Individualised Proteogenomics Applied to Analysis of Amino Acid Variants in Malignant Melanoma

Marisa Schmitt[1], Tobias Sinnberg[2], Katrin Bratl[1], Claus Garbe[2], Boris Macek[1], Nicolas C. Nalpas[1]

[1] Quantitative Proteomics, University of Tuebingen, Tuebingen, Germany

[2] Division of Dermatooncology, University of Tuebingen, Tuebingen, Germany

Keywords: Proteogenomics, Cancer, Melanoma, BRAFi Resistance, Mass Spectrometry, Nucleotide Sequencing

Correspondence to:
Prof. Dr. Boris Macek
Chair, Quantitative Proteomics
Director, Proteome Center Tuebingen
Interfaculty Institute for Cell Biology
University of Tuebingen
Auf der Morgenstelle 15
72076 Tuebingen
Germany
Phone: +49/(0)7071/29-70556
E-Mail: boris.macek@uni-tuebingen.de

Dr. Nicolas Claude Nalpas
Quantitative Proteomics
Proteome Center Tuebingen
Interfaculty Institute for Cell Biology
University of Tuebingen
Auf der Morgenstelle 15
72076 Tuebingen
Germany
Phone: +49/(0)7071/29-70552
E-Mail: nicolas.nalpas@ifiz.uni-tuebingen.de

Abstract

Analysis of patient-specific nucleotide variants is a cornerstone of personalised medicine. Although only 2% of the genomic sequence is protein-coding, mutations occurring in these regions have the potential to influence protein structure and may have severe impact on disease aetiology. Of special importance are variants that affect modifiable amino acid residues, as protein modifications involved in signal transduction networks cannot be analysed by genomics. Proteogenomics enables analysis of proteomes in context of patient- or tissue-specific non-synonymous nucleotide variants. Here we developed an individualised proteogenomics workflow and applied it to study resistance to BRAF inhibitor vemurafenib in malignant melanoma cell line A375. This approach resulted in high identification and quantification of non-synonymous nucleotide variants, transcripts and (phospho)proteins. We integrated multi-omic datasets to reconstruct the perturbed signalling networks associated with BRAFi resistance, prioritise key actionable nodes and predict drug therapies with potential to disrupt BRAFi resistance mechanism in A375 cells. Notably, we showed that AURKA inhibition is effective and specific against BRAFi resistant A375 cells. Furthermore, we investigated nucleotide variants that interfere with protein modification status and potentially influence signal transduction networks. Mass spectrometry (MS) measurements confirmed variant-driven modification changes in approximately 50 proteins; among them was the transcription factor RUNX1 displaying a variant on a known phosphorylation site S(Ph)276L. We confirmed the loss of phosphorylation site by MS and demonstrated the impact of this variant on RUNX1 interactome. Our study paves the way for large-scale application of proteogenomics in melanoma.

Introduction

Accumulation of mutations is one of the hallmarks of cancer cells and malignant melanoma is a type of cancer with the highest frequency of somatic mutations (1). Recent investigations showed that mutations of signalling targets in malignant melanoma are associated with poor clinical outcome, specifically in the mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (ERK) pathway that affects abnormal cellular growth (2). The RAS/BRAF/MEK/ERK pathway is mutated to an oncogenic form in 30% of all cancers, with non-synonymous BRAF mutations in up to 50% of cutaneous melanomas (3). The predominant BRAF mutation is within the kinase domain with a single nucleotide substitution of valine to glutamic acid at amino acid 600 (4). This mutation can result in a 500-fold increased, dimerization-independent activation of BRAF, and thus leads to a constitutive activation of downstream signalling in cancer cells (3,5). Targeted inhibition of the mutated BRAF kinase with selective inhibitors like vemurafenib, dabrafenib or encorafenib (BRAFi) results in a reduction of MAPK pathway signalling (5). However, almost all patients rapidly develop resistance to BRAFi monotherapy after a period of approximately five months (2). The considerable majority of BRAF resistance development is caused by molecular or genetic alterations that lead to MAPK pathway reactivation. The identification of multiple cellular mechanisms of resistance has greatly improved the understanding of malignancy and clinical outcomes of $BRAF^{V600E}$ metastatic melanoma e.g. by the introduction of combined BRAF and MEK inhibition. However, variants that alter the corresponding protein modification status and therefore influence resistance, remain largely elusive. In addition, the precise effect of nsSNVs, insertions and deletions (InDels) and frameshift variants at the proteome and PTM level are still largely unknown.

The past decade has seen a revolution in high-throughput sequencing technologies, which provide information on DNA/RNA sequence, gene structure and expression (6). Mass spectrometry (MS)-based proteomics is experiencing a technological revolution similar to that of the high-throughput sequencing. The current state-of-the-art "shotgun" proteomics workflows are capable of routine, comprehensive analysis of proteomes (7,8) and posttranslational modifications (PTMs) such as phosphorylation (9,10). However, most of the standard proteomics approaches identify peptides and proteins by matching MS/MS spectra against protein databases derived from public repositories (e.g. UniProt) that are not "individualised", i.e. do not contain sequence information specific for an individual patient, tissue or cell line. Commonly used protein databases therefore inherently prevent identification of individual non-synonymous variants. Proteogenomics addresses this issue by combining nucleotide and protein sequencing information, thus enabling simultaneous study and integration of DNA sequence, RNA expression and splicing, protein isoform abundance, as well as localisation of protein PTMs in personalised fashion (11,12).

Here we applied an individualised proteogenomics approach to a single immortalised human melanoma cell line, A375, in its parental as well as in BRAFi resistant state, to analyse non-synonymous variants and their impact on signal transduction networks in context of acquired resistance to kinase inhibitors. Integration of matching genomics, transcriptomics and (phospho)proteomics datasets allowed the reconstruction of signal transduction networks specific to individual phenotypes. Subsequently, we were able to prioritise a number of drugs based on their disruptive potential on the signal transduction networks associated with resistance to BRAFi. We further investigated the impact of non-synonymous amino acid variants on phosphorylation sites and their putative functional effects on cell signalling network.

Results

To study the impact of amino acids variants on signal transduction networks, we selected a widely exploited melanoma cell line harbouring the $BRAF^{V600E}$ mutation, A375. The cell line was established with two different phenotypes, drug-sensitive ("S" phenotype) and -resistant ("R" phenotype) against the BRAF inhibitor vemurafenib, as described previously (13). Both phenotypes were subjected to whole exome sequencing, RNA-sequencing, as well as proteomics and phosphoproteomics evaluation (**Figure S1A**). These omics measurements were then integrated to reconstruct signalling networks that are disturbed in BRAFi-resistant A375 cells and to investigate the impact of variants altering protein modification status.

BRAFi-sensitive and -resistant A375 cells were measured with high coverage across different "omic" platforms

We investigated the mutational landscape, transcript expression, as well as abundance of proteins and phosphorylation sites in context of BRAFi resistance in A375 cells. Number and type of non-synonymous nucleotide variants detected in the whole exome sequencing (WES) analysis was similar across phenotypes (**Figure 1A** and **Table S1.1**). We observed a high overlap (93.82%) of nucleotide variants between A375 R and S (**Figure S1B**), most of which have already been reported in Cosmic and/or dbSNP databases. Variants were characterised based on the reference to alternate nucleotide variant change, which revealed a higher exchange frequency of adenine to guanine (and vice versa), as well as cytosine to thymine (and vice versa). These variants represented approximately 64% of the total nucleotide changes and were consistent across phenotypes (**Figure S1C**). These non-synonymous variants were incorporated into their corresponding protein sequences, generating several thousand additional protein sequence

isoforms contained within the individualised alternate protein sequence database. Despite this large increase in the number of protein isoform sequences, the database search space increased by only 2% (**Figure 1B**). Over 11,000 expressed genes and 42,000 transcripts were identified by RNA-sequencing using Oxford Nanopore Technology (**Figure 1C**), of which the vast majority was shared between phenotypes. The individualised alternate protein sequence database was subsequently used, together with the reference protein sequence database, for the processing of deep proteomics and phosphoproteomics data from A375 S and A375 R. High resolution MS led to the quantification of more than 9,000 protein groups (**Figure 1D**), most of which were shared between phenotypes. While the number of transcripts and protein groups was similar between phenotypes, over twice as many phosphopeptides were exclusively identified in A375 R compared to A375 S. This points to the importance of protein phosphorylation in resistance to BRAF inhibition in A375 cells.

Key signalling pathways are perturbed in BRAFi-sensitive and -resistant melanoma A375 cells

We next performed a quantitative comparison of BRAFi-sensitive versus -resistant phenotypes for A375 using the allelic frequency of nucleotide variants, transcript expression, as well as protein- and phosphorylation site abundance (Significance A with s0, see **Supplementary Information)**. Investigation of the allelic frequency in A375 R versus A375 S revealed 271 upregulated and 281 downregulated non-synonymous nucleotide variants (**Figure 2A** and **Table S1.1**). To narrow down the driver mutations in context of resistance to BRAFi, we overlapped the somatic non-synonymous nucleotide variants identified in A375 R against the study from Long and colleagues (14), who analysed somatic mutations in 10 malignant melanoma patients (**Table S2.1**). While we did not observe any shared variants between A375 R and patients' data, we did find three shared genes harbouring different somatic variants (**Figure 2B**). Of note, the overlap of somatic mutations

between patients in the Long *et al.* study was also very limited, which highlights the need for personalised approach to investigate cancer.

RNA-seq analysis identified 1,064 significantly changing transcripts between A375 R and A375 S phenotypes (613 upregulated and 451 downregulated) (**Figure 2C** and **Table S1.2**). We conducted an overrepresentation analysis of Reactome pathways using transcripts found significantly upregulated in A375 R or A375 S (**Figure 2D** and **Table S1.3**). This analysis established the involvement of CREB signalling, RSK activation and ERK/MAPK targets within A375 S, whereas signalling by PDGF, glycoprotein hormones and collagen degradation are found in A375 R.

MS analysis identified 302 significantly regulated proteins between phenotypes **(Figure 2E** and **Table S1.4)**, including key signalling proteins such as BRAF, proteins from the RAS family and several mitogen-activated protein kinases. A similar overrepresentation analysis based on significantly changing proteins showed that nucleosome assembly is overrepresented in A375 R, while pathways connected to plasma lipoproteins and immune system are overrepresented in A375 S (**Figure 2F** and **Table S1.5**). Only 22 phosphorylation sites were found differentially abundant between phenotypes, including HMGA2 and RUNX1 (**Figure S2A** and **Table S1.6**).

Interestingly, the comparison between transcriptome and proteome revealed a low correlation in abundance ratios of A375 R versus A375 S (**Figure S2B**). Only seven transcript/protein pairs were found significantly changing in the same direction, including LIMA1, SMTN and SDCBP. Taken together, our data show that transcriptomics and proteomics provided distinct functional information in context of BRAFi-resistance, thus emphasising the need for proteogenomics to characterise the underlying signalling pathways.

Amino acid variants are detectable within key signalling pathways and proteins

The aforementioned individualised protein database, used during the processing of the MS data, allowed us to detect a number of variant peptides, i.e. peptides harbouring reference or alternate amino acid. More amino acid variants were detected in drug-resistant cells compared to drug-sensitive cells on the proteome and phosphoproteome level (**Figure S2C**). To address whether mutational hotspots occur within specific pathways, we performed pathway overrepresentation based on identified protein isoforms harbouring amino acid variants (**Figure S2D**). Within drug-resistant cells, VEGF and ERbB signalling pathways were exclusively overrepresented for variant protein isoforms, whereas HIPPO signalling pathway, Rho GTPase cycle and lysosome were specific for A375 S cells.

Next, we investigated the most commonly mutated genes in melanoma or in context of BRAFi resistance (**Figure S2E**). Half of these genes were found to harbour at least one nucleotide variant. However, only the BRAF V600E variant could be identified at both the WES and proteome levels. Most of the commonly mutated genes were quantified at the transcriptome and proteome levels. Protein abundance changes based on label-free quantification revealed non-significant trends between phenotypes. For example, KMTD2 had a higher intensity in drug-sensitive A375 S cells, whereas ERBB2 showed a higher abundance in A375 R cells. Among the resistance marker genes found within the Long et al. study, only MAP2K3 (MEK3) was found mutated in our dataset (R26T). Among these most commonly mutated proteins in melanoma, several were phosphorylated on known as well as new modification sites (**Table S1.6**). These results highlight the capability of proteogenomic approach to detect amino acid variants on expressed proteins, which in turn reveal an overrepresentation of key cell signalling pathways.

The perturbed signalling network in BRAFi resistant A375 cells can be targeted by several drugs

We next integrated the significantly changing non-synonymous nucleotide variants, transcripts, proteins and phosphorylation sites into a protein-protein interaction network (**Figure 3A, S3A** and **Table S1.7**). Only the top 200 entries were retained on the basis of their number of connections obtained from the BioGRID database. A functional annotation analysis of these 200 entries revealed an overrepresentation in proteins involved in several immune system pathways, several general cancer-related pathways, as well as MAPK, Ras and PI3K/Akt signalling pathways (**Figure S3B** and **Table S1.8**). Interestingly, less than 50% of the entries in this signalling network harbour one or more amino acid variants (**Figure 3A**). The size of the entries scales up with the impact factor obtained through our custom bioinformatic workflow and represents their importance in context of cancer, melanoma and melanoma resistance (see **Supplementary Information**). A list of drugs was retrieved from DrugBank database and their targets highlighted within this network. Several entries were highlighted through this approach, such as AURKA, BRAF and SULT1A1; all of them are known in context of cancer, are mutated in A375 cells, can be targeted by drugs and are upregulated in A375 R (protein level).

Potential drugs were prioritised based on their target specificity, as well as on the number of degrees (connections) their targets have with the rest of the signalling network (**Figure 3B**). Among the drugs that have the potential to disrupt this resistance-specific signalling network were Alisertib and AT9283 to inhibit AURKA, FG-2216 and PX-478 to target HIF1A, and Tanespimycin to suppress HSP90. We experimentally validated the action of the compound Alisertib on BRAFi-resistant and -sensitive A375 cell lines (**Figure 3C** and **S3C**). BRAFi-resistant A375 cells were sensitive to AURKA inhibition (AURKAi) with Alisertib, regardless of the absence/presence of BRAFi vemurafenib. Conversely, A375 S tolerated Alisertib (in absence of

BRAFi) and were able to proliferate compared to A375 R cells. Our results highlight the effectiveness and specificity of AURKAi against BRAFi-resistant A375 cells and demonstrates that the integration of proteogenomic datasets has potential to predict effective and individualised drug therapy in context of BRAFi resistance.

Multiple amino acid variants directly affect protein phosphorylation status

We subsequently focused on MS-identified amino acid variants based on their reported involvement in cancer and melanoma, as well as their effect on protein phosphorylation status (**Figure S3D** and **Table S1.9**). We identified 44 variant protein isoforms that harboured a phosphorylation event on (or in instant vicinity of) the variant site in A375 cells (**Figure 3D and S3E**). For the variant directly affecting the phosphorylation site, these comprised both, reference variant peptides that lost the phosphorylation site, and alternate variant peptides that gained a phosphorylation site. Notably, among the phosphorylated alternate peptides, more than half displayed a tendency for higher abundance in the resistant phenotype. Several of these variants were among the top impact ranked variants due to their involvement in melanoma and effect on phosphorylation status (**Figure S3E**). While, these phosphorylated variant peptides were not significantly changing in abundance between A375 R and S, we speculate that they may still have a functional effect, which might not necessarily be connected to BRAFi resistance.

Loss of a phosphorylation site on RUNX1 alters its interactome and transcriptional activity

We investigated further one variant with high impact in context of cancer, melanoma and phosphorylation status that was found on the RUNX1 protein, a key transcription factor involved in cell proliferation, differentiation and apoptosis (15). This amino acid variant led to the loss of a known phosphorylation site due to change from serine to leucine at position 276 (**Figure 4A**). The

reference and alternate variant peptides were identified with high resolution MS in both A375 S and R cells (**Figure S4A** to **C**). The phosphorylation site is located in a highly modified region in close proximity to the transcription activation domain of RUNX1, previously reported to be involved in binding of key regulatory proteins, such as P300 (16). We hypothesised that this variant is likely to influence the interactome of RUNX1. Therefore, we generated a RUNX1 gene knockout in A375 S cells using the CRISPR/Cas9 system. Single cell clones were selected for further analysis based on their effective RUNX1 knockout (KO). As a control we used a non-targeting (NonTar) control guide sequence. A375 RUNX1 KO cells showed an insertion of 215 bp in the Exon 1 of the gene compared to reference DNA of A375 S cells (**Figure S4E**). The lack of expression of RUNX1 protein was confirmed by western blot and MS analysis (**Figure S4D**). To study the impact of the loss of a modifiable amino acid, we performed immunoprecipitation of Flag-tagged RUNX1_wt and RUNX1_S276L in RUNX1 KO SILAC labelled cells in three independent replicates (**Figure 4B** and **Table S3.1**). The interactome analysis by LC-MS/MS revealed that RUNX1 and its core binding factor CBFB were significantly enriched in both pulldowns compared to Flag-empty vector (**Figure S4F and G**). Interestingly, the known interaction partner histone deacetyltransferase HDAC1 was enriched in RUNX1_wt interactome and depleted in the RUNX1_S276L interactome (**Figure 4B**).

To confirm these findings, we performed pulldown assays with synthetic peptides harbouring the amino acid sequence for reference and alternate variant peptides of RUNX1 in A375 cells (**Figure 4C** and **Table S3.2**). As in the interactome study, HDAC1 was significantly depleted in the pulldown of alternate versus reference variant peptide indicating that the interaction between HDAC1 and RUNX1 is disturbed due to the variant. We also identified transcriptional repressor SIN3A to be significantly depleted in alternate variant peptide pulldown compared to reference

pulldown similar to HDAC1. RIN1 and PTPN23 showed the same trend as HDAC1 and SIN3A; and both proteins are known to act as regulator of RAS-mediated mitogenic activity (17,18). The proteins enriched in alternate pulldown compared to reference variant peptide pulldown were overrepresented in TGFβ signalling, melanogenesis and insulin signalling pathways (**Figure S4H** and **Table S3.3**). Taken together, we demonstrate that the loss of this known phosphorylation site has an impact on the interactome of RUNX1 (**Figure S4I**) and postulate that it leads to altered transcriptional activity (**Figure 4D**).

Discussion

In this study, we applied an individualised onco-proteogenomics approach to a widely used cell line model of melanoma (A375), a cancer that is well known for its high mutation load (1) and the potential for rewiring cellular networks (12,19). Two consortia, namely the Clinical Proteomic Tumor Analysis Consortium (CPTAC) and The Cancer Genome Atlas (TCGA), have greatly contributed to the development of onco-proteogenomics (11,20,21). However, proteogenomics studies are still relatively rare and, due to their complexity, out of reach of most proteomics (or genomics) laboratories. Here, we are going beyond state-of-the-art proteogenomics workflow, through the analysis of a single melanoma cell line in order to generate truly individualised multi-omics abundance measurement and to predict tumour-specific drug therapies.

Proteogenomic reconstructs the signalling network linked to BRAFi resistance in A375 cells

In this study, the nucleotide variant incorporation revealed very similar numbers (~10,000 variants) across A375 phenotypes, the large majority being SNVs, which is consistent with a previous study (22). We also observed characteristic nucleotide substitutions, whereby two thirds of substitutions are comprised of transitions and the rest being transversions. The C to T transition

was highly represented and is known to result from sun-light exposure, which is highly relevant for skin cancer (23). The proteome coverage we obtained for A375 cells (~9,000 protein groups) is similar to other state-of-the-art MS-based study of cancer cell lines (24). The differentially changing entries identified through our multi-omics approach were overrepresented in pathways of critical importance for melanoma resistance to BRAFi, including collagen degradation, CREB signalling, RSK activation and MAPK, Ras and PI3K/Akt signalling pathways (25). Another aspect of pivotal importance was the distinct functional information provided by the transcriptome and proteome datasets. This agrees with previous study (24) and greatly emphasises the importance of proteogenomic approaches as opposed to single-omic investigations.

We then reconstructed the signalling network associated with BRAFi-resistance in A375; i.e. using the significantly changing non-synonymous nucleotide variants, transcripts, proteins and phosphorylation sites. This approach highlighted several hubs and high impact entries, such as CUL3, XPO1, AURKA, HSPA5, H2AFX and HIF1A, of which several lead to reduced melanoma patient survival or are known for their involvement in melanoma development (26-28). Several drugs were identified and ranked based on their potential to disrupt this signalling network; notably Alisertib, a highly specific inhibitor of AURKA, which has been previously reported for its beneficial effect in combination with BRAF and MEK inhibitors in melanoma treatment (27). We experimentally validated the use of Alisertib on A375 S and A375 R cells and could show that A375 R cells are sensitive to AURKAi regardless of the absence/presence of BRAFi. Our data confirm that AURKA has a critical role in the context of resistance and may be suitable for the treatment of melanoma as reported previously (29). In addition, the prediction and subsequent validation of AURKAi as specific drug therapy to treat BRAFi resistant A375 cells verify the use of our individualised proteogenomics approach.

Proteogenomic pinpoints several peptides that are phosphorylated on the variant site

Because of the higher number of phosphorylation sites within A375 R, we investigated further the amino acid variants that affect phosphorylation status and are important in melanoma development. Around 14.8% of all amino acids in the human proteome are serine, threonine or tyrosine (30), which are predominantly modified by phosphorylation. Several studies have reported that these three amino acids are disproportionally affected by missense mutations (31,32). While these may not all be relevant in tumour cells, since not all genes are expressed at any one-time, previous studies have shown the deleterious effect of such variants (11,12). Among the variant protein isoforms with impact score medium to high**,** two thirds were known oncogenes or tumour suppressors genes; and more than 80% had a loss of known phosphorylation sites.

We identified hundreds of amino acid variants by high resolution mass spectrometry, including some that led to a change in the modification status of the protein. Our identification results are in the same range (or higher) as other studies investigating amino acid variants using custom protein sequence databases (33,34). Overrepresentation of variant protein isoforms revealed their accumulation in signalling by VEGF and ERbB signalling pathways in drug-resistant cells, whereas HIPPO signalling pathway and Rho GTPase cycle were overrepresented in sensitive cells. The overrepresented pathways common to both phenotypes, such as MAPK signalling pathway, are known to be highly activated in melanoma cells with acquired resistance (35). Interestingly, most of the MS-identified amino acid variants affecting the phosphorylation status were displaying a trend towards higher abundance in A375 R. Such discrepancy may be connected to BRAFi resistance in A375 cells as reported in other studies (36,37). Among interesting variant, RUNX1 was one of the top impact scoring entries in A375 cells due to its known role in cancer and loss of phosphorylation site at position S276.

Rewiring of signal transduction network due to loss of a known phosphorylation site on RUNX1

We experimentally validated this striking example of a loss of a known phosphorylation sites on RUNX1 and showed that this variant has an impact on the interactome of RUNX1. The transcription factor RUNX1 is mutated in 3.03% of melanoma patients and so far, 43 mutations are described in the literature for cutaneous melanoma. The variant site S276L of RUNX1 is located in a highly modified region of the protein and may influence the nearby transcriptional activation domain. Wee *et al.* showed *in vitro* that the triple phosphorylation at the sites S249, T273 and S276 are important for the interaction with the histone acetyltransferase p300 and thus lead to the regulation of gene transcription via chromatin remodelling (38). Here, we could not identify p300 in the interactome studies of RUNX1 by immunoprecipitation of overexpressed RUNX1 or synthetic peptide pulldowns. However, we identified the transcriptional activator WWTR1 (TAZ) and KAT7 and the corresponding transcriptional repressors HDAC1 and SIN3A to be changing between reference and alternate pulldown of RUNX1. The loss of the interaction to HDAC1 by mutating RUNX1 at S48, S303 and S424 to aspartic acid *in vitro* has been described previously (39). Here, we hypothesised that the interaction is associated with the modification status of the protein. The crosstalk between acetylation/deacetylation mediated and phosphorylation/dephosphorylation may alter the transcriptional activity by RUNX1**.** It is well known in the literature that RUNX1_wt switches between active and repressive state due to modifications like acetylation and phosphorylation and binding of interaction partners like HDAC1, PRMT1 and P300 (40). We postulate that RUNX1_S276L remains in the active state due the loss of binding to transcriptional repressors like HDAC1 and SIN3A, which could lead to the accumulation of acetylation on the protein itself as well as histones. This may result in stronger transcriptional activity, which should be tested in further experiments. Taken together, we

postulate that this variant, which influence the modification status of the protein, changed the interactome of RUNX1 and altered the transcriptional activity of RUNX1.

Conclusions

Proteogenomics is a powerful tool to study the mode of action of disease-associated mutations at the genome, transcriptome, proteome and PTM level. Here, we applied a proteogenomics workflow to study the melanoma cell line A375 sensitive and resistant to BRAF inhibition. The investigation and integration of multi-omic datasets allowed us to reconstruct the perturbed signalling networks associated with BRAFi resistance. This resulted in the prioritisation of key actionable nodes and the prediction of drug therapies with the potential to disrupt BRAFi resistance mechanism in A375 cells. Notably, we demonstrated the use of AURKA inhibitor as an effective and specific drug against BRAFi resistant A375 cells. We also detected the loss or gain of several phosphorylation events due to variants. We could confirm the loss of Ser276 phosphorylation site by MS as a direct consequence of variant S276L on the transcription factor RUNX1. Our results suggest that this mutation has an impact on the interactome of RUNX1 and may be responsible for change in its transcriptional activity. We believe that such proteogenomics workflow is readily applicable to other types of cancer and within cell lines as well as patient-derived samples.

Materials and Methods

Only star methods are presented below; the rest of the methods is described fully in **Supplementary Information.**

Cell culture

The human metastatic BRAF^{V600E}-mutated melanoma cell line A375 (CRL-1619, ATCC) was used in this study and authenticated by Microsynth AG. The generation of the cell line with acquired resistance to vemurafenib analogue PLX4720 (Selleckchem) (for simplicity referred to as "vemurafenib" in the Results section) was conducted as described previously (13). A375 S and R cells were grown in RPMI medium (Sigma-Aldrich) supplemented with FBS (10%, PAN Biotech) and penicillin/streptavidin (100 U/ml, PAN Biotech) at 37°C and 5% CO2.

For immunoprecipitation assays, SILAC-labelling of cells was performed as described previously (41) and detailed description of labelling of cells, CRISPR/Cas9-mediated knockout of RUNX1 and interaction assays can be found in the **Supplementary Information**.

Incorporation of non-synonymous variants into protein databases

To integrate the proteogenomic datasets, we used an in-house bioinformatics pipeline, which is coded entirely in the R programming language (42). The transcript nucleotide sequences were extracted from GRCh38 *H. sapiens* genome assembly and Ensembl transcript annotation (via BSgenome and GenomicFeatures packages). These sequences were then *in silico* translated (from start to first stop codon) into a reference protein sequences database (Biostrings package). The called variants, within Variant Call Format (VCF) files from A375 R and A375 S, were injected into each overlapping reference transcript nucleotide sequences and then *in silico* translated. The

resulting protein sequences were written into two FASTA files containing reference variant protein sequences and sample-specific alternate variant protein sequences.

Annotation of the biological impact of detected variants

In the current study, we prioritised amino acid variants based on their impact in context of BRAFi resistance in melanoma. For this purpose, known variant sites in melanoma, as well as known variant sites in cancer, were obtained from CGDS (43). These were overlapped with A375 identified variants and classified as loss/gain of sites. A list of oncogenes and tumour suppressor genes was compiled from Cosmic, ONGene, Bushman lab and Uniprot (30,44,45), whereas a list of genes harbouring somatic mutation in BRAFi resistant tumour was retrieved after reanalysis of published study (14). A375 variants found on these genes were annotated as relevant in cancer and/or BRAFi resistance.

A second impact scoring strategy was also performed to investigate protein phosphorylation-based signal transduction networks in A375 melanoma cells. Each reference/alternate variant protein sequence was annotated based on whether phosphorylation sites (S/T/Y) were lost and/or gained (IRanges package). A list of known kinase motifs was retrieved from PhosphoNetworks (46) and these motifs were searched along the reference/alternate variant protein sequences. Located kinase motifs were overlapped with variants position to determine loss/gain of the motifs. Known human phosphorylation sites were retrieved from PhosphoSitePlus and Phospho.ELM databases (47,48). This study identified variants, which affected these sites, were annotated as loss/gain of phosphorylation. In a similar fashion, known variant sites in melanoma were obtained from CGDS (43) and overlapped with the variants from A375 R and S. A list of oncogenes and tumour suppressor genes was compiled from Cosmic, ONGene, Bushman lab and Uniprot (30,44,45). Variants on these genes that were identified in A375 R and S were annotated as cancer-relevant.

A Levenshtein similarity score was calculated between reference and alternate variant protein sequences, whereby alternate sequences with less than 90% similarity to their reference were flagged.

Each amino acid within variant protein sequences were attributed a "+1" score for every overlap with an impact annotation. A summed score was then calculated for each amino acid within alternate variant sequence, and the maximum summed score was reported for that variant protein isoform. Because the score depends on the number of impacts used during the annotation, we also computed a scaled maximum score (between 0 and 1), to allow comparison between processings. Following the computation of all impacts, each variant protein isoform is ranked to allow prioritisation for follow up studies.

Extraction and digestion of proteins

Cells were harvested at 80% confluence with lysis buffer (6 M urea, 2 M thiourea, 60 mM Tris pH 8.0) complemented with protease (complete Mini EDTA-free tablets, Roche) and phosphatase inhibitors (5 mM glycerol-2-phosphate, 5 mM sodium fluoride, and 1 mM sodium orthovanadate) and 1% N-ocetylglucoside (NOG, Sigma-Aldrich) for 10 min on ice. DNA and RNA were removed from the cell lysate using benzonase (1 U/ml, Merck Millipore) for 10 min on room temperature (RT). Cell debris was cleared by centrifugation (2,800 xg, 10°C, 20 min). Proteins were precipitated from cell lysates using eight volumes of acetone (-20°C) and one volume of methanol and incubated overnight at -20°C. The resulting solution was centrifuged (2,800 xg, 10°C, 20 min) to form a cell pellet. The pellet was washed two times with 80% acetone (-21°C) and resuspended in lysis buffer without NOG. Protein concentration was measured using Bradford assay. 3 mg of extracted proteins were reduced with 10 mM of dithiothreitol (DTT) for 1 h, alkylated with 55 mM iodoacetamide for an additional hour and digested with Lys-C (Lysyl

Endopeptidase, Wako Chemicals) for 3 h at RT. After adding four volumes of 10 mM ammonium bicarbonate, proteins were digested with trypsin (Promega Corporation) overnight. To stop the digestion, 1% TFA was added.

Detailed description of high pH reverse phase chromatography can be found in the **Supplementary Information.**

Phosphopeptide enrichment

Phosphopeptides were enriched using TiO_2 beads (Titansphere, 10 µm, GL Sciences). 1 mg of beads (in 80%, 1% TFA) were added to acidified high pH fractions and incubated for 30 min in a rotation wheel. Phosphopeptide-bound TiO_2 beads were sequentially washed with 30% ACN, 1% TFA, followed by 50% ACN, 1% TFA and 80% ACN, 1% TFA. Peptides were eluted with 5% NH_4OH into 20% TFA followed by 80% ACN in 1% FA. The eluate was reduced by vacuum centrifugation, pH was adjusted to < 2.7 with TFA and peptides were desalted on C18 StageTips prior LC-MS/MS measurements.

Liquid chromatography - mass spectrometry

Peptides were measured on an EASY-nLC 1200 ultra-high-pressure system (Thermo Fisher Scientific) coupled to a quadrupole Orbitrap mass spectrometer (Q Exactive HF and HFX, Thermo Fisher Scientific, USA) via a nanoelectrospray ion source. About 1 µg of peptides was loaded on a 20-cm analytical HPLC-column (75 µm ID PicoTip fused silica emitter (New Objective); in-house packed using ReproSil-Pur C18-AQ 1.9-µm silica beads (Dr Maisch GmbH)). Peptides were eluted using a 90 min gradient for proteomic, phosphoproteomic and synthetic peptide pulldown studies and 60 min gradients for RUNX1 interaction studies. Gradient was generated by solvent A (0.1% FA) and solvent B (80% ACN in 0.1% FA) and 200 nl/min. Column temperature was kept

at 40°C. The mass spectrometer was operated in data-dependent mode, collecting MS spectra in the Orbitrap mass analyzer (60,000 resolution, 300–1650 *m/z* range) with an automatic gain control (AGC) target of 3E6 and a maximum ion injection time of 25 ms. For higher-energy collisional dissociation (HCD), the 12 most intensive peptides were selected and MS/MS spectra were recorded with a resolution of 30,000 (fill time 45 ms). For phosphoproteomic studies, top7 method was used with a resolution of 60,000 for HCD scans and maximum fill time of 220 ms. For the analysis of RUNX1 interactome (measured on Q Exactive HFX), full MS were acquired in the range of 300 - 1750 *m/z* at a resolution of 60,000 (fill time 20 ms). Twelve most abundant precursor ions from a survey scan were selected for HCD fragmentation (fill time 110 ms) and MS/MS spectra were acquired at a resolution of 30,000 on the Orbitrap analyzer. Precursor dynamic exclusion was enabled with a duration of 20 s. Synthetic peptide pulldowns were analysed with a top7 method with a resolution of 60,000 and a fill time of 110 ms.

Mass spectrometry data processing

The raw data files were processed with the MaxQuant software suite (version 1.6.8.0 and 1.5.2.8) (49). The Andromeda search engine searched MS/MS data against *H. sapiens* reference (99,354 entries) and cell line-specific alternate databases (A375 = 29,104 entries), as well as UniProt *H. sapiens* (release 2019/02/13; 95,943 entries) database and commonly observed contaminants. Carbamidomethylation of cysteine (C) was set as fixed modification and oxidation of methionine, phosphorylation at serine, threonine or tyrosine were defined as variable modifications. Trypsin/P was selected as a protease. No more than two missed cleavages were allowed. The MS tolerance was set at 4.5 ppm and MS/MS tolerance at 20 ppm for the analysis using HCD fragmentation method. The false discovery rate (FDR) for peptides and proteins was set to 1%. For label-free quantification, a minimum of one peptide was required. For quantification of proteins in the

immunoprecipitation experiments, the amino acids (Lys4)/(Arg6) and (Lys8)/(Arg10) were defined as 'medium' and 'heavy' labels for the comparison of RUNX1 overexpressed cell lines. For all other parameters, the default settings were used.

Proteogenomics integration

Our in-house proteogenomic bioinformatic pipeline was used to integrate WES and MS datasets, specifically to check which mutations were identified across omic datasets. Initially, the reference and alternate variant protein sequences were *in silico* digested according to laboratory condition; i.e. digestion with trypsin and up to two missed cleavages (cleaver package). The overlap of MS-identified peptides with *in silico* digested peptides led to classification into reference (non-mutated peptide that overlap the mutation position on reference protein), alternate (mutated peptide that overlap the mutation position on alternate protein) or unspecific (non-mutated peptide that does not overlap any mutated positions on reference protein) variant peptides. On the basis of this peptide classification, we summarised the peptides identification per variant protein isoforms, allowing coverage characterisation into reference only, alternate only, reference and alternate or unspecific. We finally focused on PTM (as implemented in the MaxQuant processing), which here consists in phosphorylation sites. Reference and/or alternate variant peptides found phosphorylated were flagged as such, as well as those were the phosphorylation occurred directly on the variant sites (either on reference or alternate variant sequences). This coverage information is exported within MaxQuant style processing results (tab-separated file as output).

Subsequently, we reconstructed the network of protein-protein (using BioGRID database) and drug-target (using DrugBank database) interactions; i.e. signalling network of BRAFi resistant A375 and RUNX1 interactome network (50,51). The specificities of the drugs, interacting with nodes from the generated network, were calculated based on all possible target reported in

DrugBank database. Drugs were prioritised further by summing the number of interactions their targets have within the network. The generated networks were exported (using igraph and RCy3 packages) into Cytoscape for further customisation (52).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

M.S., T.S., B.M. and N.C.N. designed the study. M.S. performed the proteomics experiments, while K.B. helped with the interaction proteomic screen. M.S. and N.C.N. analysed the data and performed statistical analysis. M.S. and N.C.N. wrote the manuscript with the input from all authors.

Acknowledgements

The authors acknowledge Prof. Dr. Yulia Skokowa for fruitful discussions, Prof. Dr. Stefan Stefanovic for synthetic peptides, Prof. Dr. Stefano Stifani for the pCMV_RUNX1 expression plasmid, Dr. Karsten Krug and Lena Thiess for their help in the initial stages of the project, as well as c.ATG Core Facility in Tuebingen for the WES library preparation and sequencing. This work was supported by the High Performance and Cloud Computing Group at the Center for Data Processing of the University of Tuebingen, the state of Baden-Wuerttemberg (bwHPC), Deutsches Konsortium für Translationale Krebsforschung (DKTK), German Research Foundation (DFG) grants No. INST 37/935-1 and INST 37/741-1 FUGG (to B.M.), and by intramural funding from the University of Tuebingen for the promotion of junior researchers (to N.C.N).

Data availability

The high throughput nucleotide sequencing data have been deposited in the NCBI Sequence Read Archive (53) with the bioproject accession number PRJNA616103. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (54) partner repository with the dataset identifier PXD018305. The WES bioinformatics pipeline is available online (55).

References

1. Rajasagi M, Shukla SA, Fritsch EF, Keskin DB, DeLuca D, Carmona E*, et al.* Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. Blood **2014**;124:453-62
2. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J*, et al.* Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N Eng J Med **2011**;364:2507-16
3. Flaherty KT, Robert C, Hersey P, Nathan P, Garbe C, Milhem M*, et al.* Improved Survival with MEK Inhibition in BRAF-Mutated Melanoma. New England Journal of Medicine **2012**;367:107-14
4. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S*, et al.* Mutations of the BRAF gene in human cancer. Nature **2002**;417:949-54
5. Allen EMV, Wagle N, Sucker A, Treacy D, Goetz EM, Place CS*, et al.* The genetic landscape of clinical resistance to RAF inhibition in metastatic melnaoma. Cancer Discov **2014**;4:94-109
6. Mardis ER. A decade's perspective on DNA sequencing technology. Nature **2011**;470:198-203
7. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R*, et al.* A draft map of the human proteome. Nature **2014**;509:575-81
8. Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM*, et al.* Mass-spectrometry-based draft of the human proteome. Nature **2014**;509:582-7
9. Macek B, Mann M, Olsen JV. Global and site-specific quantitative phosphoproteomics: principles and applications. Annual review of pharmacology and toxicology **2009**;49:199-221
10. Olsen JV, Mann M. Status of large-scale analysis of post-translational modifications by mass spectrometry. Molecular & cellular proteomics : MCP **2013**;12:3444-52
11. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P*, et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. Nature **2016**;534:55-62
12. Creixell P, Schoof EM, Simpson CD, Longden J, Miller CJ, Lou HJ*, et al.* Kinome-wide Decoding of Network-Attacking Mutations Rewiring Cancer Signaling. Cell **2015**;163:202-17
13. Sinnberg T, Makino E, Krueger MA, Velic A, Macek B, Rothbauer U*, et al.* A Nexus Consisting of Beta-Catenin and Stat3 Attenuates BRAF Inhibitor Efficacy and Mediates Acquired Resistance to Vemurafenib. EBioMedicine **2016**;8:132-49
14. Long GV, Fung C, Menzies AM, Pupo GM, Carlino MS, Hyman J*, et al.* Increased MAPK reactivation in early resistance to dabrafenib/trametinib combination therapy of BRAF-mutant metastatic melanoma. Nat Commun **2014**;5:5694
15. Otto F, Lübbert M, Stock M. Upstream and downstream targets of RUNX proteins. Journal of cellular biochemistry **2003**;89:9-18
16. Kitabayashi I, Yokoyama A, Shimizu K, Ohki M. Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation. EMBO J **1998**;17:2994-3004
17. Hu H, Bliss JM, Wang Y, Colicelli J. RIN1 is an ABL tyrosine kinase activator and a regulator of epithelial-cell adhesion and migration. Curr Biol **2005**;15:815-23
18. Doyotte A, Mironov A, McKenzie E, Woodman P. The Bro1-related protein HD-PTP/PTPN23 is required for endosomal cargo sorting and multivesicular body morphogenesis. Proc Natl Acad Sci U S A **2008**;105:6308-13
19. Sheynkman GM, Shortreed MR, Cesnik AJ, Smith LM. Proteogenomics: Integrating Next-Generation Sequencing and Mass Spectrometry to Characterize Human Proteomic Variation. Annu Rev Anal Chem (Palo Alto Calif) **2016**;9:521-45
20. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z*, et al.* Proteogenomic characterization of human colon and rectal cancer. Nature **2014**;513:382-7

21. Krug K, Jaehnig EJ, Satpathy S, Blumenberg L, Karpova A, Anurag M*, et al.* Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. Cell **2020**
22. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM*, et al.* An integrated map of genetic variation from 1,092 human genomes. Nature **2012**;491:56-65
23. Pfeifer GP, You YH, Besaratinia A. Mutations induced by ultraviolet light. Mutat Res **2005**;571:19-31
24. Nusinow DP, Szpyt J, Ghandi M, Rose CM, McDonald ER, 3rd, Kalocsay M*, et al.* Quantitative Proteomics of the Cancer Cell Line Encyclopedia. Cell **2020**;180:387-402 e16
25. Tian Y, Guo W. A Review of the Molecular Pathways Involved in Resistance to BRAF Inhibitors in Patients with Advanced-Stage Melanoma. Med Sci Monit **2020**;26:e920957-e
26. Kim H, Frederick DT, Levesque MP, Cooper ZA, Feng Y, Krepler C*, et al.* Downregulation of the Ubiquitin Ligase RNF125 Underlies Resistance of Melanoma Cells to BRAF Inhibitors via JAK1 Deregulation. Cell Rep **2015**;11:1458-73
27. Caputo E, Miceli R, Motti ML, Taté R, Fratangelo F, Botti G*, et al.* AurkA inhibitors enhance the effects of B-RAF and MEK inhibitors in melanoma treatment. Journal of Translational Medicine **2014**;12:216
28. Cerezo M, Rocchi S. New anti-cancer molecules targeting HSPA5/BIP to induce endoplasmic reticulum stress, autophagy and apoptosis. Autophagy **2017**;13:216-7
29. Shang Y-Y, Yao M, Zhou Z-W, Jian C, Li X, Hu R-Y*, et al.* Alisertib promotes apoptosis and autophagy in melanoma through p38 MAPK-mediated aurora a signaling. Oncotarget **2017**;8:107076-88
30. UniProt Consortium T. UniProt: the universal protein knowledgebase. Nucleic Acids Res **2018**;46:2699
31. Gentile S, Martin N, Scappini E, Williams J, Erxleben C, Armstrong DL. The human ERG1 channel polymorphism, K897T, creates a phosphorylation site that inhibits channel activity. Proc Natl Acad Sci U S A **2008**;105:14704-8
32. Reimand J, Bader GD. Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. Mol Syst Biol **2013**;9:637
33. Wen B, Xu S, Zhou R, Zhang B, Wang X, Liu X*, et al.* PGA: an R/Bioconductor package for identification of novel peptides using a customized database derived from RNA-Seq. BMC Bioinformatics **2016**;17:244
34. Wang D, Eraslan B, Wieland T, Hallstrom B, Hopf T, Zolg DP*, et al.* A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol Syst Biol **2019**;15:e8503
35. Poulikakos PI, Persaud Y, Janakiraman M, Kong X, Ng C, Moriceau G*, et al.* RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). Nature **2011**;480:387-90
36. Parker R, Vella LJ, Xavier D, Amirkhani A, Parker J, Cebon J*, et al.* Phosphoproteomic Analysis of Cell-Based Resistance to BRAF Inhibitor Therapy in Melanoma. Front Oncol **2015**;5:95-
37. Corazao-Rozas P, Guerreschi P, André F, Gabert P-E, Lancel S, Dekiouk S*, et al.* Mitochondrial oxidative phosphorylation controls cancer cell's life and death decisions upon exposure to MAPK inhibitors. Oncotarget **2016**;7:39473-85
38. Wee HJ, Voon DC, Bae SC, Ito Y. PEBP2-beta/CBF-beta-dependent phosphorylation of RUNX1 and p300 by HIPK2: implications for leukemogenesis. Blood **2008**;112:3777-87
39. Guo H, Friedman AD. Phosphorylation of RUNX1 by cyclin-dependent kinase reduces direct interaction with HDAC1 and HDAC3. J Biol Chem **2011**;286:208-15
40. Brettingham-Moore KH, Taberlay PC, Holloway AF. Interplay between Transcription Factors and the Epigenome: Insight from the Role of RUNX1 in Leukemia. Front Immunol **2015**;6:499
41. Schmitt M, Sinnberg T, Nalpas NC, Maass A, Schittek B, Macek B. Quantitative proteomics links the intermediate filament nestin to resistance to targeted BRAF inhibition in melanoma cells. Molecular & cellular proteomics : MCP **2019**

42. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
43. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA*, et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov **2012**;2:401-4
44. Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A*, et al.* The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. Br J Cancer **2004**;91:355-8
45. Liu Y, Sun J, Zhao M. ONGene: A literature-based database for human oncogenes. J Genet Genomics **2017**;44:119-21
46. Hu J, Rho HS, Newman RH, Zhang J, Zhu H, Qian J. PhosphoNetworks: a database for human phosphorylation networks. Bioinformatics **2014**;30:141-2
47. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res **2015**;43:D512-20
48. Diella F, Cameron S, Gemund C, Linding R, Via A, Kuster B*, et al.* Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. BMC Bioinformatics **2004**;5:79
49. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology **2008**;26:1367-72
50. Oughtred R, Stark C, Breitkreutz BJ, Rust J, Boucher L, Chang C*, et al.* The BioGRID interaction database: 2019 update. Nucleic Acids Res **2019**;47:D529-D41
51. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P*, et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res **2006**;34:D668-72
52. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D*, et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res **2003**;13:2498-504
53. Coordinators NR. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res **2018**;46:D8-D13
54. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ*, et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res **2019**;47:D442-D50
55. Nalpas N, Schmitt M, Macek B. Whole-exome sequencing (and nanopore RNA-seq) pipeline. 1.1: Zenodo; 2020.

Figures

Figure 1: State-of-the-art proteogenomic workflow to maximise multi-omics coverage. [A] Number of non-synonymous nucleotide variants per variant type identified by WES for A375 S and R. [B] The number of protein sequences per reference or alternate databases, as well as the overlap in search space between databases (up to two missed cleavages). [C] Gene and transcript counts identified by RNA-sequencing for A375 S and R. [D] Protein group and phosphopeptide site counts quantified by LC-MS/MS for A375 S and R.

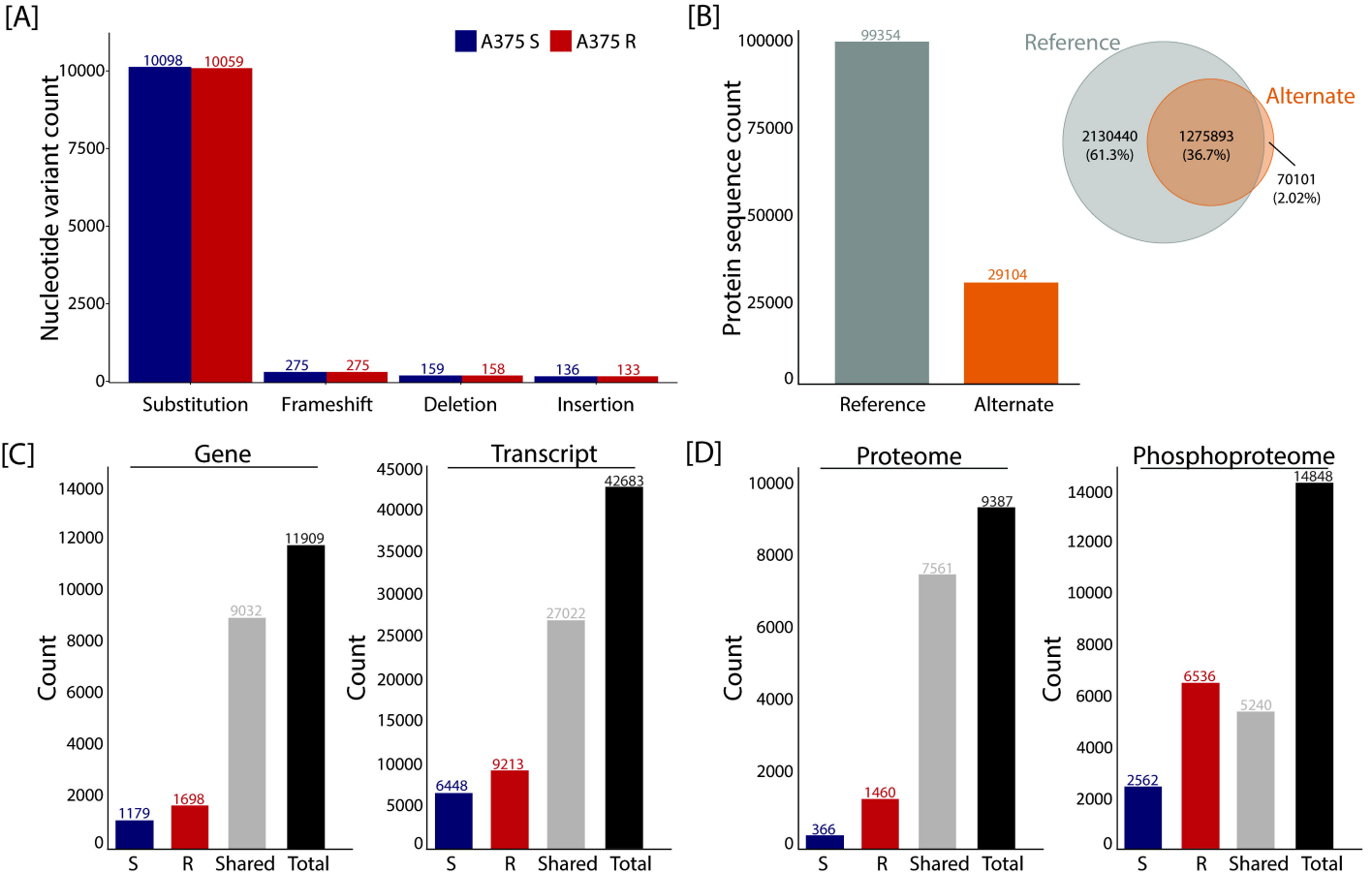
Figure 2: Comparison of drug-sensitive and resistant A375 cells at the genome, transcriptome and proteome level. [A, C and E] Comparison of A375 R vs. A375 S based on non-synonymous nucleotide variants [A], transcripts [C] and protein groups [E]. Summed abundance is the $\log_{10}$ sum of A375 R and A375 S for allelic depth [A], transcripts per million transcripts [C] and label-free quantification intensities [E]. Significantly changing entries are coloured in red and blue when up-regulated in A375 R and A375 S, respectively (Sig. A with s0, see **Supplementary Information**). [B] Venn diagram showing the overlap in genes with non-synonymous somatic mutations between A375 R (identified in this study) and the patients from Long and colleagues' study (Long et al., 2014). Within the study from Long and colleagues, the genes with somatic non-synonymous nucleotide variants are represented as shared across one or more patients. [D and F] The Reactome pathways that are overrepresented based on significantly changing transcripts [D] and protein groups [F] between A375 R (red) and A375 S (blue). Up to the top 20 Reactome pathways are displayed (ranked on FDR).

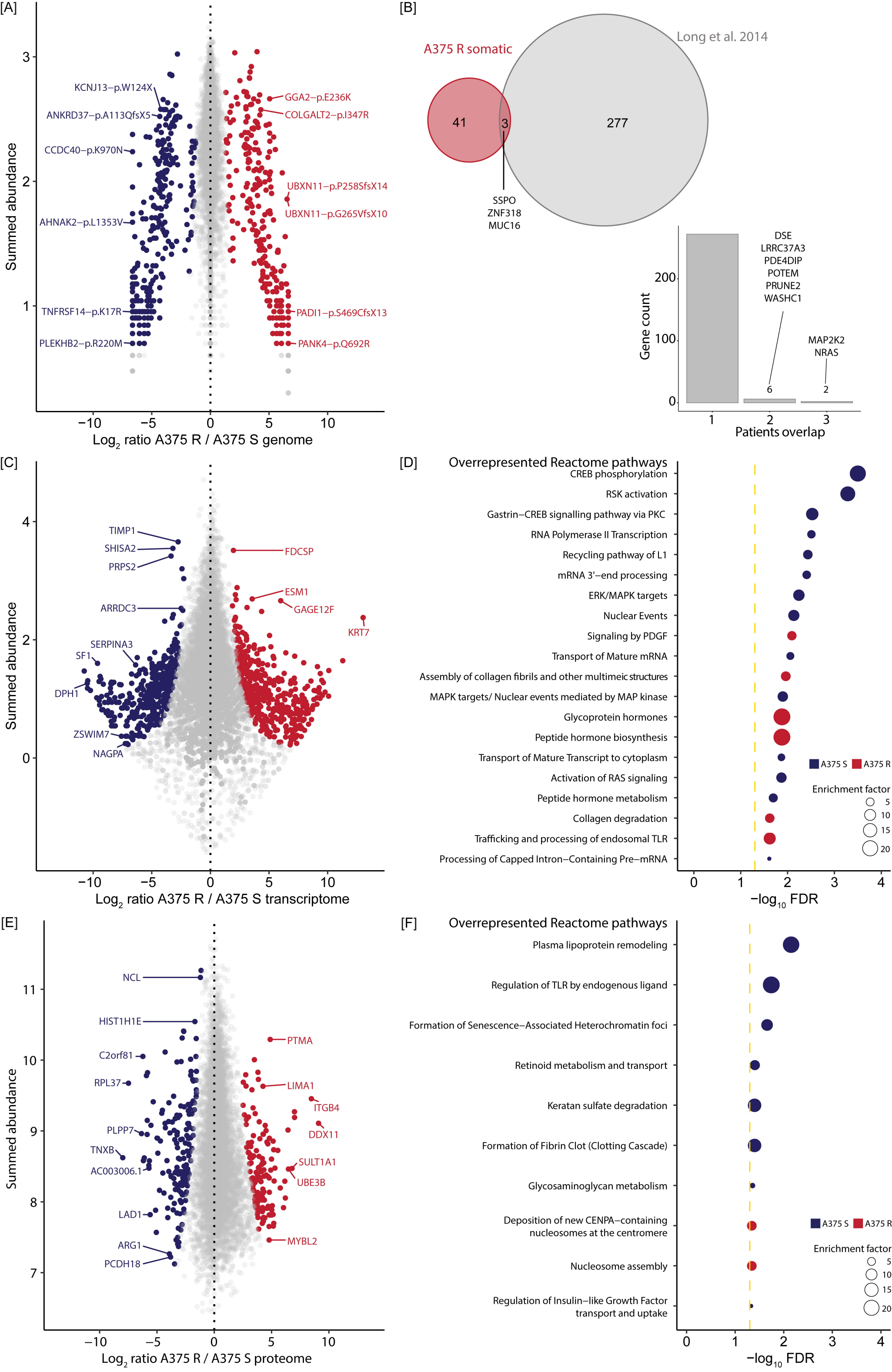
Figure 3: The disturbed signalling network in BRAFi resistant cells can be targeted by a number of drugs. [A] The interaction signalling network is generated based on significantly changing non-synonymous nucleotide variants (circle), transcripts (triangle), proteins (diamond)

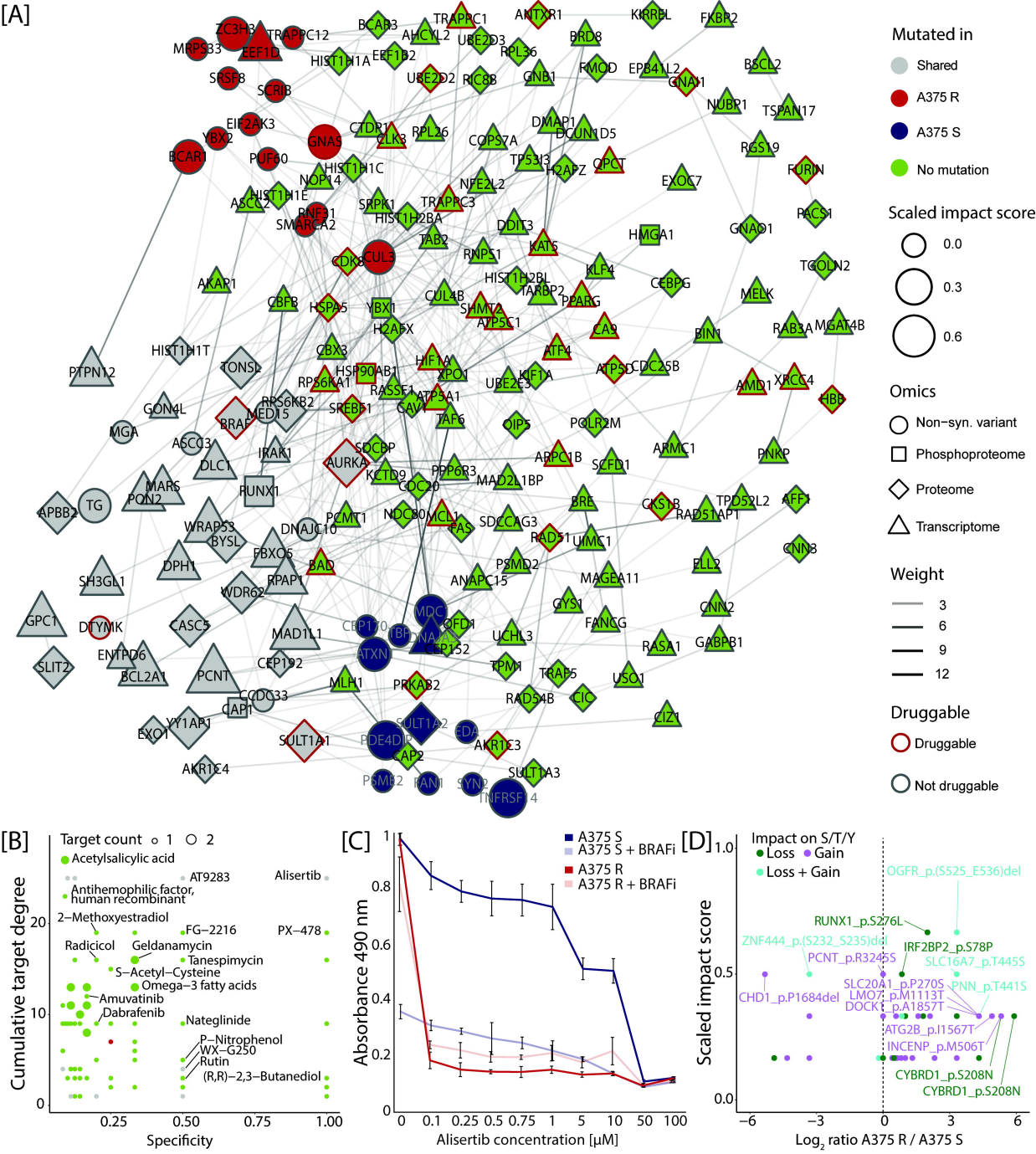
and phosphorylation sites (square). Only the top 200 entries are displayed (ranked based on their interaction degree). Entries that have a non-synonymous nucleotide variant only in A375 R are shown in red, only in A375 S are in blue, shared in both phenotypes are in grey and not mutated are in green. Entries that can be targeted by a drug are displayed with a red stroke. Entries were prioritised based on their importance in context of melanoma and BRAFi resistance (impact factor) and their node size increased accordingly. [B] The drugs (approved by FDA), interacting with entries from the interaction signalling network, are displayed based on their specificity to their target and how many connections their targets have. Colour-coding corresponds to whether one of the drug targets contains a non-synonymous nucleotide variant found only in A375 R (red), only in A375 S (blue), shared in both phenotypes (grey) and not mutated (green). [C] Cell viability assay of A375 S and A375 R cells treated with AURKA inhibitor Alisertib at the indicated concentrations or in combination with the BRAF inhibitor vemurafenib (2 µM). Cell viability was determined with MTS assay 72h after treatment start (n=6). Error bars represent standard deviations of replicates. [D] Scatter plot of the non-synonymous amino acid variants that have an impact on protein phosphorylation status either as a loss of a S/T/Y amino acids (green), gain of S/T/Y (purple) or loss/gain of S/T/Y (blue). The intensity ratio of A375 R vs. A375 S corresponds to the identified alternate peptide, while the scaled impact factor allows ranking of non-synonymous amino acid variants due to their importance in context of melanoma and phosphorylation status disruption.

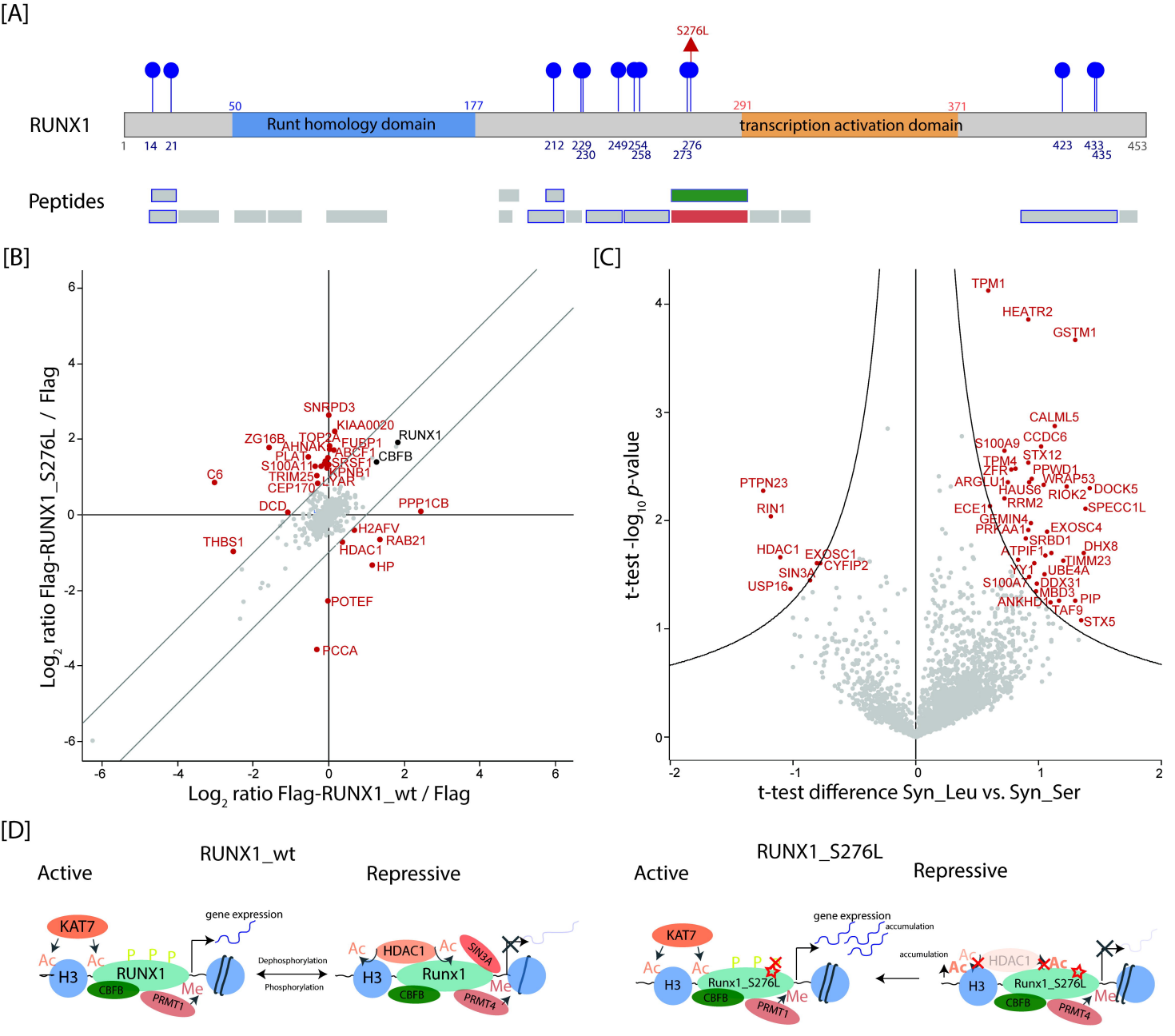
Figure 4: Loss of a known phosphorylation site leads to a changed interactome and altered transcriptional activity of RUNX1. [A] Schematic overview of the transcription factors RUNX1 protein. Numbers indicate the positions of amino acids residues within the protein. Identified phosphorylation sites are highlighted in blue and identified amino acid variants are highlighted in

red. Identified peptides by LC-MS/MS are shown in the second panel. Phosphorylated peptides are indicated with a blue border, while reference and alternate variant peptides are highlighted in green and red, respectively. [B] Interaction proteomics screen in A375 RUNX1_KO cells stably overexpressing Flag-tagged RUNX1_wt or Flag-tagged RUNX1_S276L. SILAC protein expression (log_2) of RUNX1_wt or Flag-tagged RUNX1_S276L relative to the corresponding control cell line (Flag tag only). RUNX1 and its core binding factor CBFB are marked in black. Significantly up and downregulated proteins are highlighted in red. Results represent three replicates per experiment group. [C] Volcano plot of synthetic alternate peptide (Syn_Leu) versus synthetic reference peptide (Syn_Ser) pulldowns of A375 cells. The log_2 fold change in abundance between Syn_Leu and Syn_Ser are plotted against $-log_{10}$ p-value (n=3). Black lines indicate the significance threshold based on student t-test (FDR < 0.01; S0=1.2). Significantly up and downregulated proteins are highlighted in red. [D] Schematic overview of proposed interaction of RUNX1_wt and RUNX1_S276L with main transcriptional regulators.


[A]
A375 S
A375 R
Nucleotide variant count
10098
10059
275
275
159
158
136
133
Substitution
Frameshift
Deletion
Insertion
[B]
Protein sequence count
99354
29104
Reference
Alternate
Reference
Alternate
2130440
(61.3%)
1275893
(36.7%)
70101
(2.02%)
[C]
Gene
Transcript
Count
1179
1698
9032
11909
6448
9213
27022
42683
S
R
Shared
Total
[D]
Proteome
Phosphoproteome
366
1460
7561
9387
2562
6536
5240
14848


[A]
Summed abundance
Log2 ratio A375 R / A375 S genome
KCNJ13−p.W124X
ANKRD37−p.A113QfsX5
CCDC40−p.K970N
AHNAK2−p.L1353V
TNFRSF14−p.K17R
PLEKHB2−p.R220M
GGA2−p.E236K
COLGALT2−p.I347R
UBXN11−p.P258SfsX14
UBXN11−p.G265VfsX10
PADI1−p.S469CfsX13
PANK4−p.Q692R
[B]
A375 R somatic
Long et al. 2014
41
3
277
SSPO
ZNF318
MUC16
Gene count
Patients overlap
DSE
LRRC37A3
PDE4DIP
POTEM
PRUNE2
WASHC1
MAP2K2
NRAS
6
2
[C]
Summed abundance
Log2 ratio A375 R / A375 S transcriptome
TIMP1
SHISA2
PRPS2
ARRDC3
SERPINA3
SF1
DPH1
ZSWIM7
NAGPA
FDCSP
ESM1
GAGE12F
KRT7
[D]
Overrepresented Reactome pathways
CREB phosphorylation
RSK activation
Gastrin−CREB signalling pathway via PKC
RNA Polymerase II Transcription
Recycling pathway of L1
mRNA 3'−end processing
ERK/MAPK targets
Nuclear Events
Signaling by PDGF
Transport of Mature mRNA
Assembly of collagen fibrils and other multimeic structures
MAPK targets/ Nuclear events mediated by MAP kinase
Glycoprotein hormones
Peptide hormone biosynthesis
Transport of Mature Transcript to cytoplasm
Activation of RAS signaling
Peptide hormone metabolism
Collagen degradation
Trafficking and processing of endosomal TLR
Processing of Capped Intron−Containing Pre−mRNA
A375 S
A375 R
Enrichment factor
−log10 FDR
[E]
Summed abundance
Log2 ratio A375 R / A375 S proteome
NCL
HIST1H1E
C2orf81
RPL37
PLPP7
TNXB
AC003006.1
LAD1
ARG1
PCDH18
PTMA
LIMA1
ITGB4
DDX11
SULT1A1
UBE3B
MYBL2
[F]
Overrepresented Reactome pathways
Plasma lipoprotein remodeling
Regulation of TLR by endogenous ligand
Formation of Senescence−Associated Heterochromatin foci
Retinoid metabolism and transport
Keratan sulfate degradation
Formation of Fibrin Clot (Clotting Cascade)
Glycosaminoglycan metabolism
Deposition of new CENPA−containing nucleosomes at the centromere
Nucleosome assembly
Regulation of Insulin−like Growth Factor transport and uptake
A375 S
A375 R
Enrichment factor
−log10 FDR


[A]
Mutated in
Shared
A375 R
A375 S
No mutation
Scaled impact score
0.0
0.3
0.6
Omics
Non−syn. variant
Phosphoproteome
Proteome
Transcriptome
Weight
3
6
9
12
Druggable
Druggable
Not druggable
[B]
Target count
1
2
Acetylsalicylic acid
AT9283
Alisertib
Antihemophilic factor, human recombinant
2−Methoxyestradiol
FG−2216
PX−478
Radicicol
Geldanamycin
Tanespimycin
S−Acetyl−Cysteine
Omega−3 fatty acids
Amuvatinib
Dabrafenib
Nateglinide
P−Nitrophenol
WX−G250
Rutin
(R,R)−2,3−Butanediol
Cumulative target degree
Specificity
[C]
A375 S
A375 S + BRAFi
A375 R
A375 R + BRAFi
Absorbance 490 nm
Alisertib concentration [μM]
[D]
Impact on S/T/Y
Loss
Gain
Loss + Gain
OGFR_p.(S525_E536)del
RUNX1_p.S276L
ZNF444_p.(S232_S235)del
IRF2BP2_p.S78P
PCNT_p.R3245S
SLC16A7_p.T445S
SLC20A1_p.P270S
PNN_p.T441S
LMO7_p.M1113T
CHD1_p.P1684del
DOCK1_p.A1857T
ATG2B_p.I1567T
INCENP_p.M506T
CYBRD1_p.S208N
CYBRD1_p.S208N
Scaled impact score
Log2 ratio A375 R / A375 S


[A]
S276L
RUNX1
Runt homology domain
transcription activation domain
50
177
291
371
1
14
21
212
229
230
249
254
258
276
273
423
433
435
453
Peptides
[B]
Log2 ratio Flag-RUNX1_S276L / Flag
Log2 ratio Flag-RUNX1_wt / Flag
SNRPD3
KIAA0020
RUNX1
ZG16B
TOP2A
FUBP1
AHNAK
ABCF1
CBFB
PLA1
SRSF1
S100A11
KPNB1
C6
TRIM25
LYAR
CEP170
DCD
PPP1CB
THBS1
H2AFV
RAB21
HDAC1
HP
POTEF
PCCA
[C]
t-test -log10 p-value
t-test difference Syn_Leu vs. Syn_Ser
TPM1
HEATR2
GSTM1
CALML5
CCDC6
S100A9
STX12
TPM4
PPWD1
ZFR
WRAP53
ARGLU1
HAUS6
DOCK5
RIOK2
PTPN23
RRM2
SPECC1L
RIN1
ECE1
GEMIN4
EXOSC4
PRKAA1
SRBD1
DHX8
HDAC1
EXOSC1
ATPIF1
TIMM23
CYFIP2
YY1
UBE4A
SIN3A
S100A7
DDX31
USP16
MBD3
PIP
ANKHD1
TAF9
STX5
[D]
RUNX1_wt
Active
Repressive
gene expression
KAT7
Ac
H3
RUNX1
CBFB
PRMT1
Me
P
Dephosphorylation
Phosphorylation
HDAC1
SIN3A
Runx1
PRMT4
RUNX1_S276L
Runx1_S276L
accumulation