

RESEARCH

Open Access



# Frequent variations in cancer-related genes may play prognostic role in treatment of patients with chronic myeloid leukemia

Alexander V. Lavrov<sup>1,2\*</sup>, Ekaterina Y. Chelysheva<sup>3</sup>, Svetlana A. Smirnikhina<sup>1</sup>, Oleg A. Shukhov<sup>3</sup>, Anna G. Turkina<sup>3</sup>, Elmira P. Adilgereeva<sup>1</sup> and Sergey I. Kutsev<sup>1,2</sup>

From The 7th International Young Scientists School "Systems Biology and Bioinformatics" (SBB'2015) Novosibirsk, Russia. 22-25 June 2015

## Abstract

**Background:** Genome variability of host genome and cancer cells play critical role in diversity of response to existing therapies and overall success in treating oncological diseases. In chronic myeloid leukemia targeted therapy with tyrosine kinase inhibitors demonstrates high efficacy in most of the patients. However about 15 % of patients demonstrate primary resistance to standard therapy. Whole exome sequencing is a good tool for unbiased search of genetic variations important for prognosis of survival and therapy efficacy in many cancers. We apply this approach to CML patients with optimal response and failure of tyrosine kinase therapy.

**Results:** We analyzed exome variations between optimal responders and failures and found 7 variants in cancer-related genes with different genotypes in two groups of patients. Five of them were found in optimal responders: rs11579366, rs1990236, rs176037, rs10653661, rs3803264 and two in failures: rs3099950, rs9471966. These variants were found in genes associated with cancers (*ANKRD35*, *DNAH9*, *MAGEC1*, *TOX3*) or participating in cancer-related signaling pathways (*THSD1*, *MORN2*, *PTCRA*).

**Conclusion:** We found gene variants which may become early predictors of the therapy outcome and allow development of new early prognostic tests for estimation of therapy efficacy in CML patients. Normal genetic variation may influence therapy efficacy during targeted treatment of cancers.

**Keywords:** Genetic variation, Chronic myeloid leukemia, TKI resistance, Whole exome sequencing, Biomarkers, Prognostic factors

## Background

Chronic myeloid leukemia (CML) accounts for about 10 % of all leukemias. The primary mutation in blood stem cell is induced by the translocation between chromosomes 9 and 22 leading to the formation of the so-called Philadelphia chromosome (Ph<sup>+</sup>) – an aberrant chromosome 22. As a

result of this translocation two genes are fused: *ABL* (chromosome 9) and *BCR* (chromosome 22) [1]. *BCR/ABL* gene produces atypical tyrosine kinase – BCR/ABL which activates MAPK pathway and induces cell proliferation, blocks apoptosis and leads to genome instability resulting in further development of the disease.

Till late 1990s the prognosis for the patients with CML was poor. The disease progressed gradually to the accelerated phase, blast crisis and finally death. Invention of BCR/ABL tyrosine kinase inhibitor (TKI) made the disease well controlled and dramatically improved patient's survival. The first TKI – imatinib – has become the standard therapy of CML since its FDA approval in 2001. Patient's

\* Correspondence: alexandervlavrov@gmail.com

<sup>1</sup>Laboratory of Mutagenesis, Federal State Budgetary Institution "Research Centre for Medical Genetics", Moskvorechie, 1, Moscow 115478, Russia

<sup>2</sup>Department of Molecular and Cellular Genetics, State Budgetary Educational Institution of Higher Professional Education "Russian National Research Medical University named after N.I. Pirogov" of Ministry of Health of the Russian Federation, Moscow, Russia

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

survival considerably improved from 7.5 years after diagnosis before imatinib invention to 17.5 years now-a-days [2]. However some patients still remain resistant to TKI therapy. The rate of primary resistance is in the range of 5–20 % [3, 4]. According to ELN recommendations 2013 [5] CML patients on TKI therapy are assigned to optimal responders, failures and “warning” group. Optimal responders may continue the same treatment, failures need treatment changes, “warning” group needs careful observation with the readiness to change treatment options. Currently available prognostic scores – Sokal and EUTOS may help to predict patients with worse clinical outcome on TKIs although some studies failed to demonstrate predictive value of EUTOS score for different clinical outcomes including long-term overall survival rate [6, 7]. Therefore additional clinically relevant prognostic markers for CML patients are extremely important. High genetic variability of cancer cells in CML is difficult for analysis and selection of prognostic markers.

In this work we perform whole exome sequencing (WES) of newly diagnosed CML patients and analyze exomic variations which can serve as early predictive markers of the high and low efficacy of TKI therapy.

## Results

### Characteristics of the patients

Twenty three patients were available with follow-up data at 6 months after beginning TKI treatment (11 males and 12 females; mean age – 42.4 years). Two patients (8.7 %) were classified as failures, 7 patients (30.4 %) were assigned to warning group and other 14 patients demonstrated optimal response according to ELN recommendations 2013. Blood samples from 8 patients were taken for further analysis: 4 optimal responders and 4 patients with TKI inefficacy (2 failures and 2 patients from the warning group). Frontline therapy was imatinib in 3 optimal responders and 3 non responders, and nilotinib in 1 responder and 1 non responder. Two failures were switched to nilotinib and dasatinib after 2.5 and 4 months of imatinib therapy. The detailed description of the patients included in the study is available in the Additional file 1: Table S1. Blood samples were drawn before starting TKI therapy when vast majority of nucleated blood cells were leukemic. Paired blood samples were drawn from 3 optimal responders after 6 months of TKI therapy during relapse for analysis of normal exomes. All 11 blood samples were sequenced.

### Whole exome sequencing

In total about 55 millions of nucleotides in targeted sequences were covered in average 20×. Sequences were aligned to hg19. Ion Torrent Variant Caller plugin was used to call variants. In each of 11 samples we found from 33,657 to 38,636 homo- and heterozygous variants (SNVs), including single nucleotide variants and insertions and deletions. All the variations are available at

European Nucleotide Archive (<http://www.ebi.ac.uk/ena/data/view/PRJEB9857>). There were about 5300 unique variants in each sample. All 116,508 variants were merged in one table for further annotation and analysis. Variants were found in 17,741 genes. In each gene we found from at least one variant in one or more samples up to 345 variants (*MUC16*). The full list of genes with number of found variants is available in Additional file 2: Table S2. Pathway enrichment analysis using Gorilla (ver. Mar 28, 2015) [8] revealed possibly damaged by SNVs cellular pathways: antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway, TAP-independent ( $p = 1.8 \times 10^{-7}$ ), interferon-gamma-mediated signaling pathway ( $p = 9.6 \times 10^{-8}$ ) and O-glycan processing ( $p = 1.9 \times 10^{-13}$ ); functions: peptide antigen binding ( $p = 1.8 \times 10^{-10}$ ), MHC class II receptor activity ( $p = 2.1 \times 10^{-7}$ ); and structures: MHC protein complex ( $p = 3.7 \times 10^{-11}$ ), integral component of lumenal side of endoplasmic reticulum membrane ( $p = 5.2 \times 10^{-11}$ ), Golgi lumen ( $p = 5 \times 10^{-13}$ ) and ER to Golgi transport vesicle membrane ( $p = 7.7 \times 10^{-10}$ ). The full list of the enrichment analysis results is available in Additional file 1: Table S3.

### Variant analysis

We were looking for variants which were present in all samples of one group and absent in all samples from the opposite group. We performed such a filtering for optimal responders and for the failures separately. In leukemic cells of all four optimal responder we found 11 333 variants in total. These variants included both homozygous and heterozygous SNPs. 152 of them were not present in either of leukemic samples from patients with failure of the response. We further filtered missense and splice-site variants and also variants in regulatory regions and we got only 27 variants. By manual verification of sequenced reads in IGV browser [9, 10] we confirmed 21 variant. Based on GeneCards descriptions (available in October, 2014) we selected 11 variants in genes known to be associated with cancers (*ANKRD35* – rs11579366, *ATP7B* – rs732774, *COSM147691*, *IGHV4-31* – rs61995642, *ANPEP* – rs25651, *DNAH9* – rs1990236, *CCDC165* – rs632423, *MAGEC1* – rs176037) or cancer-associated pathways like MAPK (*FCRL3* – rs7522061, *LRP2* – rs2229263, *KLB* – rs4975017) or cellular processes frequently deteriorated in cancer like DNA integrity control (*FANCD2* – rs3864017). We selected 3 variants (*ANKRD35* – rs11579366, *DNAH9* – rs1990236, *MAGEC1* – rs176037) positive for either SIFT [11] or PolyPhen [12] score to validate them with Sanger sequencing.

We also selected additional variant previously not described in dbSNP135 which we confirmed manually with IGV browser. It was in breast cancer associated *TOX3* (rs10653661 according to dbSNP141). This variant was also selected for confirmation by Sanger sequencing.

Another filtering option was to select only homozygotes in all 4 optimal responders. Two out of 10 of them were confirmed by manual verification. One of them was annotated in the gene involved in angiogenesis, apoptosis, activation of TGF-beta (*THSD1* – rs3803264), and was also selected.

We applied exactly the same filtering pipe-line to the set of data from leukemic cells in the group of failures. We found 8 969 variants in all four failures. Only 51 of them were not present in optimal responders. Filtering missense and splice-site variants left 15 variants. We verified manually 7 variants. GeneCards annotations did not allow to perform any specific filtering based on the gene functions – some genes seem to play important role in RNA metabolism (*DHX37* – rs11558556), cell growth and proliferation (*ADRA1A* – rs1048101, CM064954) or immune system (*PTCRA* – rs9471966) and little is known about others (*KIAA0226L* – rs1408184, *CCDC11* – rs35193847, *MORN2* – rs3099950, *TMX4* – rs2294248). Two variants (*MORN2* – rs3099950, *PTCRA* – rs9471966) with deleterious/damaging effect according to SIFT or PolyPhen were selected for further Sanger sequencing.

Scheme of all filtering steps is available in Additional file 1: Figure S1.

#### Validation of the findings with Sanger sequencing

After filtering and annotation we selected 7 variants for Sanger sequencing. Five of them (rs11579366, rs1990236, rs176037, rs10653661, rs3803264) were typical for optimal responders, two (rs3099950, rs9471966) – for failures. Sanger sequencing confirmed WES findings except one discordant result. Sanger sequencing detected heterozygote for rs10653661 in *TOX3* in one failure previously reported by WES as homozygous reference. Detailed analysis of the WES reads revealed low coverage of this particular locus and allelic variant (insertion of TTTCT) was present in the only read. Obviously it was not reported by automatic variant calling and it was not also considered during manual verification with IGV. The only read with mutation is always a very unreliable finding to be considered as a real variant.

#### Discussion

Targeted tyrosine-kinase inhibitors dramatically changed efficacy of CML treatment and overall survival rates. However there are about 20 % of CML patients failing to achieve optimal response to TKI therapy. In this work we applied WES in order to find genomic variants which could explain differences in TKI inefficacy. WES is very popular in studying cancers including myeloid neoplasia, however there are no published studies about the differences between groups of CML patients with optimal and non-optimal response to standard TKI therapy [13]. This approach has the benefits of the unbiased search of the

independent prognostic markers. CML patients receiving TKI therapy were assigned to either optimal responders or non-optimal/failures according to ELN criteria 2013. WES was performed in 8 patients and paired recovered samples. Bioinformatic analysis revealed classic long-tail distribution of the genes according to the number of variants found in them. Annotation of the top-100 mutated genes (from 41 to 345 mutations per gene) revealed that these genes are important for cellular processes and functions connected with the activity of the immune system. Knowledge about involvement of specific pathways allows developing gene-expression panels for signaling pathway profiling which may serve as an independent molecular marker [14].

However our aim was to find direct single-gene variations which can explain and predict response to the therapy. Multiple approaches can be applied to prioritize possible targets among thousands of SNVs found in individual exomes. We focused on the possible prognostic effect and expected to find variants with strong association with one of the treatment outcomes. We filtered exonic variants (27 in optimal responders and 15 in failures) which were either missense or otherwise potentially involved in gene functioning. We selected either missense variants with predicted damaging effect on protein function, or upstream/downstream regulatory variants. These manually confirmed variants in cancer-associated genes were selected for further Sanger resequencing. Below we describe all genes with found variants.

*ANKRD35* (Ankyrin Repeat Domain-Containing Protein 35) is a protein coding gene. Genecards annotated in October 2014 its association with chronic lymphocytic leukemia. This gene is expressed in normal human tissues, including bone marrow, whole blood and white blood cells but nothing is known about its function and role in cancer pathogenesis of its variants [15]. We found *g.I:145562293 G>C* substitution (rs11579366) in responders but not in failures. This variant has different frequencies of alleles in different populations. According to “The 1000 Genomes Project (Phase 3)” the allele frequencies are  $G = 0.2912$  in GBR but  $G = 0.8802$  in ACB (data obtained on April 28, 2015). It is a promising prognostic factor because with that high population frequency we though found opposite homozygotes in failures and optimal responders. This *G>C* substitution has “probably damaging” PolyPhen score (0.993).

*DNAH9* (Dynein, Axonemal, Heavy Chain 9) encodes the heavy chain subunit of axonemal dynein, a large multi-subunit molecular motor. *DNAH9* interacts with B-cell CLL/lymphoma 6 (*BCL6*). *BCL6* keeps B-cells in a proliferative state in the germinal center, and down-regulation of *BCL6* is required for B-cell maturation and differentiation. The dysregulation of *BCL6* is implicated in lymphomagenesis [16]. A total of 28 records are available in ClinVar with different variants of *DNAH9* found

in melanoma and lung cancer (April 28, 2015). Heterozygous mutation in splicing site of *DNAH9* (*NM\_001372.3:c.10242+5G>A*) was reported in one of the cases with atypical CML [17]. We found a missense variant (rs1990236) *g.17:11865462G>A*. This variant has a highly deleterious SIFT score (0.02). Its population frequency is from as low as 0.07 in YRI up to 0.3 in PEL being 0.2 in European populations.

*MAGEC1* (Melanoma Antigen Family C, 1) is a protein coding gene associated with myeloma [18], digestive tract carcinomas [19], glioblastoma [20] and melanoma [21]. It is a member of the melanoma antigen gene (MAGE) family. The proteins of this family are tumor-specific antigens that can be recognized by autologous cytolytic T-lymphocytes [22]. *MAGEC1* helps cancer cells in multiple myeloma to escape from immune-surveillance, and its expression stimulates proliferation of cancer cells [23]. The variant we found (rs176037) is the *g.X:140993642 C>T* substitution and it has a deleterious SIFT score of 0.01. T allele frequency is about 0.5 in European populations.

We also chose variant not previously described in databases even if they did not influence the coding sequence directly. *TOX3* (High Mobility Group Box Family Member 3) is the protein coding gene. It contains a HMG-box, and may bend and unwind DNA and modify chromatin structure. Its high expression is associated with poor prognosis in breast cancer [24]. Its variant rs3803662 is associated with breast cancer [25] and multiple primary cancers [26]. *TOX3* protects against cell death by inducing antiapoptotic and repressing proapoptotic transcripts. It stimulates transcription from the estrogen-responsive or *BCL-2* promoters. Among optimal responders we found one homozygote and other heterozygotes for rs10653661 – insertion of 5 nucleotides in 3'-UTR *g.16:52472421\_2insTTTCT*. This insertion was not described in dbSNP135 which we used for annotating variants before their filtration. And we chose this variant as a new, not previously described finding. This variant got its accession in dbSNP141 and population frequencies are now available for it: from 0.24 in CDX to 0.85 in YRI with 0.42–0.44 in Europeans. The gene plays an important role in cellular processes involved in cancer and its variants are associated with different cancers and diseases.

We also filtered homozygotes present in all optimal responders while all failures were wild-type homozygotes. One of the variants was annotated as upstream gene variant in *THSD1* (Thrombospondin, Type I, Domain Containing 1). It is a surface marker of hematopoietic progenitors and endothelial cells [27]. *THSD1* is methylated and down regulated in colorectal cancer [28] and esophageal squamous cell carcinoma. In all optimal responders we found rs3803264 in a homozygous form

(AA) while all failures were wild-type homozygotes (GG). This *g.13:52971893G>A* substitution is synonymous in most of the *THSD1* transcripts (NM\_018676, NM\_199263), but NM\_018676 can generate a shorter protein (uc010tgz.2 length = 473 instead of uc001vgo.3 length = 852) in which case the variant is considered as upstream gene variant. It has a high population frequency of 0.65 in Europeans.

All the above described variants were present in the same genotype in paired samples of optimal responders before the beginning and after 6 months of TKI therapy. This is why we consider these variants to be present in the constitutive genome but not leukemic-specific somatic mutations.

In failures we found only two variants following the same criteria as described above for the optimal responders.

*MORN2* (Membrane Occupation and Recognition Nexus Repeat Containing 2) is a protein coding gene. There are five MORN-repeat containing genes in humans. MORN motifs are necessary for binding junctophilins to plasma membrane and endoplasmic reticulum [29]. Human *MORN2* facilitates phagocytosis-mediated restriction of Mycobacterium tuberculosis in macrophages and is important for LC3-associated phagocytosis [30]. *MORN4* interacts with class III myosins and may be transported to filopodia of moving cells [31]. We found rs3099950 in heterozygous form in all failures but in none of the responders. This missense substitution *g.2:39109558G>A* has been registered with different frequencies in populations from 0.0 in many to 0.18 in PJL and has the frequency of ~0.14 in Europeans.

*PTCRA* (Pre T-Cell Antigen Receptor Alpha) encodes a protein that is found in immature T-cells. *PTCRA* is the member of the pre-T-cell receptor complex, which regulates early T-cell development. This gene along with many others undergoes subsequent demethylation, remethylation and repression in appropriate stages of differentiation. *PTCRA* is required for T-cell receptor rearrangement and after this stage of T-cell development it needs to be silenced in mature CD4+ cells to avoid MHC class I binding [32]. *PTCRA* is also a member of NOTCH1 signaling pathway which is an important regulator of T-cell differentiation [33]. It was shown that *PTCRA* is expressed in acute promyelocytic leukemia significantly [34]. All failures in our study were heterozygotes for rs9471966 – *g.6:42891022G>A* substitution. It has both SIFT and PolyPhen deleterious scores. Allele A has the lowest frequency in CHS – 0.01 and the highest in ESN 0.44, and with ~0.25 in Europeans. We consider it to be a good target for both prognostic purposes and explanation of therapy failure.

All seven variants have high population frequency in at least some populations. Since we are not looking for the variants associated with CML but rather with its

sensitivity to TKI treatment, high population frequency of the potential markers and dramatic difference in their frequency between optimal responders and failures is very promising for finding a highly sensitive and specific prognostic marker. We suggest that TKI therapy may be more successful for patients carrying the following markers all together or in some combination: rs11579366 (C/C or C/G), rs1990236 (G/A), rs176037 (T/T or C/T), rs10653661 (A/ATTCT or ATTCT/ATTCT), rs3803264 (A/A) and reference variants for rs3099950 (G/G) and rs9471966 (G/G). However real prognostic value of these markers and their possible combinations can be clarified only in further prospective studies.

### Conclusions

In conclusion we performed whole exome sequencing of 8 CML patients. Four of them demonstrated optimal response to standard TKI therapy and another 4 were considered as failures. We filtered out from whole exome data 7 variants (6 SNVs and one 5 nucleotide insertion) which reflect heterogeneity of treatment efficacy in patients with CML and may help to select patients with primary resistance to TKI therapy. As far as genes with the selected variants associate with other types of cancers and found variants have predicted damaging effect their possible functional effect should be carefully studied. To further validate functional role and prognostic value of the variants it is necessary to perform retrospective and prospective studies in groups of CML patients.

### Methods

Newly diagnosed patients with Ph+ CML in chronic phase at diagnosis have been included in this study at Federal State-Funded Institution National Research Center for Hematology of the Ministry of Healthcare of the Russian Federation (Moscow, Russia) starting from June, 2012. Peripheral blood samples were collected at diagnosis after signing informed consent "Exome and transcriptome analysis of leukemic cells in chronic myeloid leukemia". The study was conducted according to the provisions of the Declaration of Helsinki. All patients received standard treatment with TKIs and were examined at regular time periods – 3, 6, 9 and 12 months during the 1-st year of observation. The examination included clinical evaluation, complete blood count, cytogenetic and molecular tests which measured *BCR-ABL* expression by quantitative PCR. Patients were assigned to the groups of either optimal responders or failures according to ELN recommendations 2013 [5]. For the purpose of our study we evaluated response at 6 months. All failures were checked for the absence of mutations in the *ABL*-kinase domain of the *BCR-ABL* by sequencing. DNA extracted from blood samples was used for WES.

WES was performed using Ion Torrent Personal Genome Machine (Thermo Fisher Scientific, Waltham, MA, USA) with 200 bp reads. Exome enrichment was performed using Ion TargetSeq™ Exome Kit which captures about 30,000 protein-coding and RNA genes and predicted microRNA-binding sites. Sequencing analysis was done in Torrent Suite software v.3.6.2. Single nucleotide variants (SNVs) were annotated with ANNOVAR [35], Partek Genomics Suite (Partek, St. Louis, Missouri, USA) and Variant Effect Predictor tool [36]. Variants identified by whole exome sequencing, were validated by Sanger sequencing. Methods in more details are described in Additional file 1.

### Availability of supporting data

The data set supporting the results of this article is available in the European Nucleotide Archive repository, # PRJEB9857 <http://www.ebi.ac.uk>.

The data sets supporting the results of this article are included within the article and its additional files: Additional files 1 and 2 which are referenced the main text.

### Additional files

**Additional file 1: Supplementary Methods. Table S1.** Characteristics of patients. **Table S3.** Pathway enrichment analysis. **Figure S1.** Filtering variants. (DOC 268 kb)

**Additional file 2: Table S2.** List of Genes and number of variants in them. (CSV 1370 kb)

### Competing interests

AVL received funding of travel from Thermo Fisher Scientific Corporation. SAS received funding of travel, accommodations, or expenses from Novartis Farma and Sage Ltd. EYC served as a consultant for Novartis, participated on the Speakers' bureau of Novartis, Bristol-Myers Squibb. AGT served as a consultant/advisor, participated on the Speakers' bureau, and received honoraria from Bristol-Myers Squibb, Novartis and Pfizer. SIK participated on the Speakers' bureau of Novartis and Pfizer. All remaining authors have declared no conflicts of interest.

### Authors' contributions

AVL, EYC, AGT, SIK designed the study, AVL, SAS, EPA performed experiments, AVL, AGT, SIK analyzed results, AVL, SAS, EYC, AGT, SIK wrote the article; EYC, OAS, AGT treated patients in clinical trial and collected clinical data. All authors read and approved the final manuscript.

### Acknowledgements

We thank A.O. Abdullaev from National Research Center for Hematology (Moscow, Russia) and A.V. Misurin from Genotechnologia (Moscow, Russia) for *BCR/ABL* kinase domain sequencing and A.V. Polyakov' laboratory (Research Centre for Medical Genetics, Moscow, Russia) for Sanger sequencing. Publication of this work was supported by Russian Ministry of Education and Science (agreement # 8113 from 28/08/2012).

### Declarations

This article has been published as part of *BMC Genetics* Volume 17 Supplement 1, 2016: Selected articles from the 7th International Young Scientists School "Systems Biology and Bioinformatics" (SBB'2015): Genetics. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcgenetics/supplements/17/S1>.

**Author details**

<sup>1</sup>Laboratory of Mutagenesis, Federal State Budgetary Institution "Research Centre for Medical Genetics", Moskvorechie, 1, Moscow 115478, Russia.  
<sup>2</sup>Department of Molecular and Cellular Genetics, State Budgetary Educational Institution of Higher Professional Education "Russian National Research Medical University named after N.I. Pirogov" of Ministry of Health of the Russian Federation, Moscow, Russia. <sup>3</sup>Scientific and Advisory Department of Chemotherapy of Myeloproliferative Disorders, Federal State-Funded Institution National Research Center for Hematology of the Ministry of Healthcare of the Russian Federation, Moscow, Russia.

Published: 27 January 2016

**References**

- Deininger MW, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. Review article. 2000. p. 3343–56.
- Yin W, Penrod JR, Maclean R, Lakdawalla DN, Philipson T. Value of survival gains in chronic myeloid leukemia. *Am J Manag Care*. 2012;18:S257–64.
- Bixby D, Talpaz M. Seeking the causes and solutions to imatinib-resistance in chronic myeloid leukemia. *Leukemia*. 2011;25:7–22.
- Quintás-Cardama A, Kantarjian HM, Cortes JE. Mechanisms of primary and secondary resistance to Imatinib in chronic myeloid leukemia. 2009. p. 16.
- Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, Apperley JF, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood*. 2013;122:872–84.
- Marin D, Ibrahim AR, Goldman JM. European Treatment and Outcome Study (EUTOS) score for chronic myeloid leukemia still requires more confirmation. *J Clin Oncol*. 2011;29:3944–5.
- Yamamoto E, Fujisawa S, Hagihara M, Tanaka M, Fujimaki K, Kishimoto K, et al. European Treatment and Outcome Study score does not predict imatinib treatment response and outcome in chronic myeloid leukemia patients. *Cancer Sci*. 2014;105:105–9.
- Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. 2009;10:48.
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013;14:178–92.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–6.
- Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res*. 2001;11:863–74.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248–9.
- Smirnikhina SA, Lavrov AV, Adilgerieva EP, Turkina AG, Kutsev SI. Clinical importance of whole-exome studies using next generation sequencing in myeloid neoplasia. *Clinical Oncohematology Fundamental research and clinical practice (RUS)*. 2013;6(1):12–20.
- Borisov NM, Terekhanova N V, Aliper AM, Venkova LS, Smimov PY, Roumiantsev S, et al. Signaling pathway activation profiles make better markers of cancer than expression of individual genes. *Oncotarget* 2014; Aug 23.
- Klopocki E, Schulze H, Strauss G, Ott C-E, Hall J, Trotier F, et al. Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am J Hum Genet*. 2007;80:232–40.
- Miles RR, Crockett DK, Lim MS, Elenitoba-Johnson KSJ. Analysis of BCL6-interacting proteins by tandem mass spectrometry. *Mol Cell Proteomics*. 2005;4:1898–909.
- Spinelli R, Pirola A, Redaelli S, Sharma N, Raman H, Valletta S, et al. Identification of novel point mutations in splicing sites integrating whole-exome and RNA-seq data in myeloproliferative diseases. *Mol Genet Genomic Med*. 2013;1:246–59.
- He L, Ji J, Liu S, Xue E, Liang Q, Ma Z. Expression of cancer-testis antigen in multiple myeloma. *J Huazhong Univ Sci Technolog Med Sci*. 2014;34:181–5.
- Chen Y-T, Panarelli NC, Piotti KC, Yantiss RK. Cancer-testis antigen expression in digestive tract carcinomas: frequent expression in esophageal squamous cell carcinoma and its precursor lesions. *Cancer Immunol Res*. 2014;2:480–6.
- Akiyama Y, Ashizawa T, Komiyama M, Miyata H, Oshita C, Omiya M, et al. YKL-40 downregulation is a key factor to overcome temozolomide resistance in a glioblastoma cell line. *Oncol Rep*. 2014;32:159–66.
- Koh SS, Wei J-P, Li X, Huang RR, Doan NB, Scolyer RA, et al. Differential gene expression profiling of primary cutaneous melanoma and sentinel lymph node metastases. *Mod Pathol*. 2012;25:828–37.
- de Carvalho F, Alves VLF, Braga WMT, Xavier CV, Colleoni GWB. MAGE-C1/CT7 and MAGE-C2/CT10 are frequently expressed in multiple myeloma and can be explored in combined immunotherapy for this malignancy. *Cancer Immunol Immunother*. 2013;62:191–5.
- Atanackovic D, Hildebrandt Y, Jadcak A, Cao Y, Luetkens T, Meyer S, et al. Cancer-testis antigens MAGE-C1/CT7 and MAGE-A3 promote the survival of multiple myeloma cells. *Haematologica*. 2010;95:785–93.
- Seksenyan A, Kadavallore A, Walts AE, de la Torre B, Berel D, Strom SP, et al. TOX3 is expressed in mammary ER(+) epithelial cells and regulates ER target genes in luminal breast cancer. *BMC Cancer*. 2015;15:22.
- Elematore I, Gonzalez-Hormazabal P, Reyes JM, Blanco R, Bravo T, Peralta O, et al. Association of genetic variants at TOX3, 2q35 and 8q24 with the risk of familial and early-onset breast cancer in a South-American population. *Mol Biol Rep*. 2014;41:3715–22.
- Park SL, Caberto CP, Lin Y, Goodloe RJ, Dumitrescu L, Love S-A, et al. Association of cancer susceptibility variants with risk of multiple primary cancers: the population architecture using genomics and epidemiology study. *Cancer Epidemiol Biomarkers Prev*. 2014;23:2568–78.
- Takayanagi S, Hiroyama T, Yamazaki S, Nakajima T, Morita Y, Usui J, et al. Genetic marking of hematopoietic stem and endothelial cells: identification of the Tmtsp gene encoding a novel cell surface protein with the thrombospondin-1 domain. *Blood*. 2006;107:4317–25.
- Khamas A, Ishikawa T, Mogushi K, Iida S, Ishiguro M, Tanaka H, et al. Genome-wide screening for methylation-silenced genes in colorectal cancer. *Int J Oncol*. 2012;41:490–6.
- Takekuma H, Komazaki S, Nishi M, Iino M, Kangawa K. Junctophilins: a novel family of junctional membrane complex proteins. *Mol Cell*. 2000;6:11–22.
- Abnave P, Mottola G, Gimenez G, Boucherit N, Trouplin V, Torre C, et al. Screening in planarians identifies MORN2 as a key component in LC3-associated phagocytosis and resistance to bacterial infection. *Cell Host Microbe*. 2014;16:338–50.
- Mecklenburg KL, Freed SA, Raval M, Quintero OA, Yengo CM, O'Tousa JE. Invertebrate and vertebrate class III myosins interact with MORN repeat-containing adaptor proteins. *PLoS One*. 2015;10, e0122502.
- Rodríguez RM, Suarez-Alvarez B, Mosén-Ansorena D, García-Peydró M, Fuentes P, García-León MJ, et al. Regulation of the transcriptional program by DNA methylation during human  $\alpha\beta$  T-cell development. *Nucleic Acids Res*. 2015;43:760–74.
- Mingueneau M, Kreslavsky T, Gray D, Heng T, Cruse R, Ericson J, et al. The transcriptional landscape of  $\alpha\beta$  T cell differentiation. *Nat Immunol*. 2013;14:619–32.
- Chapiro E, Delabesse E, Asnafi V, Millien C, Davi F, Nugent E, et al. Expression of T-lineage-affiliated transcripts and TCR rearrangements in acute promyelocytic leukemia: implications for the cellular target of t(15;17). *Blood*. 2006;108:3484–93.
- Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38, e164.
- McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*. 2010;26:2069–70.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

