific doses on breast cancer radiotherapy.

ACKNOWLEDGEMENTS

This work was financed with the aid of the SUT grant BKM/506/RAU1/2016/t.30 (AP) and BK/204/RAU1/2017/9 (JP). Calculations were carried out using GeCONiI infrastructure (POIG.02.03.01-24-099/13).

REFERENCES

[1] Alison Abbott: *Researchers pin down risks of low-dose radiation: large study of nuclear workers shows that even tiny doses slightly boost risk of leukaemia*, Nature **523** (2015), 17–19.

[2] Monika Ray et al.: *Comparison of low and high dose ionising radiation using topological analysis of gene coexpression networks*, BMC genomics **13** (2012), 190.

[3] Michael C. Joiner: *A simple α/β-independent method to derive fully isoeffective schedules following changes in dose per fraction*, International Journal of Radiation Oncology* Biology* Physics **58** (2004), 871–875.

[4] Anna Papiez et al.: *Integrating Expression Data from Different Microarray Platforms in Search of Biomarkers of Radiosensitivity.*, IWBBIO, 2014, pp. 484–493.

[5] W. Evan Johnson et al.: *Adjusting batch effects in microarray expression data using empirical Bayes methods*, Biostatistics **8** (2007), 118–127.

[6] Teunis J. Terpstra: *The asymptotic normality and consistency of Kendall's test against trend, when ties arepresent in one ranking*, Proc. Kon. Ned. Akad. v. Wetensch. A **55** (1952), 327–333.

[7] Aimable R. Jonckheere: *A distribution-free k-sample test against ordered alternatives*, Biometrika **41** (1954), 133–145.

[8] Dan Lin et al.: *Modeling dose-response microarray data in early drug development experiments using R: order-restricted analysis of microarray data*, Springer Science & Business Media, 2012.

[9] Michael Ashburner et al.: *Gene Ontology: tool for the unification of biology*, Nature genetics **25** (2000), 25–29.

[10] Alan Dabney et al.: *qvalue: Q-value estimation for false discovery rate control*, R package version **1** (2010).

[11] Steffen Grossmann et al.: *Improved detection of overrepresentation of Gene-Ontology annotations with parent–child analysis*, Bioinformatics **23** (2007), 3024–3031.

[12] David Croft et al.: *The Reactome pathway knowledgebase*, Nucleic acids research **42** (2014), D472–D477.

[13] Lijian Shao et al.: *Hematopoietic stem cell injury induced by ionizing radiation*, Antioxidants & redox signaling **20** (2014), 1447–1462.

[14] Thomas Grewal et al.: *Differential regulation of RasGAPs in cancer*, Genes & cancer **2** (2011), 288–297.

[15] Robert A. Brodsky et al.: *Resistance to apoptosis caused by PIG-A gene mutations in paroxysmal nocturnal hemoglobinuria*, Proceedings of the National Academy of Sciences **94** (1997), 8756–8760.

[16] Patrick R. Schmid et al.: *Making sense out of massive data by going beyond differential expression*, Proceedings of the National Academy of Sciences **109** (2012), 5594–5599.

[17] Stephan M. Huber et al.: *Ionizing radiation, ion transports, and radioresistance of cancer cells*, Frontiers in physiology **4** (2013), 212.