



# Editorial: Next Generation Sequencing Based Diagnostic Approaches in Clinical Oncology

Anton Buzdin<sup>1,2,3,4\*</sup>, Ira Ida Skvortsova<sup>5,6,7</sup>, Xinmin Li<sup>8</sup> and Ye Wang<sup>9</sup>

<sup>1</sup> Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia, <sup>2</sup> World-Class Research Center “Digital Biodesign and Personalized Healthcare”, Sechenov First Moscow State Medical University, Moscow, Russia, <sup>3</sup> Translational Genome Bioinformatics Laboratory, Moscow Institute of Physics and Technology (National Research University), Moscow, Russia, <sup>4</sup> Research Department, OmicsWay Corp., Walnut, CA, United States, <sup>5</sup> Therapeutic Radiology and Oncology, Medical University of Innsbruck, Innsbruck, Austria, <sup>6</sup> Group for Experimental and Translational Radiooncology, Tyrolean Cancer Research Institute, Innsbruck, Austria, <sup>7</sup> PathoBiology Group, European Organization for Research and Treatment of Cancer (EORTC), Brussels, Belgium, <sup>8</sup> Department of Pathology & Laboratory Medicine, University of California Los Angeles (UCLA) Technology Center for Genomics & Bioinformatics, Los Angeles, CA, United States, <sup>9</sup> Clinical Laboratory, Qingdao Central Hospital, The Second Affiliated Hospital of Medical College of Qingdao University, Qingdao, China

**Keywords:** cancer diagnostics, genomics, transcriptomics, bioinformatics, biomarkers, theranostics, next generation sequencing, RNA sequencing

## OPEN ACCESS

### Edited and reviewed by:

Claudio Sette,  
Catholic University of the Sacred  
Heart, Rome, Italy

### \*Correspondence:

Anton Buzdin  
buzdin@oncobox.com

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Oncology

**Received:** 30 November 2020

**Accepted:** 14 December 2020

**Published:** 28 January 2021

### Citation:

Buzdin A, Skvortsova I, Li X and  
Wang Y (2021) Editorial: Next  
Generation Sequencing Based  
Diagnostic Approaches  
in Clinical Oncology.  
*Front. Oncol.* 10:635555.  
doi: 10.3389/fonc.2020.635555

## Editorial on the Research Topic

### Next Generation Sequencing Based Diagnostic Approaches in Clinical Oncology

Next Generation Sequencing (NGS) technologies transformed cancer genetics by providing unprecedented access to big genomic and transcriptomic data (1, 2). Curation of cancer genetic profiles has led to many successful applications crosslinking mutations or polymorphisms with individual tumor response to therapies. Obviously valuable, these approaches however still cannot generate clinically actionable information for most of the cancer patients. In turn, high throughput RNA (transcriptome) analysis can be considered a rising star that may complement mutational screenings (1, 3). A combination of both approaches may be synergistic for many tasks in molecular diagnostics in oncology (4–6). Cancer mutation and expression biomarkers can not only help to set a diagnosis but also to identify appropriate personalized molecular-based treatment (7–9). Furthermore, novel bioinformatic approaches enable squeezing more and more clinically meaningful data from large genetic datasets (10).

Wang et al. reviewed current progress in using RNA sequencing (RNAseq) methods in cancer research, prognosis, and molecular diagnostics. The techniques of RNAseq have progressed rapidly from bulk RNAseq, laser-captured micro-dissected (LCM) RNAseq, and single-cell (SS) RNAseq to digital spatial RNA profiling, spatial transcriptomics, and direct *in situ* sequencing. These different technologies have their unique strengths, weaknesses, and suitable applications in the field of clinical oncology. For example, bulk RNAseq is a cost-effective mature technology that can be used for confident whole-transcriptome biomarker recovery and targeted capture of gene fusions. LCM-RNAseq interrogates tissue heterogeneity by dissecting cell type specific populations, whereas SS-RNAseq can characterize different cell types inhabiting and surrounding tumors. Zhigalova et al.

applied RNAseq to characterize T-cell receptor (TCR) and immunoglobulin repertoire in an HKP1 *Kras*<sup>G12D</sup>*p53*<sup>-/-</sup> syngeneic mouse model of lung cancer after anti-PD-1 treatment, and observed decreased TCR diversity in response to therapy. However, repertoire diversity was then restored in progressing disease but remained decreased in good responders to therapy in both CD4+ and CD8+ subsets, thus suggesting their potential usefulness as prognostic biomarkers.

Borger et al. used RNAseq and bioinformatic analysis with Oncofinder (11) for characterization of molecular processes accompanying ALPPS, a novel two-staged hepatectomy that dramatically accelerates liver regeneration and enables extensive liver tumor resection. ALPPS specific signature included activation of cell survival branch of IGF1R signaling pathway, proliferation branch of ILK Pathway, and the IL-10 Pathway, whereas the transcriptional branch of the Interferon pathway was downregulated ( $p < 0.05$ ). The PAK- and ILK-associated pathways were also activated at an earlier time point, reflecting acceleration of liver regeneration ( $p < 0.001$ ). Molecular therapies influencing those pathways could potentially improve the performance of ALPPS in the future.

Poddubskaya et al. used bulk transcriptomics to identify effective experimental drug combination for an advanced recurrent *ALK*-positive lung cancer patient using Oncobox algorithm (12) based on ranking of tumor-upregulated drug target genes. A 48 y.o. male patient received personalized off-label combination of crizotinib + bevacizumab + docetaxel that led to additional stabilization for 22 months. The patient survival after developing resistance to *ALK* inhibitor was longer for 16 months than previously reported average survival for such cases, thus evidencing effectiveness of RNA-guided prescription of cancer therapies in terms of survival and quality of life.

By experimentally *in vitro* validating meta-analysis of RNAseq data for esophageal squamous cell carcinoma (ESCC), Tang et al. identified *LINC01614* non-coding (NC) RNA that was a poor prognosis biomarker, being upregulated in ESCCs. The knockdown of *LINC01614* expression significantly inhibited the migration of ESCC cells by restricting EMT. He et al. found that another NC RNA, circular RNA *hsa\_circ\_0007843*, in colon cancer SW480 cells can serve as a molecular sponge by interacting with the microRNA miR-518c-5p that interacts with matrix metalloproteinase 2 (MMP2). Overexpression of *hsa\_circ\_0007843* promoted tumor cell growth, invasion, and migration, whereas its downregulation had opposite effects. The possible explanation for those effects is removing the miR-518c-5p inhibitory block on MMP2 translation.

Finally, Borisov and Buzdin concentrate on the phenomenon that machine learning (ML) methods that had obviously revolutionized many fields are still poorly applicable to molecular oncology. Since plausible explanation is the small size of related datasets, the authors consider using dynamic data trimming (13) to filter for more relevant and informative feature sets to apply ML for finding robust transcriptomic biomarkers using available clinically annotated datasets, e.g. (14).

Conversely, Matrone et al. reported application of cancer DNA sequencing to identify the primary origin of metastases in

the patient's thyroid gland in case of concomitant lung adenocarcinoma. The patient was diagnosed with both poorly differentiated thyroid cancer for the huge involvement of the neck and concomitant lung adenocarcinoma. Clinical features, imaging evaluation and available tumor markers couldn't support a well-defined diagnosis, and the histologic features of the thyroid and lung biopsies confirmed the figure of two different tumors. However, the NGS analysis showed a G12C mutation in *KRAS* gene in both tissues, which is highly prevalent in lung but not thyroid cancers. Therefore, the lung origin of the disease was deduced, and the patient was addressed to the appropriate therapeutic strategy.

Yuan et al. explored *ERBB2* exon 20 insertion that is a refractory oncogenic driver, by panel NGS of 59 or 1,021 genes, in 112 lung cancer patients, including 18 patients receiving afatinib treatment. There were 66% of patients receiving *TP53* co-mutation, and *FOXA1* was the most prevalent co-amplified gene (in 5.5% of the cases). Patients with co-occurring *TP53* mutation showed approximately twice shorter overall survival (OS): median OS of 14.5 versus 30.3 months,  $p = 0.04$ . *ERBB2* exon 20 insertion also related to shorter progression-free survival (PFS): median PFS of 1.2 versus 4.3 months,  $p < 0.05$ .

In a subgroup of 44 cases with *ALK* gene fusions out of total 1349 lung cancer patients, Liu et al. identified *ALK* fusion partners using a customized NGS panel. The most common partner was well-known gene *EML4*, but also a new *ALK* fusion partner *HMBOX1* was identified. The copy number alterations were found in ~30% of the cases, and the most commonly amplified genes were *MDM2* and *TERT*.

Furthermore, Belardinilli et al. used target gene panel sequencing to characterize 639 formalin-fixed paraffin-embedded (FFPE) metastatic colorectal cancer (mCRC) specimens and identified blocks of significantly cooccurring mutations. This led to a novel stratification of mCRC patients in eight groups characterized by specific mutational association patterns, which was validated on a literature dataset of The Cancer Genome Atlas (TCGA) project.

By performing whole-genome sequencing, Gao et al. investigated genetic alterations in spinal schwannoma and identified mutations in *ATM*, *CHD4*, *FAT1*, *KMT2D*, *MED12*, *NF2*, and *SUFU* genes, and homozygous deletion was observed in *NF1*, *NF2*, and *CDKN2C*. Hippo signaling pathway was most significantly affected by the mutations identified. Shahid et al. used 54-gene panel for target DNA sequencing of 26 acute myeloid leukemia (AML) samples by using ultra-deep NGS with ~5,000-coverage. Novel somatic mutations were identified, including those repeated in several genetically unrelated cancers, i.e. *STAG2* L526F and *BCORL1* A400V.

Of note, Nikitin et al. showed that Lynch Syndrome-related germline mutations (LS-mutations) in DNA mismatch repair pathway genes are statistically significantly associated with breast cancer (BC). To this end, they performed Targeted NGS of genes *MLH1*, *MSH2*, *MSH6*, *EPCAM*, and *PMS2* in a cohort of 492 healthy donors, 711 patients with hereditary BC, and 60 patients with sporadic BC. ~10% patients with hereditary BC had at least

one germline mutation, and ~5% had predicted pathogenic mutations in these genes.

Gao et al. explored in detail the influence of sample preparation method on the results of target gene panel NGS. A 22-gene panel with 103 hotspots was used to detect mutations in paired FFPE and fresh frozen tissue specimens from 118 patients with colorectal cancer. They found that 99% of the patients one or more detectable variants, with 226 variants in FFPE and 221 in fresh tissue. Of the totally 129 individual variants identified, 96 variants were common for both FFPE and fresh biosamples, 27 were specific for FFPE, and 6 for fresh tissue specimens. These findings suggest that when there is a choice, fresh frozen specimens should be most probably considered as the preferred type of biosamples because of a number (27/129) of apparently artifact variants identified specific for the FFPE tissue samples.

Another technological aspect of NGS was investigated by Petrackova et al. who used statistical tests to assess the sequencing coverage thresholds for a robust variant calling, and published a coverage calculator. Using the sequencing error only, the recommended minimum depth of coverage was 1,650 along with a threshold of no less than 30 variant reads for a targeted clinical NGS mutation analysis. This points to an important problem of data quality standardization in clinical

NGS. However, in practice this threshold can be only applied to target panel gene sequencing methods, whereas it is so far unrealistic for the whole-genome or even exome sequencing approaches.

Finally, epigenetic aspects of using NGS in oncology were considered by Jovčevska who reviewed current progress in sequencing and machine learning applications for the analysis of DNA methylation in glioblastoma.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## FUNDING

AB contribution was supported by the Russian Science Foundation grant 18-15-00061. YW was sponsored by the National Natural Science Foundation of China (No. 81670822), and Qingdao Key Research Project (No. 19-6-1-3-nsh).

## REFERENCES

- Buzdin A, Sorokin M, Garazha A, Glusker A, Aleshin A, Poddubskaya E, et al. RNA sequencing for research and diagnostics in clinical oncology. *Semin Cancer Biol* (2020) 60:311–23. doi: 10.1016/j.semcancer.2019.07.010
- Buzdin A, Sorokin M, Garazha A, Sekacheva M, Kim E, Zhukov N, et al. Molecular pathway activation – New type of biomarkers for tumor morphology and personalized selection of target drugs. *Semin Cancer Biol* (2018) 53:110–24. doi: 10.1016/j.semcancer.2018.06.003
- Rodon J, Soria JC, Berger R, Miller WH, Rubin E, Kugel A, et al. Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat Med* (2019) 25:751–8. doi: 10.1038/s41591-019-0424-4
- Negro G, Aschenbrenner B, Brezar SK, Cemazar M, Coer A, Gasljevic G, et al. Molecular heterogeneity in breast carcinoma cells with increased invasive capacities. *Radiol Oncol* (2020) 54:103–18. doi: 10.2478/raon-2020-0007
- Skvortsova I. Special Issue “Enigmatic tumor properties associated with metastatic spread” seminars in cancer biology, volume XX. *Semin Cancer Biol* (2020) 60:iii–iv. doi: 10.1016/j.semcancer.2019.10.021
- Skvortsova II, Span PN. Editorial: Advances in Biological Understanding of Tumor Radiation Resistance. *Front Oncol* (2020) 10:754. doi: 10.3389/fonc.2020.00754
- Suntsova M, Gaifullin N, Allina D, Reshetun A, Li X, Mendeleeva L, et al. Atlas of RNA sequencing profiles for normal human tissues. *Sci Data* (2019) 6:36. doi: 10.1038/s41597-019-0043-4
- Tsimberidou AM, Fountzilias E, Nikanjam M, Kurzrock R. Review of precision cancer medicine: Evolution of the treatment paradigm. *Cancer Treat Rev* (2020) 86(102019):1–11. doi: 10.1016/j.ctrv.2020.102019
- Moisseev A, Albert E, Lubarsky D, Schroeder D, Clark J. Transcriptomic and genomic testing to guide individualized treatment in chemoresistant gastric cancer case. *Biomedicines* (2020) 8(3):67. doi: 10.3390/biomedicines8030067
- Buzdin A, Sorokin M, Poddubskaya E, Borisov N. High-Throughput Mutation Data Now Complement Transcriptomic Profiling: Advances in Molecular Pathway Activation Analysis Approach in Cancer Biology. *Cancer Inform* (2019) 18:1176935119838844. doi: 10.1177/1176935119838844.
- Buzdin AA, Zhavoronkov AA, Korzinkin MB, Venkova LS, Zenin AA, Smirnov PY, et al. Oncofinder, a new method for the analysis of intracellular signaling pathway activation using transcriptomic data. *Front Genet* (2014) 5:55:55. doi: 10.3389/fgene.2014.00055
- Tkachev V, Sorokin M, Garazha A, Borisov N, Buzdin A. Oncobox Method for Scoring Efficiencies of Anticancer Drugs Based on Gene Expression Data. In: *Methods in molecular biology (Clifton, N.J.)*. (2020) p. 235–55. doi: 10.1007/978-1-0716-0138-9\_17
- Tkachev V, Sorokin M, Borisov C, Garazha A, Buzdin A, Borisov N. Flexible data trimming improves performance of global machine learning methods in omics- based personalized oncology. *Int J Mol Sci* (2020) 21(3):713. doi: 10.3390/ijms21030713
- Borisov N, Sorokin M, Tkachev V, Garazha A, Buzdin A. Cancer gene expression profiles associated with clinical outcomes to chemotherapy treatments. *BMC Med Genomics* (2020) 13:111. doi: 10.1186/s12920-020-00759-0

**Conflict of Interest:** Author AB was employed by the company OmicsWay Corp.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Buzdin, Skvortsova, Li and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.