BMC Cancer

RESEARCH

Open Access

Personalized targeted therapy prescription in colorectal cancer using algorithmic analysis of RNA sequencing data

Maxim Sorokin[1,2,3], Marianna Zolotovskaia[2], Daniil Nikitin[3,4], Maria Suntsova[5], Elena Poddubskaya[1,6], Alexander Glusker[1], Andrew Garazha[3], Alexey Moisseev[1], Xinmin Li[7*], Marina Sekacheva[5], David Naskhletashvili[8], Alexander Seryakov[9], Ye Wang[10*] and Anton Buzdin[2,4,5*]

Abstract

Background: Overall survival of advanced colorectal cancer (CRC) patients remains poor, and gene expression analysis could potentially complement detection of clinically relevant mutations to personalize CRC treatments.

Methods: We performed RNA sequencing of formalin-fixed, paraffin-embedded (FFPE) cancer tissue samples of 23 CRC patients and interpreted the data obtained using bioinformatic method Oncobox for expression-based rating of targeted therapeutics. Oncobox ranks cancer drugs according to the efficiency score calculated using target genes expression and molecular pathway activation data. The patients had primary and metastatic CRC with metastases in liver, peritoneum, brain, adrenal gland, lymph nodes and ovary. Two patients had mutations in *NRAS*, seven others had mutated *KRAS* gene. Patients were treated by aflibercept, bevacizumab, bortezomib, cabozantinib, cetuximab, crizotinib, denosumab, panitumumab and regorafenib as monotherapy or in combination with chemotherapy, and information on the success of totally 39 lines of therapy was collected.

Results: Oncobox drug efficiency score was effective biomarker that could predict treatment outcomes in the experimental cohort (AUC 0.77 for all lines of therapy and 0.91 for the first line after tumor sampling). Separately for bevacizumab, it was effective in the experimental cohort (AUC 0.87) and in 3 independent literature CRC datasets, n = 107 (AUC 0.84–0.94). It also predicted progression-free survival in univariate (Hazard ratio 0.14) and multivariate (Hazard ratio 0.066) analyses. Difference in AUC scores evidences importance of using recent biosamples for the prediction quality.

Conclusion: Our results suggest that RNA sequencing analysis of tumor FFPE materials may be helpful for personalizing prescriptions of targeted therapeutics in CRC.

*Correspondence:
Xinmin Li
XinminLi@mednet.ucla.edu
Ye Wang
yewang@qdu.edu.cn
Anton Buzdin
buzdin@oncobox.com

Full list of author information is available at the end of the article

Background

Colorectal cancer (CRC) is globally the fourth most common cancer with approximately 1.8 million new cases diagnosed in 2018 [1]. CRC overall survival was increasing slowly during two last decades [2] and remains now at the level of ~3–5 years [3, 4] with the major factors being patient age [5] and tumor stage [6]. Several targeted cancer drugs with different molecular specificities were approved for the treatment of CRC, such as bevacizumab [7], aflibercept [8], regorafenib [9], cetuximab [10], and panitumumab [11]. New potential therapeutic molecules that act as the inhibitors of MEK, MET, RAS, RAF and PD-1/PDL-1 proteins are currently undergoing clinical trials in CRC (reviewed in [12]). Current chemotherapeutic drug prescription in CRC is generally based on the results of clinical imaging (computer tomography, magnetic resonance, positron emission tomography) and on histopathological analysis [13].

However, molecular guidance for treatment of metastatic disease was also informative and successful in several CRC subtypes. For example, activating mutations in *KRAS*, *NRAS* and *BRAF* genes are used as negative predictors for EGFR-targeted therapies [14], and microsatellite instability is regarded as an indication for immunotherapy and platinum drugs prescription [15]. In turn, in *BRAF* mutant cases combined targeting of BRAF, EGFR, and MEK by the specific therapeutics can be effective [16]. HER2-specific monoclonal antibodies and low-molecular mass tyrosine kinase inhibitors are the option to treat patients with *HER2*-amplified and wild-type *KRAS/NRAS* tumors [17, 18]. Taken together, the abovementioned genetic alterations cover approximately 50–60% of all CRC cases, leaving aside the rest 40–50% of patients [19–23]. Thus, more molecular biomarkers are needed for guiding CRC treatment, especially for the cases when the current diagnostic mutations can't identify therapeutics or when patients don't respond on the standardly recommended medicines/lines of treatment.

Gene expression profiles of tumor tissues can be considered as promising emerging biomarkers for cancer molecular diagnostics and personalized prescription of targeted drugs [24]. Both single gene expression profiles [25], differential gene sets/diagnostic signatures [26], and molecular pathway activation levels [27] may be useful to predict treatment outcomes for the individual tumors with regard to any associated cancer drug or treatment regimen. RNA sequencing is currently recognized as the method of choice for high throughput assessment of gene expression [28]. Further comparison of cancer gene expression levels with the normal tissues is important for delineating specific mechanisms of cancer onset, progression and responsiveness to clinical interventions [29, 30].

However, RNA sequencing profiles can be obtained using different equipment, reagents and protocols and sometimes can be poorly compatible thus making their direct comparison problematic [31, 32]. In order to obtain comparable results, ideally the same experimental platform should be used for all biosamples in a given study [32]. We recently published a collection of RNA sequencing profiles of human healthy tissues called Atlas of Normal Tissue Expression (ANTE) [33]. It contains 159 tissue samples of human healthy donors and provides relevant reference groups for twenty human organs including colon.

Finding appropriate normal expression profiles is crucial for the molecular pathway activity analysis and for modeling tumor-specific drug efficacy. For example, a bioinformatic method Oncobox calculates "balanced efficiency score" (BES) for cancer therapeutics based on a parallel analysis of drug target gene expressions and target molecular pathway activation levels in each individual tumor [34–37]. As the output, it returns personalized rating of targeted drugs [38]. This approach was demonstrated to be effective in an ongoing prospective clinical investigation [39] and also for prescription of experimental therapies in advanced cholangiocarcinoma, lung cancer, granulosa cell ovarian cancer, and gastric cancer cases [35, 40–42]. As investigated on a retrospective cohort of gastric cancer patients, it allowed robust prediction of the efficacy of ramucirumab, a vascular endothelial growth factor (VEGF) receptor-specific therapeutic monoclonal antibody [43]. However, the Oncobox method performance so far hasn't been tested for predicting targeted drug efficiencies in the CRC patient cohorts.

In this paper, we report original clinically annotated RNA sequencing profiles for 23 CRC tumors for which we used Oncobox platform to predict individual treatment outcomes based on RNA sequencing data and compared these predictions with the real tumor response records.

Materials and methods

Tissue samples

All experimental biosamples were formalin fixed, paraffin embedded (FFPE) tumor tissue blocks. Prior further analyses, all biosamples were evaluated by a pathologist to confirm the tumor tissue origin and only the specimens with the content of tumor cells greater than 50%

were further investigated. Samples were obtained from various medical institutions of Russia and Lithuania (Supplementary file S1). The patients were 5 men and 18 women (range 37–72 years). 13 patients had sigmoid colon cancer, 1 – cecum cancer, 1 – rectal cancer, 8 – colon cancer (nos). The samples were clinically annotated with the information about the patient's diagnosis and clinical history. For all the biosamples, informed written consents to participate in this study (including publication of anonymized clinical, transcriptomic, and genetic data) were collected from the patients or their legal representatives. The consent procedure and the design of the study were approved by the ethical committee of Vitamed Clinic, Moscow, protocol date 16.10.17. Tumor responses to each line of therapy were characterized according to RECIST criteria for CRC [44]. For those 39 lines of therapy, the following treatment outcomes were collected. *Control over disease* was registered for 12 cases (8 *stable disease* and 4 *partial response* outcomes), and 27 therapies resulted in *progression of the disease.*

For the above 39 lines collected, targeted drugs were used alone or in combination with other chemotherapeutic schemes. The following targeted drugs were administered during these treatments: bevacizumab (n=23), cetuximab (n=3), panitumumab (n=2), regorafenib (n=3), crizotinib (n=2), cabozantinib (n=2), denosumab (n=1), bortezomib (n=1), and aflibercept (n=2) (Table S1).

Preparation of libraries and RNA sequencing

To isolate RNA, 10 μM - thick paraffin slices were trimmed from each FFPE tissue block using microtome. RNA was extracted from FFPE slices using QIAGEN RNeasy FFPE Kit following the manufacturer's protocol. RNA 6000 Nano or Qubit RNA Assay kits were used to measure RNA concentration. RNA Integrity Number (RIN) was measured using Agilent 2100 bio-Analyzer. For depletion of ribosomal RNA and library construction, KAPA RNA Hyper with rRNA erase kit (HMR only) was used. Different adaptors were used for multiplexing samples in one sequencing run. Library concentrations and quality were measured using Qubit ds DNA HS Assay kit (Life Technologies) and Agilent Tapestation (Agilent). RNA sequencing was done at Department of Pathology and Laboratory Medicine, University of California Los Angeles, using Illumina HiSeq 3000 equipment for single-end sequencing, 50 bp read length, for approximately 30 million (mln) raw reads per sample). Data quality check was done on Illumina SAV. De-multiplexing was performed with Illumina Bcl2fastq2 v 2.17 program. Sequencing data were deposited in NCBI Sequencing Read Archive (SRA) under accession ID PRJNA663280.

Processing of RNA sequencing data

RNA sequencing FASTQ files were processed with STAR aligner [45] in "GeneCounts" mode with the Ensembl human transcriptome annotation (Build version GRCh38 and transcript annotation GRCh38.89). Ensembl gene IDs were converted to HGNC gene symbols using Complete HGNC dataset (https://www.genenames.org/, database version from 2017 to 13). Totally, expression levels were established for 36 596 annotated genes with the corresponding HGNC identifiers. Statistics concerning mapping quality and reads number is stored in Supplementary file S2.

Publicly available gene expression datasets

The following publicly available gene expression CRC datasets were used: TCGA, GSE19862, GSE19860, GSE104645, Syn2623706. TCGA cohort contains 17 patients treated with the first-line targeted drugs (bevacizumab, cetuximab, panitumumab, regorafenib). 10 patients demonstrated progressive disease, 1 – stable disease, 4 – partial response, and 2 – complete response. PFS information was available for 117 patients. GSE19860 dataset contains expression data for 12 tumor samples from patients with metastatic or recurrent CRC who were treated with bevacizumab in combination with chemotherapy (five treatment responders and seven non-responders). GSE19862 dataset contains expression profiles for seven responders and seven non-responders to bevacizumab therapy. All patients had metastatic or recurrent CRC and received bevacizumab therapy as the first- or second-line treatment. GSE104645 dataset has 193 gene expression profiles of patients with metastatic CRC, 183 of them were "pre-treatment" samples. The first-line bevacizumab treatment outcome is available for 81 patients (75 responders, 6 non-responders). Also, 80 patients with wild type status of *KRAS* gene (*KRAS*^wt) were annotated by response to "second-line" cetuximab (54 responders, 26 non-responders). PFS data were available for 85 patients, who received bevacizumab, and for 83 patients who received cetuximab. We also used 768 gene expression profiles from Syn2623706 dataset annotated with PFS data.

Molecular pathway analysis

Pathway activation levels were established using Oncobox analytic software [34] for 1682 molecular pathways extracted from the extracted from the public databases Reactome [46], NCI Pathway Interaction Database [47], Kyoto Encyclopedia of Genes and Genomes [48], HumanCyc [49], Biocarta [50] and Qiagen Pathway Central (available at https://www.qiagen.com/us/shop/genes-and-pathways/pathway-central/). The molecular pathways were visualized using Oncobox pathway visualization/reconstruction tool [34, 36].

The pathway activation level (PAL) scores were calculated according to Oncobox method [37]. Tumor gene expression profiles were normalized on the geometric mean tissue expression profile of the control groups to calculate PAL for each individual sample investigated[33]. PAL scores were calculated as follows

$$PAL_p = \sum_n ARR_{np} * \lg(CNR_n) / \sum_n |ARR_{np}|,$$

where PAL_p is *PAL* for pathway p, CNR_n is *case-to-normal ratio*, the ratio of gene *n* expression level in a tumor sample under study to an average level for the control group; *ARR* (*activator/repressor role*) is Boolean flag that depends on function of gene *n* product in pathway *p*. ARR is −1 if gene product *n* inhibits pathway *p*; 1 if *n* activates pathway; 0 if *n* has ambiguous or unclear role in a pathway; 0.5 or −0.5, if *n* is rather activator of a pathway or its inhibitor, respectively.

The random permutation test

To test whether a given number of common differential genes or pathways between the three intersecting datasets is significant, 10 000 random intersections were performed. In every case, three random samples from three corresponding gene/pathway sets of the respective datasets were taken. Then these random samples were intersected for each iteration and 10 000 random sets of common genes/pathways were obtained. P-value of intersection significance was calculated as a fraction of random sets with equal or higher number of common genes/pathways than in the experimental observations.

Ranking targeted drugs and survival analysis

Ranking of target cancer drugs using Balanced Efficiency Scores (BES) was performed as described previously [27, 38, 40]. The Oncobox software returned personalized list of target drugs in descent order of predicted efficacy. The observed clinical responses were used for validation of Oncobox predictions using ROC AUC analysis. ROC AUC was calculated using R ROCR package. Patient survival analysis and plotting were performed using R packages *survival*, *survminer* and *ggplot2*. *P*-value threshold was set to 0.05.

Results

Clinical data

We enrolled twenty three retrospective patients with diagnosed primary or metastatic CRC who received at least one line of therapy with targeted cancer drugs, and to whom clinical response data were available. There were eighteen female and five male 37–72 years old patients enrolled (mean age 46 y.o.). Among them eighteen had distant metastases at the time of the enrollment (Table 1;

Supplementary file S1). In total, results of 39 different lines of therapy were collected, on the average 2 lines per each patient enrolled (Supplementary file S1). With relation to the date of obtaining tumor biopsy, among them 23 represented first line of chemotherapy.

RNA sequencing results

The CRC tissue specimens were the formalin fixed paraffin embedded (FFPE) tissue blocks stored in clinical diagnostic laboratory for 4–49 months before extraction of RNA. RNA sequencing resulted in fourteen sets of 29–41 million sequencing reads, on the average ~35 million reads per library. Among them, there were ~6–15 million reads uniquely mapped on known human Ensembl genes, on the average ~10 million gene-mapped reads per library. All the CRC gene expression profiles obtained here met the quality control criterion for the RNA sequencing protocol used of returning at least 2.5 million uniquely gene-mapped reads per successful library [33].

To assess data reproducibility, we added a technical control of one CRC sample (CC-9 sample) that was RNA-sequenced in four replicates using independent RNA extraction and library preparation procedures.

The gene expression profiles were then analyzed to assess if the profiles obtained are congruent with the biological nature of the biosamples under study. To this end we used principal component analysis (PCA) plot and hierarchical clustering dendrogram to visualize distributions of the experimental CRC profiles obtained here with the profiles of healthy human colon and liver tissues from ANTE database previously obtained with the same equipment and protocol [33], Fig. S1. We observed especially tight clustering for the CC-9 replicates with mean Spearman's correlation coefficient 0.965 (Fig. S1B). The experimental CRC samples formed clear-cut clusters that were most closely located with the normal colon profiles, whereas the liver samples formed outgroup clusters (Fig. S1A, C). These results suggest that the experimental RNA sequencing profiles obtained were technically reproducible and reflected biological properties of the respective biosamples (Fig. S1). We also calculated BES for each of these four replicates to assess their reproducibility for all targeted drugs. Mean Spearman correlation coefficient was 0.926 for comparison of BES among the replicates, standard deviation values are shown in Supplementary table S2.

Genes and molecular pathways linked with CRC treatment prognosis

We performed a search for novel gene expression based prognostic biomarkers. First, we analyzed differential gene expression in order to identify individual genes correlated with poor prognosis in colorectal cancer patients.

Sorokin *et al. BMC Cancer*   (2022) 22:1113

Page 5 of 19

Table 1 Outline of patient clinical and molecular information for the first line targeted therapies

ID	Sample	Targeted drug	Non-target therapy	Response	PFS [*]	Rank [**]	BES [***]	T	N	Distant metastases	RAS [****]
CC-1	Primary colorectal cancer	regorafenib		progressive disease		3	160	-62.1 NA	NA	NA	Mut
CC-2	Primary colorectal cancer	bevacizumab	FOLFOX	progressive disease		3	104	-0.9 3	2	Yes	Mut
CC-4	Primary colorectal cancer	panitumumab	FOLFIRI	progressive disease		2	139	-9.5 4	2	No	WT
CC-5	Primary colorectal cancer	bevacizumab	FOLFIRI	progressive disease		5	81	-0.8 NA	2	Yes	Mut
CC-6	Primary colorectal cancer	cetuximab	FOLFIRI	progressive disease		3	90	-2.6 4	2	Yes	WT
CC-18	Primary colorectal cancer	cetuximab	XELOX	progressive disease		8	137	-9.3 3	1	Yes	WT
CC-34	Primary colorectal cancer	bevacizumab	FOLFOX	progressive disease		4	19	4.4 4	2	Yes	Mut
CC-35	Liver metastasis	bevacizumab	FOLFIRI	partial response		12	54	0.1 3	1	Yes	WT
CC-41	Liver metastasis	bevacizumab	FOLFOX	stable disease		9	5	7.9 4	0	No	WT
CC-44	Primary colorectal cancer	bevacizumab	FOLFOX	stable disease		11	2	23.5 3	0	Yes	WT
CC-45	Peritoneal metastasis	bevacizumab	FOLFIRINOX	stable disease		20	13	5.1 4	1	Yes	Mut
CC-64	Brain metastasis	bevacizumab	FOLFIRI	progressive disease		5	6	24.3 3	0	Yes	WT
CC-65	Primary colorectal cancer	bevacizumab	FOLFIRINOX	partial response		6	22	3.1 4	2	Yes	Mut
CC-72	Liver metastasis	aflibercept	FOLFIRI	progressive disease		3	166	-107.9 3	2	Yes	Mut
CC_91	Small Intestine metastasis	bevacizumab	capecitabine	stable disease		13	6	9.0 3	1	Yes	Mut
CC-94	Adrenal Gland metastasis	bevacizumab	FOLFIRI	partial response		5	27	6.6 3	1	Yes	WT
CC-95	Liver metastasis	bevacizumab	FOLFOX	progressive disease		4	99	0.3 3	1	Yes	WT
CC-105	Primary colorectal cancer	regorafenib		progressive disease		3	113	-3.0 3	1	Yes	WT
CC-107	Primary colorectal cancer	bevacizumab	FOLFOX	stable disease		4	22	1.7 3	2	Yes	Mut
CC-109	Peritoneal metastasis	bevacizumab	FOLFIRI	stable disease		8	1	15.7 4	0	Yes	WT
CC-111	Peritoneal metastasis	bevacizumab	FOLFOX	progressive disease		2	165	-18.3 3	0	Yes	Mut

[*] Progression-free survival

[**] Rank of Oncobox Balanced Efficiency Score for a given targeted drug in a patient

[***] Balanced Efficiency Score of a drug

[****] Presence of known activating mutations in *KRAS* or *HRAS*

Sorokin *et al. BMC Cancer*   (2022) 22:1113

Page 6 of 19

Table 2 *P*-values of log-rank test and Hazard ratio for pathways, which were associated with PFS in three datasets investigated (experimental, TCGA, and Syn2623706)

Dataset	experimental	TCGA	Syn2623706	experimental	TCGA	Syn2623706
Pathway	Hazard ratio			p-value of log-rank test		
KEGG Bladder cancer Pathway	6.21	1.47	0.72	1.84E-02	4.80E-02	1.84E-02
NCI Signaling events mediated by HDAC Class II Pathway	0.16	0.63	0.56	6.11E-05	1.68E-02	6.11E-05
KEGG cGMP PKG signaling Pathway	6.21	1.46	1.56	1.54E-03	4.30E-02	1.54E-03
NCI E cadherin signaling in keratinocytes Pathway	6.21	1.71	1.56	1.64E-03	5.92E-03	1.64E-03
KEGG Aldosterone regulated sodium reabsorption Pathway	5.38	1.50	1.54	2.52E-03	3.11E-02	2.52E-03

We analyzed separately three datasets: the experimental dataset and two external datasets for validation: colorectal TCGA [51] and Syn2623706 [52]. Patients in the experimental cohort had lower PFS time, when compared to TCGA and Syn2623706. The latter may be due to increased proportion of the advanced cancers: the experimental cohort had 69% stage IV patients, in contrast to only 14.3% and 3.8% in the TCGA and Syn2623706 cohorts, respectively. For each patient we took progression-free survival (PFS) time in months and performed Cox proportional hazards survival analysis [53]. For each gene we separated patients in two cohorts: patients with DESeq2 normalized gene expression higher than median in all patients and patients with gene expression lower than median, and compared PFS time between these two groups in Cox model. For each gene we calculated log-rank p-value and Hazard Ratio that indicates relative impact of gene in overall risk of recurrence [54]. We extracted 1787, 1352 and 2770 differentially expressed genes for experimental, TCGA and Syn2623706 datasets, respectively (Supplementary table S3). However, these gene lists showed random intersection pattern ($p = 0.38$).

To further investigate molecular processes linked with the CRC prognosis, we then performed molecular pathway activation analysis using the Oncobox software [55].

In all three datasets (experimental, TCGA, and Syn2623706), we performed Cox survival analysis for each molecular pathway in the same way as for the individual gene products. Patients were stratified into two groups according to PAL value (less and higher than median PAL). We obtained log-rank test p-value < 0.05 for 83 pathways in the experimental dataset, 110 pathways in TCGA dataset, and 687 pathways in the Syn2623706 dataset. Five pathways were common for all three datasets ($p = 0.0311$ according to random permutation test, see Methods for details, Tables 2, Supplementary file S4). *KEGG Bladder cancer pathway*, *EGG cGMP PKG signaling pathway*, *NCI E cadherin signaling in keratinocytes pathway*, and *KEGG Aldosterone regulated sodium reabsorption pathway* were associated with unfavorable prognosis, whereas *NCI Signaling events mediated by HDAC Class II Pathway* was linked with longer PFS. Tumor stage may have an impact on PFS, therefore we also performed multivariate Cox analysis, which included stage

as a variable. All the above pathways remained significant also in the multivariate analysis, except for *KEGG Aldosterone regulated sodium reabsorption pathway* (Supplementary table S4)".

We separately analyzed top 10 pathways associated with poor and favorable prognosis in the experimental dataset. For that we calculated the difference between mean PAL in patients with poor prognosis and mean PAL in patients with favorable prognosis (Supplementary table S4). The pathways with the highest/lowest PAL difference are shown on Fig. 1. The most strongly downregulated pathways in experimental patients with favorable prognosis were "*reactome SHC1 events in ERBB2 signaling Main Pathway*", "*reactome CREB phosphorylation through the activation of Ras Main Pathway*", "*NCI Signaling events regulated by Ret tyrosine kinase Main Pathway*", "*NCI IL5 mediated signaling events Main Pathway*", "*reactome Regulation of KIT signaling Main Pathway*", "*reactome Interleukin 7 signaling Main Pathway*", "*KEGG Bladder cancer Main Pathway*", "*cAMP Pathway Chemotaxis*", "*NCI Endothelins Pathway (cAMP biosynthetic process)*", "*reactome Interferon gamma signaling Main Pathway*". *KEGG Bladder cancer pathway* was associated with poor prognosis in two datasets (experimental and TCGA), but was associated with favorable prognosis in the Syn2623706 dataset. However, Syn2623706 dataset contains data for only 5793 genes, and only two of them are connected to *KEGG Bladder cancer pathway*, thus the corresponding PAL values for Syn2623706 could be irrelevant. Kaplan-Meier plots for *KEGG Bladder cancer Main Pathway* are shown on Fig. 2.

The most strongly upregulated pathways in experimental patients with favorable prognosis were "reactome p130Cas linkage to MAPK signaling for integrins Main Pathway", "NCI Beta2 integrin cell surface interactions Main Pathway", "reactome Integrin alphaIIb beta3 signaling Main Pathway", "biocarta multi step regulation of transcription by pitx2 Main Pathway", "KEGG Complement and coagulation cascades Main Pathway", "biocarta wnt signaling Main Pathway", "NCI Beta3 integrin cell surface interactions Main Pathway", "NCI p73 transcription factor network Pathway apoptosis and DNA repair", "NCI Signaling events mediated by HDAC Class II Main Pathway", "KEGG FoxO signaling Main Pathway". NCI

Sorokin *et al. BMC Cancer* (2022) 22:1113

Page 7 of 19

	P-value	PAL difference
reactome_p130Cas_linkage_to_MAPK_signaling_for_integrins_pathway	0.036	-43.7
NCI_Beta2_integrin_cell_surface_interactions_pathway	0.049	-29.4
reactome_Integrin_alphaIIb_beta3_signaling_pathway	0.036	-28.4
biocarta_multi_step_regulation_of_transcription_by_pitx2_pathway	0.007	-28.1
KEGG_Complement_and_coagulation_cascades_pathway	0.049	-27.7
biocarta_wnt_signaling_pathway	0.018	-23.7
NCI_Beta3_integrin_cell_surface_interactions_pathway	0.036	-20.4
NCI_p73_transcription_factor_network_apoptosis_and_DNA_repair	*0.009*	*-11*
NCI_Signaling_events_mediated_by_HDAC_Class_II_pathway	0.018	-10.8
KEGG_FoxO_signaling_pathway	0.007	-9.7

Favorable prognosis — **Poor prognosis**

	P-value	PAL difference
reactome_Interferon_gamma_signaling_pathway	0.04	18.3
NCI_Endothelins_Pathway(cAMP_biosynthetic_process)	0.018	18.3
cAMP_Pathway_Chemotaxis	0.018	18.6
KEGG_Bladder_cancer_pathway	*0.018*	*18.7*
reactome_Interleukin_7_signaling_pathway	0.027	19
reactome_Regulation_of_KIT_signaling_pathway	0.036	20
NCI_IL5_mediated_signaling_events_pathway	0.018	20.6
NCI_Signaling_events_regulated_by_Ret_tyrosine_kinase_pathway	0.018	24.6
reactome_CREB_phosphorylation_through_the_activation_of_Ras_pathway	0.0002	31.4
reactome_SHC1_events_in_ERBB2_signaling_pathway	0.006	47.6

Fig. 1 Top 10 upregulated and inhibited pathways sorted by PAL values for comparison between poor and favorable prognosis CRC samples. Pathways upregulated in favorable prognosis group are shown on the top, pathways upregulated in poor prognosis group shown on the bottom. Pathways validated using TCGA dataset are shown in italic

p73 transcription factor network Pathway apoptosis and DNA repair was the pathway with the highest PAL difference between patients with poor and favorable prognosis in the experimental dataset, which was also significant in TCGA dataset (Figs. 1 and 2). The same trend was observed also in the Syn2623706 dataset; however, this difference did not reach statistical significance (p=0.098).

The pathway activation schemes for *NCI p73 transcription factor network Pathway apoptosis and DNA repair, KEGG Bladder cancer Pathway* were visualized and shown on Fig. 3. The *KEGG bladder cancer pathway* activation in poor prognosis CRCs was mostly due to increased expression of functional nodes HRAS/KRAS/NRAS, ARAF/BRAF/RAF1, MAPK1/MAPK3, and RPS6KA5 (Fig. 3 A-C). Interestingly, this pathway that was upregulated in poor-prognosis cancers lacks known molecular targets for drugs investigated in this study (Table 3).
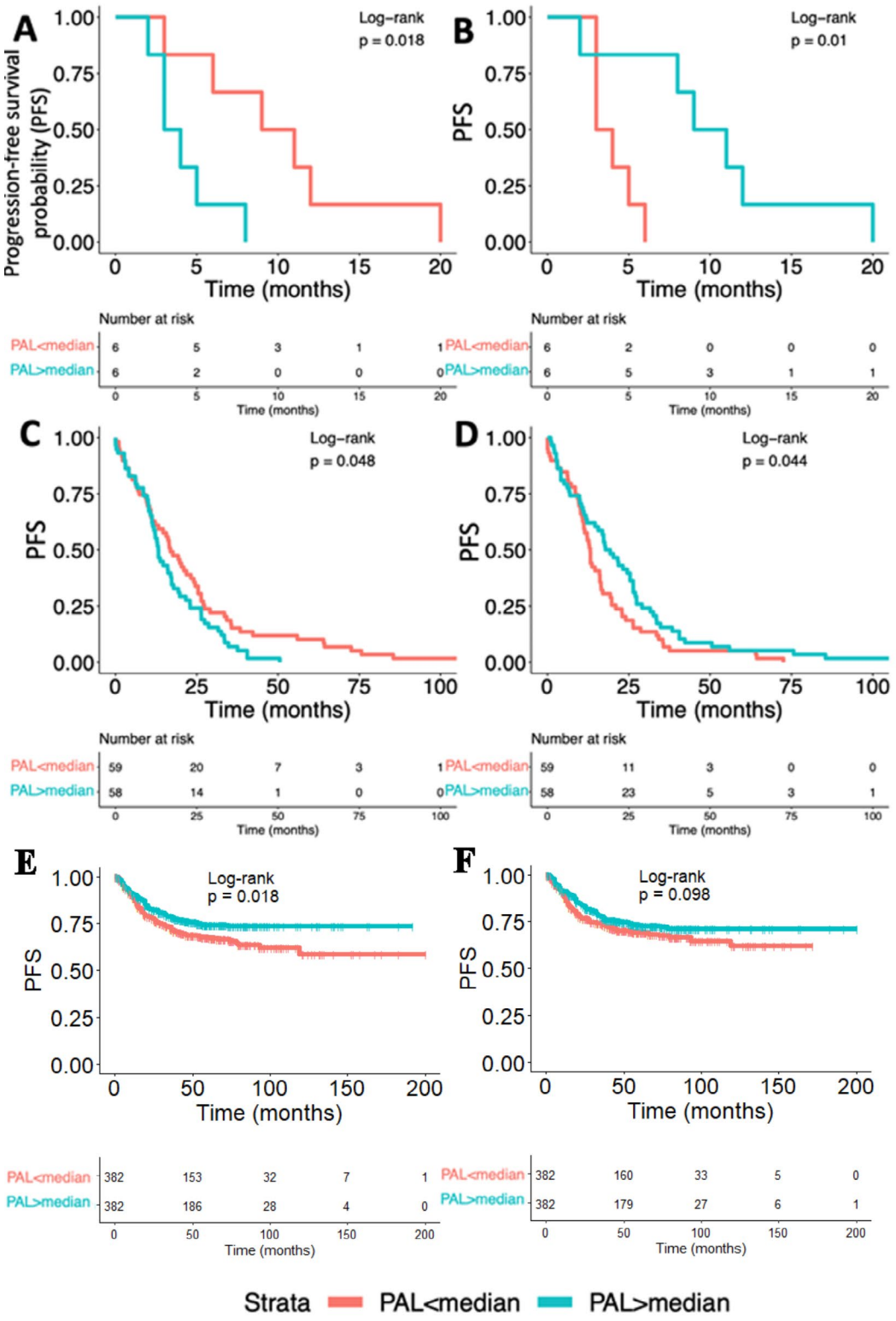
Sorokin *et al. BMC Cancer*   (2022) 22:1113

Page 8 of 19

Fig. 2 Kaplan-Meier plots and risk tables for the validated pathways. **(a)** – experimental samples, *KEGG Bladder cancer Main Pathway*. **(b)** - experimental samples, *NCI p73 tr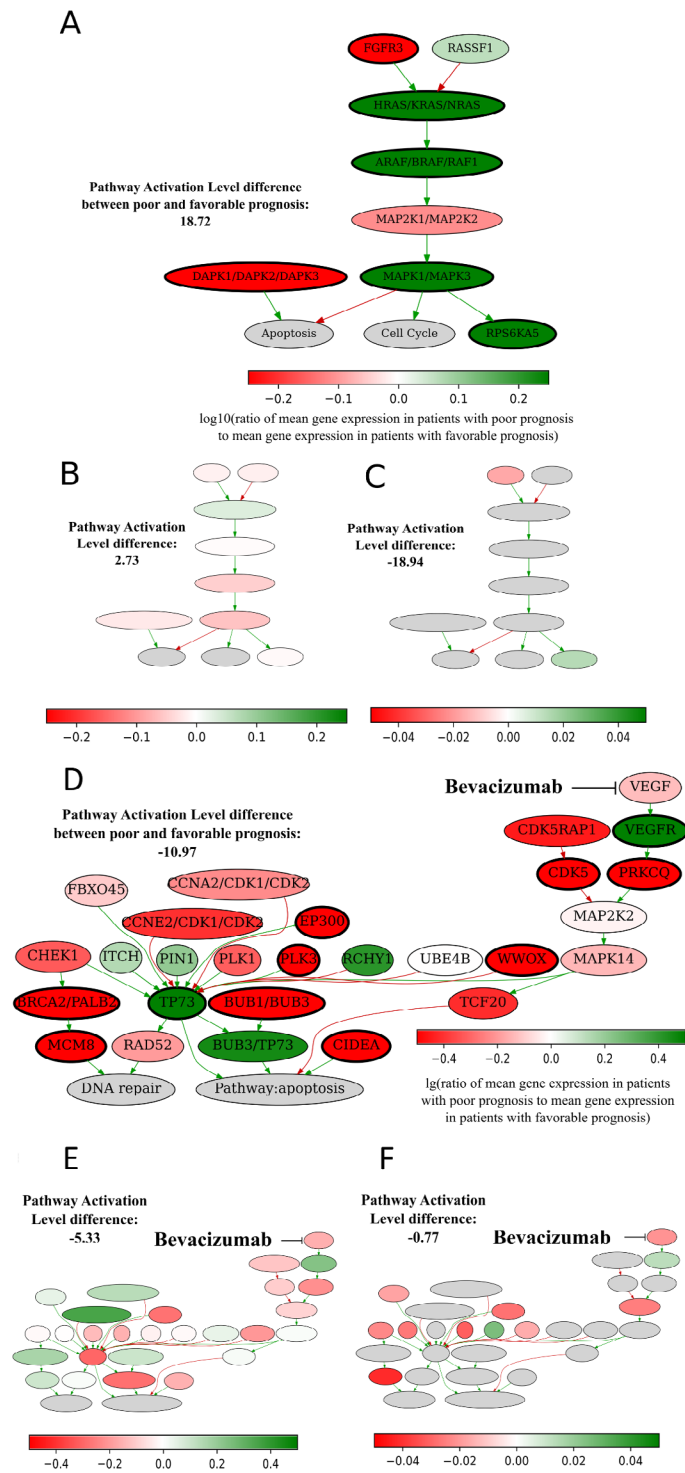anscription factor network Pathway apoptosis and DNA repair*. **(c)** – TCGA samples, *KEGG Bladder cancer Main Pathway*. **(d)** – TCGA samples, *NCI p73 transcription factor network Pathway apoptosis and DNA repair*. **(e)** – Syn2623706 samples, *KEGG Bladder cancer Main Pathway*. **(e)** – Syn2623706 samples, *NCI p73 transcription factor network Pathway apoptosis and DNA repair*

Sorokin *et al. BMC Cancer*   (2022) 22:1113

Page 9 of 19

Fig. 3 Activation schemes of *KEGG Bladder cancer Main Pathway* (**a-c**) and *NCI p73 transcription factor network Pathway apoptosis and DNA repair* (**d-f**) in experimental (**ad**), TCGA (**be**) and Syn2623706 (cf.) CRC samples. Color indicates log10-transformed ratio of mean gene expression values in patients with poor prognosis normalized on gene expression in patients with favorable prognosis. Bold margins indicate most strongly up/downregulated nodes. Patients were divided into poor and favorable prognosis groups by median survival for experimental and TCGA datasets. Because median survival was undefined for Syn2623706 dataset, poor and favorable prognosis was determined relatively median of survival time for non-censored patients

Table 3 Characteristic of targeted drugs investigated in the experimental group

Drug generic name	Molecular target(s)	Reference for molecular target	Drug use in CRC	Reference for drug use
Bevacizumab	VEGF	[56]	Approved for advanced CRC	[7]
Aflibercept	VEGFA, VEGFB, PGF	[57]	Approved for CRC	[58]
Bortezomib	PSMB5, PSMB1	[59]	Off-label for CRC	[60]
Cabozantinib	MET, KDR	[61]	Off-label for CRC	[62]
Cetuximab	EGFR	[63] [64]	Approved for metastatic CRC	[65]
Crizotinib	ALK, MET	[66]	Off-label for CRC	[67]
Panitumumab	EGFR	[68]	Approved for metastatic CRC	[69]
Regorafenib	RET, VEGFR1, VEGFR2, VEGFR3, KIT, PDGFR-alpha, PDGFR-beta, FGFR1, FGFR2, TIE2, DDR2, TrkA, Eph2A, RAF-1, BRAF, SAPK2, PTK5, and Abl	[70]	Approved for metastatic CRC	[9]
Denosumab	TNFSF11	[71]	Off-label for CRC	[72]

The *NCI p73 transcription factor network Pathway apoptosis and DNA repair pathway* was, in contrast, enhanced in favorable prognosis patients in three datasets, but showed borderline significance for Syn2623706 (Fig. 2). On its pathway activation scheme, we observed that this pathway was activated in favorable prognosis group mostly due to increased expression of nodes TP73, BUB1/BUB3, BRCA2/PALB2, MCM8, CIDEA, WWOX, EP300, PLK3, CDK5 and PRKCQ (Fig. 3DE). In turn, this pathway contains VEGF - molecular target of bevacizumab, one of most common CRC drugs (Table 3), which was not among the top activated nodes in favorable prognosis CRCs still was upregulated in them in three datasets (Fig. 3D-F). This favorable prognosis association in the experimental dataset is in line with the clinical records that 15 experimental CRC patients received bevacizumab during the first line of treatment.

Validation of algorithmic prediction of drug response for CRC patients

Oncobox platform is unique in its possibility of algorithmic individual prioritizing of targeted cancer drugs using high-throughput gene expression data and pathway activation profiles [24, 27, 38]. The therapeutics are ranked according to their simulated abilities to inhibit aberrantly regulated molecular pathways and drug target genes [38]. We used this platform to build personalized rating of targeted drugs for each individual CRC sample under investigation. To this end, the RNA sequencing profiles obtained for the experimental CRC samples were individually compared with the set of six healthy colon tissue profiles from the ANTE collection [33]. We chose this type of normalization instead of other collections available from TCGA [51] and GTEX [73] project databases because of identical equipment and sequencing protocols. This was also reflected by closer PCA clustering pattern of experimental CRC profiles and ANTE norms compared to TCGA and GTEX norms (Supplementary Figure S1A).

By comparing tumor versus normal samples, the output modeled drug efficiency value termed *balanced efficiency score* (BES) was calculated for 159 targeted therapeutics present in the Oncobox database [74]. This allowed to individually rank the drugs according to their BES values. The higher BES suggests higher predicted efficacy of a targeted drug for an individual tumor; overall, positive values suggest potential usefulness of a drug to the patient, whereas zero or negative values predict lack of benefit for the patient treatment [40]. Thus, rank of BES score for a given drug among other drugs investigated in an individual patient can serve as a measure of predicted efficacy for this drug compared to the others [40]. Distributions of BES for drugs investigated here and of their ranks are shown for all lines of therapy (Fig. 4AB) and separately for the first line of therapy (Fig. 4CD).

Both standardly approved and off-label drugs used as the experimental therapy were present in our experimental group (Table 2, Supplementary Table S1). Only standardly approved drugs were included in the first line treatments, whereas off-label drugs were prescribed in the further lines. Interestingly, off-label therapeutics formed a small group of treatments with relatively high BES values and ranks (Fig. 4AB).

We then visualized BES values and ranks distributions in relation to the patient response status for all lines of therapy (Fig. 4EF) and separately for the first line of therapy (Fig. 4GH). In both cases drugs prescribed for the responders had relatively higher BES scores (Fig. 4EG, respectively) and ranks (Fig. 4FH, respectively).

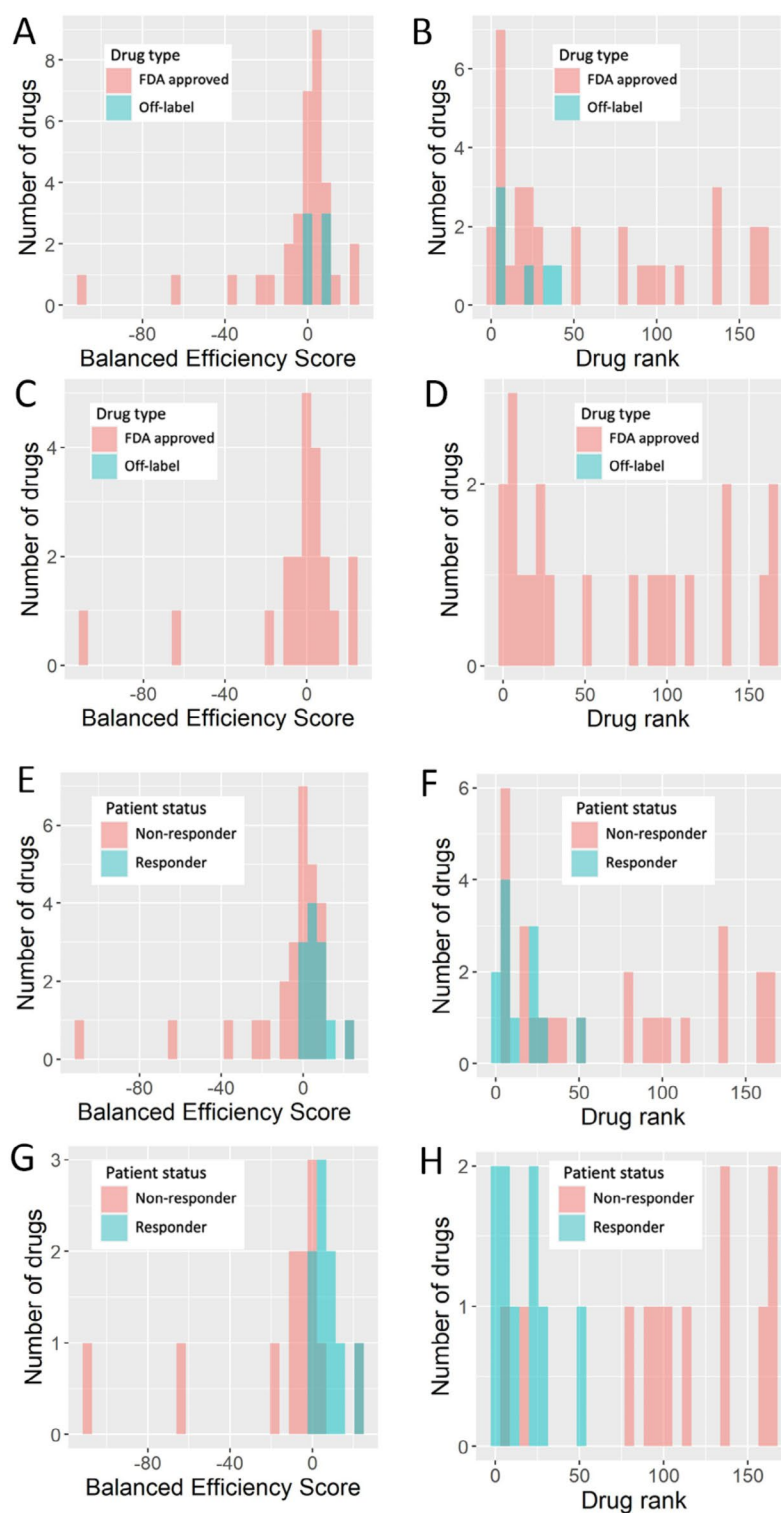We then statistically compared ranks of Oncobox BES values with the registered clinical outcomes in the

Fig. 4 Distribution of BES values and ranks in experimental cohort for all lines of therapy (**a-e**, n=30; **b-f**, n=30) and for first lines of therapy (**c-g**, n=12; **d-h**, n=12). Drug ranks were BES ranks among all drugs in a given patient. Color in **a-d** denotes drugs approved for CRC and experimental therapeutics. Color in **a-d** denotes drugs approved for CRC and experimental therapeutics. Color in **e-h** denotes treatment response. *Responders* group included patients with partial response and stable disease, *non-responders* group included progressive disease patients

experimental patient cohort (Supplementary Table S1). To this end we used the *area under the ROC curve* (ROC AUC) metric that is broadly applicable for assessing biomarker quality in oncology [75–80]. The AUC value is the overall characteristic of biomarker robustness that positively correlates with the quality of a biomarker and varies depending on its sensitivity and specificity in a range between 0.5 and 1 [81]. The standard biomarker quality threshold is 0.7 and greater AUC values indicate good-quality biomarkers, and vice versa [82].

For the available 39 lines of therapy in the experimental cohort (Supplementary Table S1) we compared Oncobox drug ranks versus treatment outcome records (classified as either *control over disease* (positive outcome) or *progressive disease* (negative outcome) (Fig. 4E-H). We found that the Oncobox BES rank as the biomarker could effectively predict CRC response on treatment by targeted therapeutics with AUC=0.77 (Supplementary Figure S2A). For the first-line only targeted therapy data, there were positive outcomes for 9 patients (6 stable disease, 3 partial responses), and the negative outcomes for 12 patients (progressive disease). AUC-based performance of Oncobox drug score ranks in this case was higher than for the analysis of all available outcomes (AUC=0.91; Supplementary figure S2B). We additionally tested BES predictive capacity in bevacizumab-only subset of the experimental data. BES predicted bevacizumab response with AUC 0.82 in the first-line therapy group (9 responders, 6 non-responders), and AUC 0.64 in non-first-line therapy group (1 responder, 7 non-responders). This suggests that using biopsy specimens obtained recently before RNA sequencing and BES analysis can result in more informative reports compared to using older biosamples where further lines of therapy were ongoing, because every additional therapy can reshape tumor gene expression profile. Indeed, BES AUC was only 0.64 for combined non-first line treatments.

We then investigated in the same way performance of the Oncobox drug ranking algorithm for the previously published literature datasets of CRC expression profiles annotated with the targeted drug treatment outcomes. We identified clinical response-annotated datasets GSE19860 and GSE19862 with Affymetrix Human Genome U133 Plus 2.0 microarray gene expression profiles for 12 and 14 CRC patients, respectively, who were treated with bevacizumab. Patients from dataset GSE19860 were treated by the combination therapy of bevacizumab plus mFOLFOX6, whereas patients from dataset GSE19862 received bevacizumab therapy.

We calculated bevacizumab BES based on normalized gene signal intensity values [83] and observed its high performance in predicting drug response for those datasets as well (AUC 0.94 and 0.90 for GSE19860 and GSE19862, respectively; Fig. 5AB).

We further validated BES predictive capacity on publicly available gene expression dataset GSE104645. We found that BES value was an effective response predictor for bevacizumab (AUC=0.84) and was strongly associated with PFS (HR=0.53, CI 0.33–0.84, log rank test p-value 0.0057, Fig. 5CE). BES rank for bevacizumab showed the same trend (AUC=0.73) and was strongly associated with PFS (HR=2.06, CI 1.29–3.29, log rank test p-value 0.0019, Fig. 5DF). However, BES was ineffective for predicting results of second-line treatment with cetuximab (p-value ~0.48, data not shown), and in this case showed no association with PFS (p-value ~0.89, data not shown), which agrees with our previous results on lower efficacy for non-first line treatments predictions (Supplementary Figure S2).

Alternatively, this can be the result of a lower predictive capacity of BES towards EGFR-targeted monoclonal antibodies, in contrast to bevacizumab.

We then investigated performance of Oncobox drug ranking for another relevant clinically annotated gene expression dataset extracted from the TCGA project database. We found that the Oncobox BES rank in the same settings as for the previous patient cohorts predicted treatment response with AUC=0.74 (Supplementary Figure S3).

Comparison of algorithmic drug ranking and experimental progression-free survival data

We then assessed the ability of Oncobox drug score ranks to predict progression-free survival (PFS) of CRC patients following treatment with the respective targeted drugs. Second and further treatment lines were prescribed to the same patients as the first line, therefore we included only first-line treatment outcomes into this analysis. Median PFS for the first line therapy data was 5 months in the entire experimental cohort. For the PFS analysis we stratified patients with the available "first line" targeted therapy outcomes in two groups: (*i*) those who received drug with BES>0 (predicted as effective), or (*ii*) with BES≤0 (predicted as ineffective), summarized on Table 1. The threshold for BES was established in the previous studies [74, 84]. Univariate survival analysis showed the significant difference in PFS between those two cohorts: patients who received in the first line drugs with BES>0 demonstrated higher PFS (Hazard ratio=0.14, log-rank *p*=0.00038), Fig. 6 A.

We then compared the significance of drug score ranks for PFS with the possible impact of patient age, sex, cancer stage and presence of *RAS* family gene mutation using Cox proportional hazards multivariate analysis (Fig. 6B). This type of analysis also confirmed PFS differences between the two cohorts of patients with BES>0 and BES≤0 (*p*<0.001). *RAS* family mutation status was insignificant which agrees
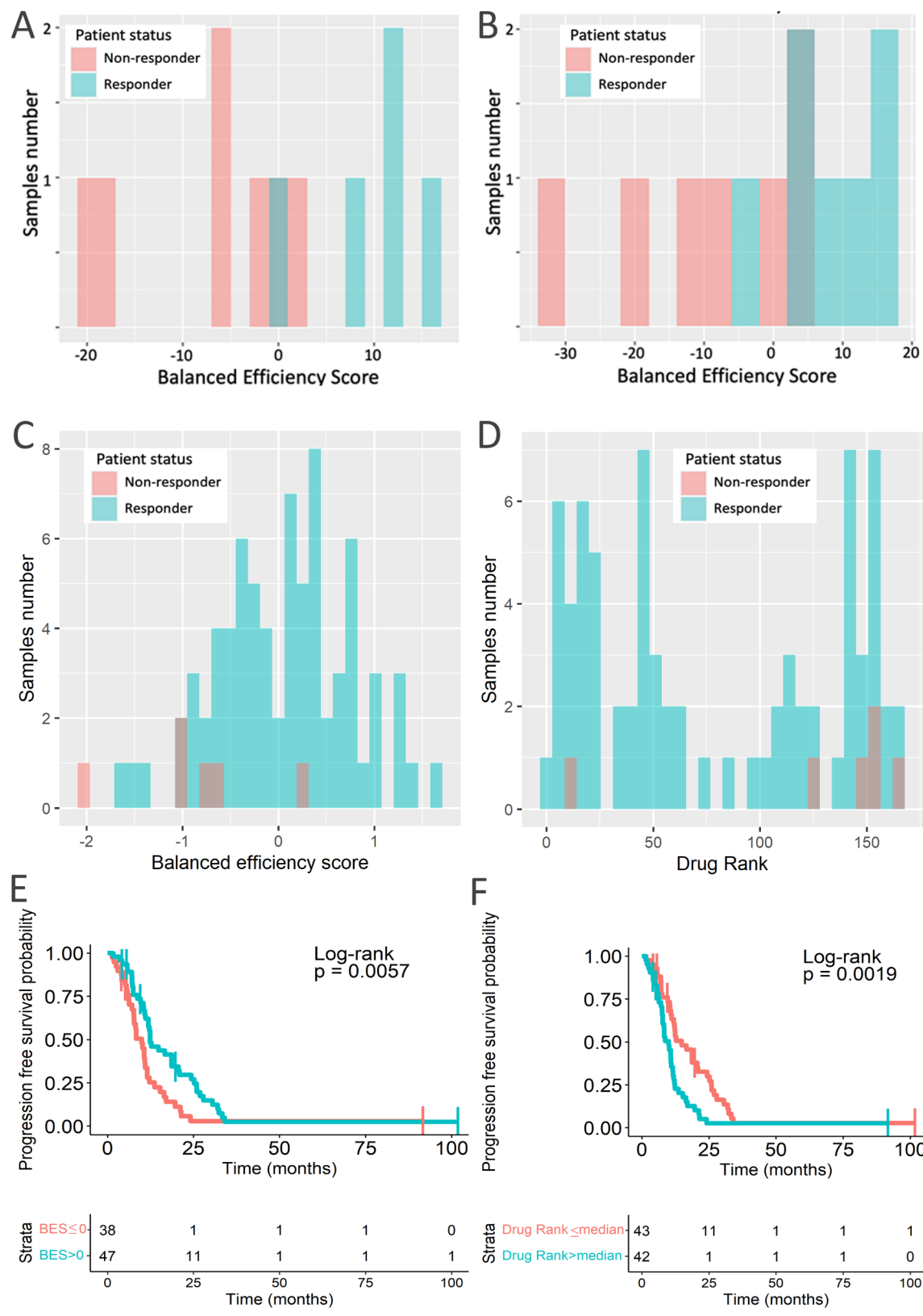
Fig. 5 BES performance for the literature CRC gene expression profiles. **(a)** GSE19860 dataset, AUC=0.94, 5 *responders* and 7 *non-responders*.**(b)** GSE19862 dataset, AUC=0.90, 7 *responders* and *7 non-responders*. **(c)** GSE104645 dataset, AUC=0.84, 75 *responders* and 6 *non-responders*. **(d)** GSE104645 dataset by Drug Rank AUC=0.73, 75 *responders* and 6 *non-responders*. **(ef)** PFS analysis of GSE104645 cohort stratified by BES value **(e)** or rank **(f)**
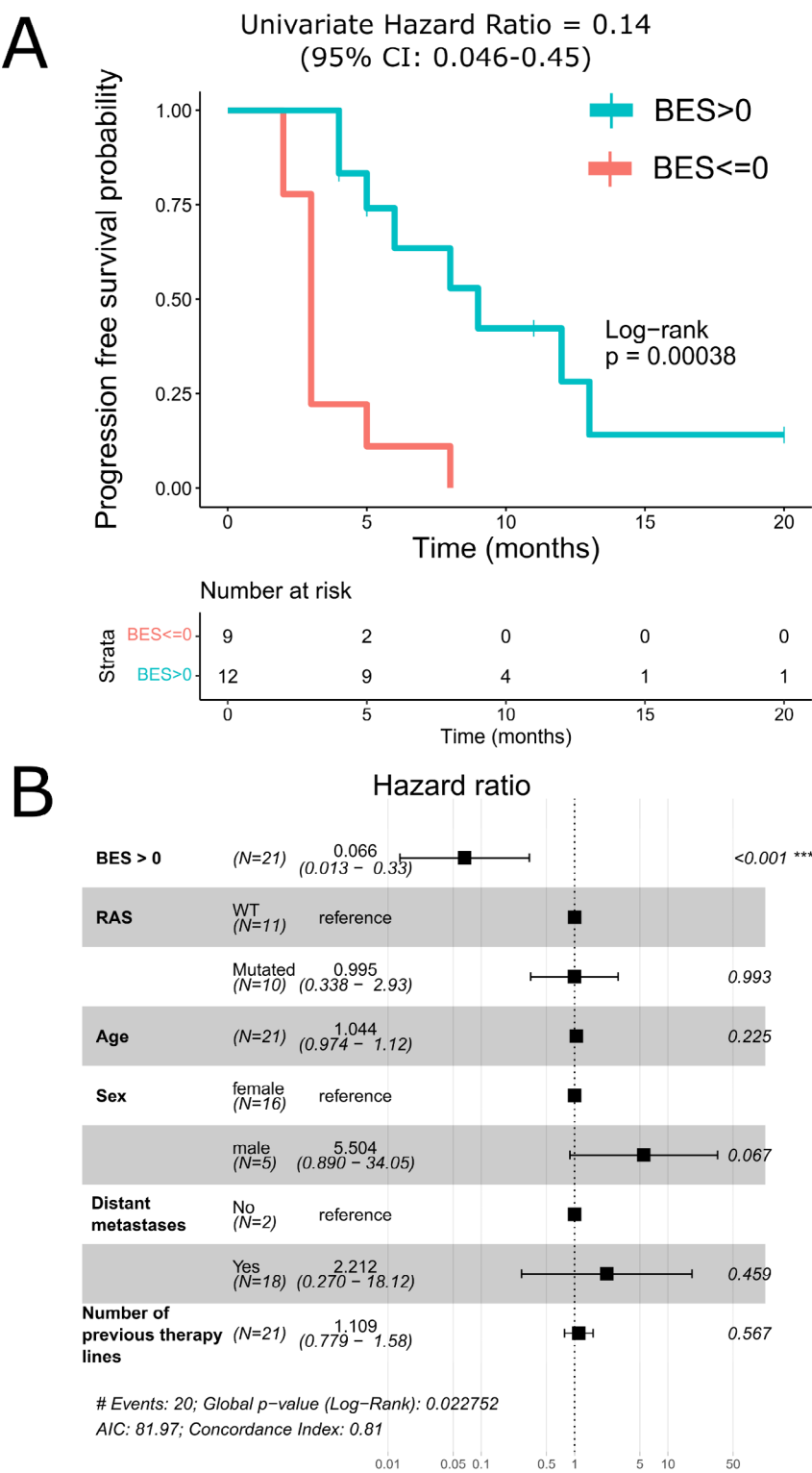
Fig. 6 Survival analysis of experimental dataset for association of PFS and drug BES value. **(a)** – Kaplan-Meier plot based on grouping by BES score (n=21). **(b)** – results of the Cox proportional hazards multivariate analysis (n=21 lines). AIC- Akaike Information Criterion. Hazard Ratios (black squares) for the features are presented on X-scale (log-transformed), whiskers indicate 95% confidence interval. Right column shows Cox proportional-hazards model p-values for Hazard Ratios

with the previous report [85] that on the background of adjuvant chemotherapy (as in this study) *RAS* family mutation is not linked with PFS in CRC. Number of previous lines of therapy was not significant in this analysis possibly due to low variability of this parameter between the patients: 9 patients did not receive prior therapy, 6 patients received one line, 3 patients – 2 lines, 1 patient – 3 lines and 2 patients – 5 lines (Table S1).

Discussion

In this paper we experimentally investigated RNA sequencing gene expression profiles for fourteen CRC patient FFPE tumor biosamples linked with the data on clinical response for 39 lines of therapy containing nine different targeted cancer drugs. Molecular data, such as gene expression and molecular pathway activation levels, is a rich source of information for attempting to personalize cancer drug prescriptions [29, 86, 87]. By using bioinformatic platform Oncobox [34, 36, 39] we modeled here the possible outcomes of nine targeted drugs based on RNA sequencing data and compared the output results with the observed clinical response data.

To this end, the Oncobox drug score (BES) ranks were compared with the therapy response records. We found that the BES ranks could effectively predict tumor response for the experimental CRC patients for the sets of all available therapeutic lines (AUC=0.77) and for the first lines of therapy after obtaining tumor biopsies (AUC=0.91). This is in line with the previous literature on the Oncobox method ability to predict gastric cancer outcomes after treatment with ramucirumab, a VEGF receptor-specific targeted therapeutic, also for the FFPE-derived RNA sequencing profiles. In that case AUC was 0.75 for using ramucirumab as the monotherapy, or 0.7 for its use in combinations with other non-targeted chemotherapy drugs [88]. The performance of drug score ranks for the first line of therapy was significantly higher than for all lines of therapy, thus suggesting that using more recent biopsy specimens is highly influential for the quality of drug treatment response predictions using RNA sequencing data.

Metastatic lesions may have strongly different molecular features than the primary tumors. Liu et al. analyzed matched primary and metastatic CRC samples and did not find significant differences in oncogenic mutations and copy number variations, however identified gene networks associated with CRC metastasis, which may have an impact on treatment outcomes [89]. To further investigate this issue, we performed an additional analysis and excluded metastatic tumors from our experimental dataset (three out of twelve samples). In this case, we obtained AUC 0.7 for all therapy lines analysis, 0.92

for the first-line therapy only, and 0.83 for bevacizumab as the first-line therapy. These results did not differ significantly from the analysis of the entire experimental dataset.

Pentheroudakis et al. previously aimed at identification of a gene expression signature associated with progression-free survival after bevacizumab treatment in CRC. The authors developed potentially effective two-gene signatures, which however were not validated independently [90]. We also performed progression free survival (PFS) analysis for the experimental patients to assess the ability of algorithmic drug score to predict PFS in CRC patients. It was found a significant PFS predictor in both univariate (Hazard ratio=0.14, log-rank p=0.00038) and multivariate (Hazard ratio=0.066, log-rank p<0.001) survival analysis: patients who received drug with BES>0 showed higher PFS. However, the current experimental sampling was limited, and further clinical investigations are needed to provide sufficient line of evidence for introducing algorithmic RNA-based drug scoring to the clinical routine.

The results obtained are in agreement with the hypotheses (*i*) that RNA sequencing data can help personalizing cancer drug prescriptions, and (*ii*) that not only the fresh or freshly-frozen specimens can be used for the analysis, whereas the FFPE-derived materials can be also informative. The latter is important in terms of feasibility of storing tumor biosamples, as FFPE tissue blocks can be stored at room temperature for years, still serving as the source of RNA of acceptable quality for gene expression studies [33]. While RNA is usually degraded in FFPE biosamples, RNAseq still can be used to produce gene expression profiles that give a highly correlated figure with the profiles obtained for the respective fresh tissues biosamples, e.g. as shown by Ben-Moshe et al. [89].

The method of algorithmic drug scoring completely relies on the gene profiling data; thus, RNA sequencing procedure should be standardized for obtaining reproducible results. Another limitation is that due to its rationale it can only rank cancer drugs with characterized molecular specificities (called targeted drugs), whereas many useful non-targeted chemotherapeutics remain outside the analysis.

When comparing treatment responders versus non-responders we also found five molecular pathways that can be indicative of good or poor CRC treatment outcome (Supplementary Table S4).

The most strongly activated pathway connected with poor prognosis was the *KEGG Bladder cancer Main Pathway* and the most strongly activated pathway linked with the favorable prognosis was the *NCI p73 transcription factor network Pathway apoptosis and DNA repair*. The latter was activated in the favorable prognosis group and had several upregulated tumor suppressor

genes promoting DNA repair and apoptosis [91]: BUB1/BUB3, BRCA2/PALB2, MCM8, CIDEA, WWOX, EP300, PLK3, CDK5 and PRKCQ (Fig. 3). It had upregulated VEGF (target of bevacizumab, Table 3) which indicates that favorable prognosis is connected with activation of molecular signaling that can be targeted by bevacizumab. This is also in line with previous study by Kotulak et al., where the authors revealed decreased expression of p73 in colorectal cancer [92]. Moreover, Uboveja et al. showed that p73 plays a critical role in suppression of colon cancer metastasis [93], which in turn may be associated with prognosis. On the contrary, *KEGG Bladder cancer Main Pathway*, which was activated in the poor prognosis group, had no molecular targets of any drug explored in this study. Thus, contrast activation features of these pathways in CRCs may be important for drug response. However, this observation needs to be validated in further independent research using cohorts with a more controlled set of patients with less diverse tumor characteristics e.g. focused on particular molecular subtype as described previously [97].

Taken together, our results support the hypothesis that utilizing targeted therapies in CRC can be improved by using RNA sequencing based algorithmic analysis. However, current study has certain limitations: first, different types of colon and rectal cancers were analyzed altogether. Second, nine targeted drugs were prescribed to the patients in experimental cohort. In future studies, separate analyses for individual drugs and specific cancer subtypes should be conducted to evaluate predictive values of gene expression profiling for prescribing therapy in colorectal cancer.

Conclusion

In this study we analyzed experimental and publicly available gene expression profiles with linked targeted drug treatment outcomes. Oncobox bioinformatical platform was used to simulate targeted drug efficiencies in individual patients. Oncobox demonstrated high predictive capacity, especially when freshly obtained biopsies were used to predict sensitivity to bevacizumab. Our results suggest that gene expression profiling of tumor FFPE materials may be helpful for personalizing prescriptions of targeted therapeutics in CRC.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-022-10177-3.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Availability of data and materials
Supplemental material for this article is available online. The original sequencing data were deposited to NCBI SRA with accession number PRJNA663280 allowing a free access: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA663280/.

Declarations

Ethics approval and consent to participate
The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Vitamed Clinic, Moscow, protocol date 16.10.17.OR. Informed written consents to participate were obtained from all subjects involved in the study. All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication
Informed written consents for publication were obtained from all subjects involved in the study.

Competing interests
MSor, AGar, and AB have a financial relationship with OmicsWay Corp. The remaining authors have nothing to disclose.

Author details
[1]I.M. Sechenov First Moscow State Medical University, 119991 Moscow, Russia
[2]Moscow Institute of Physics and Technology, 141701 Moscow Region, Russia
[3]OmicsWay Corp, 91789 Walnut, CA, USA
[4]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, 117997 Moscow, Russia
[5]World-Class Research Center "Digital biodesign and personalized healthcare", Sechenov First Moscow State Medical University, 119991 Moscow, Russia
[6]Clinical Center Vitamed, 121309 Moscow, Russia
[7]Department of Pathology and Laboratory Medicine, University of California, 90095 Los Angeles, CA, USA
[8]N.N. Blokhin Russian Cancer Research Center, 115478 Moscow, Russia
[9]Medical Holding SM-Clinic, Moscow, Russia

[10]Core Laboratory, The Affiliated Qingdao Central Hospital of Qingdao University, Qìngdao, China

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.
2. Aguiar Junior S, Oliveira MM de, Silva DRME, Mello CAL de, Calsavara VF, Curado MP. SURVIVAL OF PATIENTS WITH COLORECTAL CANCER IN A CANCER CENTER. Arq Gastroenterol. 2020;57:172–7.
3. Kumar S, Burney IA, Zahid KF, Souza D, Belushi PC, Mufti MAL. TD, et al. Colorectal Cancer Patient Characteristics, Treatment and Survival in Oman–a Single Center Study. Asian Pac J Cancer Prev. 2015;16:4853–8.
4. Maajani K, Khodadost M, Fattahi A, Shahrestanaki E, Pirouzi A, Khalili F, et al. Survival Rate of Colorectal Cancer in Iran: A Systematic Review and Meta-Analysis. Asian Pac J Cancer Prev. 2019;20:13–21.
5. Baghestani AR, Daneshvar T, Pourhoseingholi MA, Asadzade H. Survival of colorectal cancer patients in the presence of competing-risk. Asian Pac J Cancer Prev. 2014;15:6253–5.
6. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet (London England). 2014;383:1490–502.
7. Cohen MH, Gootenberg J, Keegan P, Pazdur R. FDA drug approval summary: bevacizumab plus FOLFOX4 as second-line treatment of colorectal cancer. Oncologist. 2007;12:356–61.
8. André T, Chibaudel B. [Aflibercept (Zaltrap(®)) approved in metastatic colorectal cancer]. Bull Cancer. 2013;100:1023–5.
9. Ettrich TJ, Seufferlein T. Regorafenib. In: Recent Results in Cancer Research. Springer New York LLC. 2018. p. 45–56. https://doi.org/10.1007/978-3-319-91442-8_3.
10. Garrett CR, Eng C. Cetuximab in the treatment of patients with colorectal cancer. Expert Opin Biol Ther. 2011;11:937–49.
11. Giusti RM, Shastri KA, Cohen MH, Keegan P, Pazdur R. FDA drug approval summary: panitumumab (Vectibix). Oncologist. 2007;12:577–83.
12. Xie Y-H, Chen Y-X, Fang J-Y. Comprehensive review of targeted therapy for colorectal cancer. Signal Transduct Target Ther. 2020;5:22.
13. Link W. Principles of Cancer Treatment and Anticancer Drug Development. Cham: Springer International Publishing; 2019.
14. Afrăsânie VA, Marinca MV, Alexa-Stratulat T, Gafton B, Păduraru M, Adavidoaiei AM, et al. KRAS, NRAS, BRAF, HER2 and microsatellite instability in metastatic colorectal cancer-practical implications for the clinician. Radiology and Oncology. 2019;53.
15. Nguyen M, Tipping Smith S, Lam M, Liow E, Davies A, Prenen H, et al. An update on the use of immunotherapy in patients with colorectal cancer. Expert Rev Gastroenterol Hepatol. 2020;:1–14. https://doi.org/10.1080/17474124.2021.1845141.
16. Kanat O, Ertas H, Caner B. Contemporary treatment approaches for metastatic colorectal cancer driven by BRAF V600 mutations. World J Gastrointest Oncol. 2020;12:1080–90. https://doi.org/10.4251/wjgo.v12.i10.1080.
17. Tosi F, Sartore-Bianchi A, Lonardi S, Amatu A, Leone F, Ghezzi S, et al. Long-term Clinical Outcome of Trastuzumab and Lapatinib for HER2-positive Metastatic Colorectal Cancer. Clin Colorectal Cancer. 2020. https://doi.org/10.1016/j.clcc.2020.06.009.
18. Sartore-Bianchi A, Lonardi S, Martino C, Fenocchio E, Tosi F, Ghezzi S, et al. Pertuzumab and trastuzumab emtansine in patients with HER2-amplified metastatic colorectal cancer: The phase II HERACLES-B trial. ESMO Open. 2020;5. https://doi.org/10.1136/esmoopen-2020-000911.
19. Bruera G, Pepe F, Malapelle U, Pisapia P, Mas AD, Giacomo D, Di, et al. Prevalence of KRAS, NRAS and BRAF mutations detected by massive parallel sequencing and differential clinical outcome in metastatic colorectal cancer (MCRC) patients (pts) treated with first line FIr-B/FOx adding bevacizumab (BEV) to triplet chemotherapy. Ann Oncol. 2017;28:vi14.
20. Vaughn CP, Zobell SD, Furtado LV, Baker CL, Samowitz WS. Frequency of KRAS, BRAF, and NRAS mutations in colorectal cancer. Genes Chromosomes Cancer. 2011;50:307–12.
21. Li Z-N, Zhao L, Yu L-F, Wei M-J. BRAF and KRAS mutations in metastatic colorectal cancer: future perspectives for personalized therapy. Gastroenterol Rep. 2020;8:192–205.
22. Sanchez-Ibarra HE, Jiang X, Gallegos-Gonzalez EY, Cavazos-González AC, Chen Y, Morcos F, et al. KRAS, NRAS, and BRAF mutation prevalence, clinicopathological association, and their application in a predictive model in Mexican patients with metastatic colorectal cancer: A retrospective cohort study. PLoS ONE. 2020;15:e0235490.
23. Bylsma LC, Gillezeau C, Garawin TA, Kelsh MA, Fryzek JP, Sangaré L, et al. Prevalence of RAS and BRAF mutations in metastatic colorectal cancer patients by tumor sidedness: A systematic review and meta-analysis. Cancer Med. 2020;9:1044–57.
24. 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. 2019. https://doi.org/10.1016/j.semcancer.2019.07.010.
25. Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. Mol Cancer Ther. 2015;14:847–56.
26. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AAM, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002;347:1999–2009. https://doi.org/10.1056/NEJMoa021967.
27. 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. https://doi.org/10.1016/j.semcancer.2018.06.003.
28. SEQC/MAQC-III Consortium. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Qual Control Consortium. 2014;32. https://doi.org/10.1038/nbt.2957.
29. 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. https://doi.org/10.1038/s41591-019-0424-4.
30. 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 Inf. 2019;18:1176935119838844. https://doi.org/10.1177/1176935119838844.
31. Buzdin AA, Zhavoronkov AA, Korzinkin MB, Roumiantsev SA, Aliper AM, Venkova LS, et al. The OncoFinder algorithm for minimizing the errors introduced by the high-throughput methods of transcriptome analysis. Front Mol Biosci. 2014;1:8. https://doi.org/10.3389/fmolb.2014.00008.
32. Borisov N, Shabalina I, Tkachev V, Sorokin M, Garazha A, Pulin A, et al. Shambhala: a platform-agnostic data harmonizer for gene expression data. BMC Bioinformatics. 2019;20:66. https://doi.org/10.1186/s12859-019-2641-8.
33. 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. https://doi.org/10.1038/s41597-019-0043-4.
34. Sorokin M, Kholodenko R, Suntsova M, Malakhova G, Garazha A, Kholodenko I, et al. Oncobox Bioinformatical Platform for Selecting Potentially Effective Combinations of Target Cancer Drugs Using High-Throughput Gene Expression Data. Cancers (Basel). 2018;10:365. https://doi.org/10.3390/cancers10100365.
35. Poddubskaya EV, Baranova MP, Allina DO, Smirnov PY, Albert EA, Kirilchev AP, et al. Personalized prescription of tyrosine kinase inhibitors in unresectable metastatic cholangiocarcinoma. Exp Hematol Oncol. 2018;7:21. https://doi.org/10.1186/s40164-018-0113-x.
36. Buzdin A, Garazha A, Sorokin M, Glusker A, Aleshin A, Allina D, et al. RNA sequencing analysis for profiling activation of cancer-associated molecular pathways. J Clin Oncol. 2019;37 15_suppl:e13032–e13032.
37. Borisov N, Sorokin M, Garazha A, Buzdin A. Quantitation of molecular pathway activation using RNA sequencing data. In: Methods in molecular biology (Clifton NJ). 2019. p. In press. https://doi.org/10.1007/978-1-0716-0138-9_15.
38. Tkachev V, Sorokin M, Garazha A, Borisov N, Buzdin A. Oncobox method for scoring efficiencies of anticancer drugs based on gene expression data. In: Astakhova K, Bukhari SA, editors. Methods Mol Biol. New York: Springer US; 2020. pp. 235–55. https://doi.org/10.1007/978-1-0716-0138-9_17.
39. Poddubskaya E, Buzdin A, Garazha A, Sorokin M, Glusker A, Aleshin A, et al. Oncobox, gene expression-based second opinion system for predicting response to treatment in advanced solid tumors. J Clin Oncol. 2019;37 15_suppl:e13143–3.
40. Poddubskaya EV, Baranova MP, Allina DO, Sekacheva MI, Makovskaia LA, Kamashev DE, et al. Personalized prescription of imatinib in recurrent granulosa cell tumor of the ovary: case report. Cold Spring Harb Mol case Stud. 2019;5:a003434. https://doi.org/10.1101/mcs.a003434.
41. 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. https://doi.org/10.3390/biomedicines8030067.

42. Poddubskaya E, Bondarenko A, Boroda A, Zotova E, Glusker A, Sletina S, et al. Transcriptomics-Guided Personalized Prescription of Targeted Therapeutics for Metastatic ALK-Positive Lung Cancer Case Following Recurrence on ALK Inhibitors. Front Oncol. 2019;9:1026. https://doi.org/10.3389/fonc.2019.01026.

43. Sorokin M, Poddubskaya E, Baranova M, Glusker A, Kogoniya L, Markarova E, et al. RNA sequencing profiles and diagnostic signatures linked with response to ramucirumab in gastric cancer. Cold Spring Harb Mol case Stud. 2020;6.

44. Trillet-Lenoir V, Freyer G, Kaemmerlen P, Fond A, Pellet O, Lombard-Bohas C, et al. Assessment of tumour response to chemotherapy for metastatic colorectal cancer: accuracy of the RECIST criteria. Br J Radiol. 2002;75:903–8.

45. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21. https://doi.org/10.1093/bioinformatics/bts635.

46. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic Acids Res. 2014;42:D472–7. https://doi.org/10.1093/nar/gkt1102.

47. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the Pathway Interaction Database. Nucleic Acids Res. 2009;37 suppl_1:D674–9. https://doi.org/10.1093/nar/gkn653.

48. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.

49. Romero P, Wagg J, Green ML, Kaiser D, Krummenacker M, Karp PD. Computational prediction of human metabolic pathways from the complete human genome. Genome Biol. 2005;6:R2. https://doi.org/10.1186/gb-2004-6-1-r2.

50. Nishimura D. BioCarta. Biotech Softw Internet Rep. 2001;2:117–20. https://doi.org/10.1089/152791601750294344.

51. Cancer Genome Atlas Research. Network JN, Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013;45:1113–20. https://doi.org/10.1038/ng.2764.

52. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015 2111. 2015;21:1350–6.

53. George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol. 2014;21:686–94.

54. Singh R, Mukhopadhyay K. Survival analysis in clinical trials: Basics and must know areas. Perspect Clin Res. 2011;2:145–8.

55. Buzdin AA, Sorokin M, Borisov NM, Kuzmin D, Gudkov A, Zolotovskaia MA, et al. Algorithmic annotation of functional roles for components of 3044 human molecular pathways. Front Genet. 2021;12:139.

56. Shih T, Lindley C. Bevacizumab: an angiogenesis inhibitor for the treatment of solid malignancies. Clin Ther. 2006;28:1779–802.

57. Browning DJ, Kaiser PK, Rosenfeld PJ, Stewart MW. Aflibercept for age-related macular degeneration: a game-changer or quiet addition? Am J Ophthalmol. 2012;154:222–6.

58. Sharma T, Dhingra R, Singh S, Sharma S, Tomar P, Malhotra M, et al. Aflibercept: a novel VEGF targeted agent to explore the future perspectives of anti-angiogenic therapy for the treatment of multiple tumors. Mini Rev Med Chem. 2013;13:530–40.

59. Chen D, Frezza M, Schmitt S, Kanwar J, Dou QP. Bortezomib as the first proteasome inhibitor anticancer drug: current status and future perspectives. Curr Cancer Drug Targets. 2011;11:239–53.

60. Wang S, Wang L, Zhou Z, Deng Q, Li L, Zhang M, et al. Leucovorin Enhances the Anti-cancer Effect of Bortezomib in Colorectal Cancer Cells. Sci Rep. 2017;7:682.

61. Yakes FM, Chen J, Tan J, Yamaguchi K, Shi Y, Yu P, et al. Cabozantinib (XL184), a novel MET and VEGFR2 inhibitor, simultaneously suppresses metastasis, angiogenesis, and tumor growth. Mol Cancer Ther. 2011;10:2298–308.

62. Abdelaziz A, Vaishampayan U. Cabozantinib for the treatment of kidney cancer. Expert Rev Anticancer Ther. 2017;17:577–84.

63. Lee MS, Kopetz S. Current and Future Approaches to Target the Epidermal Growth Factor Receptor and Its Downstream Signaling in Metastatic Colorectal Cancer. Clin Colorectal Cancer. 2015;14:203–18.

64. Goldberg RM. Cetuximab. Nat Rev Drug Discov. 2005;Suppl:10-1.

65. Geng F, Wang Z, Yin H, Yu J, Cao B. Molecular Targeted Drugs and Treatment of Colorectal Cancer: Recent Progress and Future Perspectives. Cancer Biother Radiopharm. 2017;32:149–60.

66. Forde PM, Rudin CM. Crizotinib in the treatment of non-small-cell lung cancer. Expert Opin Pharmacother. 2012;13:1195–201.

67. Kazandjian D, Blumenthal GM, Chen H-Y, He K, Patel M, Justice R, et al. FDA approval summary: crizotinib for the treatment of metastatic non-small cell lung cancer with anaplastic lymphoma kinase rearrangements. Oncologist. 2014;19:e5–11.

68. Lynch DH, Yang X-D. Therapeutic potential of ABX-EGF: a fully human anti-epidermal growth factor receptor monoclonal antibody for cancer treatment. Semin Oncol. 2002;29(1 Suppl 4):47–50.

69. Giusti RM, Cohen MH, Keegan P, Pazdur R. FDA review of a panitumumab (Vectibix) clinical trial for first-line treatment of metastatic colorectal cancer. Oncologist. 2009;14:284–90.

70. Grothey A, Prager G, Yoshino T. The Mechanism of Action of Regorafenib in Colorectal Cancer: A Guide for the Community Physician. Clin Adv Hematol Oncol. 2019;17(Suppl 1):1–19.

71. Westenfeld R, Ketteler M, Brandenburg VM. Anti-RANKL therapy–implications for the bone-vascular-axis in CKD? Denosumab in post-menopausal women with low bone mineral density. Nephrol Dial Transplant. 2006;21:2075–7.

72. Thosani S, Hu MI. Denosumab: a new agent in the management of hypercalcemia of malignancy. Future Oncol. 2015;11:2865–71.

73. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45:580–5.

74. Tkachev V, Sorokin M, Garazha A, Borisov N, Buzdin A, Tkachev V, Sorokin M, Garazha A, Borisov NBA, et al. Oncobox Method for Scoring Efficiencies of Anticancer Drugs Based on Gene Expression Data. Methods Mol Biol. 2020;2063:235–55.

75. Chen L, Zhou Y, Tang X, Yang C, Tian Y, Xie R, et al. EGFR mutation decreases FDG uptake in non–small cell lung cancer via the NOX4/ROS/GLUT1 axis. Int J Oncol. 2019;54:370–80. https://doi.org/10.3892/ijo.2018.4626.

76. Tanioka M, Fan C, Parker JS, Hoadley KA, Hu Z, Li Y, et al. Integrated analysis of RNA and DNA from the phase III trial CALGB 40601 identifies predictors of response to trastuzumab-based neoadjuvant chemotherapy in HER2-positive breast cancer. Clin Cancer Res. 2018;24:5292–304. https://doi.org/10.1158/1078-0432.CCR-17-3431.

77. Liu T, Cheng G, Kang X, Xi Y, Zhu Y, Wang K, et al. Noninvasively evaluating the grading and IDH1 mutation status of diffuse gliomas by three-dimensional pseudo-continuous arterial spin labeling and diffusion-weighted imaging. Neuroradiology. 2018;60:693–702. https://doi.org/10.1007/s00234-018-2021-5.

78. Zolotovskaia MA, Sorokin MI, Roumiantsev SA, Borisov NM, Buzdin AA. Pathway Instability Is an Effective New Mutation-Based Type of Cancer Biomarkers. Front Oncol. 2018;8:658. https://doi.org/10.3389/fonc.2018.00658.

79. Lezhnina K, Kovalchuk O, Zhavoronkov AA, Korzinkin MB, Zabolotneva AA, Shegay PV, et al. Novel robust biomarkers for human bladder cancer based on activation of intracellular signaling pathways. Oncotarget. 2014;5:9022–32. https://doi.org/10.18632/oncotarget.2493.

80. Borisov NM, Terekhanova NV, Aliper AM, Venkova LS, Smirnov PY, Roumiantsev S, et al. Signaling pathways activation profiles make better markers of cancer than expression of individual genes. Oncotarget. 2014;5:10198–205. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4259415&tool=pmcentrez&rendertype=abstract. Accessed 5 Oct 2015.

81. Green DM, Swets JA. Signal detection theory and psychophysics. R.E. Krieger Pub. Co; 1974.

82. Boyd JC. Mathematical tools for demonstrating the clinical usefulness of biochemical markers. Scand J Clin Lab Invest. 1997;57:46–63.

83. Harbig J, Sprinkle R, Enkemann SA. A sequence-based identification of the genes detected by probesets on the Affymetrix U133 plus 2.0 array. Nucleic Acids Res. 2005;33:e31.

84. Sorokin M, Poddubskaya E, Baranova M, Glusker A, Kogoniya L, Markarova E, et al. RNA sequencing profiles and diagnostic signatures linked with response to ramucirumab in gastric cancer. Mol Case Stud. 2020;:mcs.a004945. https://doi.org/10.1101/mcs.a004945.

85. Deng Y, Wang L, Tan S, Kim GP, Dou R, Chen D, et al. KRAS as a predictor of poor prognosis and benefit from postoperative FOLFOX chemotherapy in patients with stage II and III colorectal cancer. Mol Oncol. 2015;9:1341–7.

86. Ben-Hamo R, Jacob Berger A, Gavert N, Miller M, Pines G, Oren R, et al. Predicting and affecting response to cancer therapy based on pathway-level biomarkers. Nat Commun. 2020;11:3296.

87. Estevez-Garcia P, Rivera F, Molina-Pinelo S, Benavent M, Gómez J, Limón ML, et al. Gene expression profile predictive of response to chemotherapy in metastatic colorectal cancer. Oncotarget. 2015;6:6151–9.

88. Sorokin M, Poddubskaya E, Baranova M, Glusker A, Kogoniya L, Markarova E, et al. RNA sequencing profiles and diagnostic signatures linked with response to ramucirumab in gastric cancer. Cold Spring Harb Mol case Stud. 2020;6.

89. Liu J, Cho YB, Hong HK, Wu S, Ebert PJ, Bray SM, et al. Molecular dissection of CRC primary tumors and their matched liver metastases reveals critical role of

immune microenvironment, EMT and angiogenesis in cancer metastasis. Sci Rep. 2020;10:10725. https://doi.org/10.1038/s41598-020-67842-5.

90. Pentheroudakis G, Kotoula V, Fountzilas E, Kouvatseas G, Basdanis G, Xanthakis I, et al. A study of gene expression markers for predictive significance for bevacizumab benefit in patients with metastatic colon cancer: a translational research study of the Hellenic Cooperative Oncology Group (HeCOG). BMC Cancer. 2014;14:111. https://doi.org/10.1186/1471-2407-14-111.

91. De Zio D, Cianfanelli V, Cecconi F. New insights into the link between DNA damage and apoptosis. Antioxid Redox Signal. 2013;19:559–71.

92. Kotulak A, Wronska A, Kobiela J, Godlewski J, Stanislawowski M, Wierzbicki P. Decreased expression of p73 in colorectal cancer. Folia Histochem Cytobiol. 2016;54:166–70. https://doi.org/10.5603/FHC.a2016.0018.

93. Uboveja A, Satija YK, Siraj F, Sharma I, Saluja D. p73 – NAV3 axis plays a critical role in suppression of colon cancer metastasis. Oncogenesis. 2020;9:12. https://doi.org/10.1038/s41389-020-0193-4.

Publisher's Note