Individualized proteogenomics of malignant melanoma cell lines and tissues in response to kinase inhibition and immunotherapy

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Marisa Schmitt

aus Miltenberg

Tübingen

2020

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	28.07.2020
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Boris Macek
2. Berichterstatter:	Prof. Dr. Alfred Nordheim

Erklärung/Declaration

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

„Individualized proteogenomics of malignant melanoma cell lines and tissues in response to kinase inhibition and immunotherapy"

selbständig verfasst, nur die angebenden Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stelle als solche gekennzeichnet habe. Ich erkläre, dass die Richtlinien zur Sicherung guter wissenschaftlicher Praxis der Universität Tübingen beachtet wurden. Ich versichere, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

I hereby declare that I have produced the work entitled „Individualized proteogenomics of malignant melanoma cell lines and tissues in response to kinase inhibition and immunotherapy", submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den 03.06.2020 _____

Declaration for the parts related to SkMel28 R

After completion and evaluation of my doctoral thesis by the supervisors and during revision of the manuscript II it came to our attention that one of the melanoma cell lines used in this study (SkMel28 R) was inadvertently contaminated in collaborator's laboratory with another cell line (A375 R). Unfortunately, this affects text, figures and tables of the second part of this thesis (manuscript II) related to this cell line. However, the main conclusions related to this part are not affected. The examination committee was familiarized with this issue during my defense.

Tübingen, den 09.09.2020 _____

Table of contents

Summary

Melanoma is the most aggressive form of skin cancer, with a rapidly increasing incidence rate. Malignant melanoma is characterized by mutations in the mitogen-activated kinase (MAPK) pathway, which strongly correlate with poor prognosis of the disease. The kinase BRAF is mutated in ~48% of melanoma cases, resulting predominantly in V600E substitution that leads to constitutive activation of the BRAF kinase and downstream signaling pathways. Over the last decade, several therapeutic treatments for melanoma have been developed with improved efficiency and overall survival rates. Targeted inhibition of the mutated BRAF with selective inhibitors, such as vemurafenib or dabrafenib and immunotherapy with the immune checkpoint antibodies targeting PD-1 and CTLA-4 receptors, results in regression of the disease. However, only a minority of patients can benefit from the current therapies and most of them quickly develop resistance to the treatment. Prognostic biomarkers, resistance mechanisms and mutational profiles of melanoma are mainly studied by genomics and transcriptomics. Although only about 2% of the genome codes for proteins, variants in these region of the genome have a high potential to rewire signal transduction networks. In addition, the majority of targeted cancer therapies do not target the genome, but rather the protein itself. Thus, it is highly important to analyze proteins and their patient-specific alterations in context of personalized medicine. Mass spectrometry-based proteomics can be used to study protein-specific clinical questions and can identify molecular mechanisms of treatment-resistant melanoma. By combining personalized genomics and proteomics, in an approach called proteogenomics, it is possible to derive patient-specific protein sequence databases – that include patient-specific amino acid variants. These in turn can provide deeper and more comprehensive molecular characterization of cellular processes that underlie disease progression. Several mechanisms for acquired resistance and even cross-resistance in melanoma have been detected, but key (phospho)proteins involved in resistance, as well as mutations altering protein modification status are still largely elusive. This thesis develops and applies personalized proteogenomics workflows to study these mechanisms on the level of individual melanoma cells and patient tissues.

In the first part of this thesis, a SILAC-based quantitative (phospho)proteomics profiling of vemurafenib-resistant and -sensitive A375 melanoma cells was performed to gain new insights into molecular processes that govern resistance to BRAF inhibitors. Among down-regulated proteins in vemurafenib-resistant cell lines were multiple cytoskeletal proteins including the intermediate filament nestin. Previous studies showed that nestin is expressed in various types of solid tumors and its abundance correlates with malignant phenotype of transformed cells.

However, the role of nestin in cancer cells with regard to acquired resistance is still poorly understood. CRISPR/Cas9 knockout of the nestin gene showed that the loss of nestin leads to increased cellular proliferation and colony formation upon treatment with kinase inhibitors. Moreover, nestin depletion is associated with an invasive phenotype and acquired resistance to MEK and BRAF inhibitors. Finally, phosphoproteome analysis revealed that nestin depletion affects integrin and PI3K/AKT/mTOR pathway signaling similar to resistant cells. In this part, proteomic and phosphoproteomic changes have been determined for BRAF inhibitor resistant and sensitive cells.

In the second part of the thesis, an individualized proteogenomics approach was applied to two melanoma cell lines, A375 and SkMel28, to analyze non-synonymous mutations and their impact on signal transduction networks in context of acquired resistance to kinase inhibitors. Integration of genomics and proteomics highlighted the distinct mutational landscape of both cell lines and revealed that cancer mutations are accumulating in MAPK and ErbB signaling pathways in resistant cells. Several alternate peptides interfering with the modification status of proteins with a potential to rewire signal transduction pathways were confirmed by high resolution mass spectrometry. Among them was transcription factor RUNX1, previously connected with myeloid leukemia and breast cancer. Validation of a loss of a known phosphorylation site on RUNX1 using SILAC-based protein interaction studies suggested that this mutation has an impact on the interactome of the protein and may alter its transcriptional activity. Taken together, this part of the thesis established the individualized proteogenomic workflow for analysis of mutational profiles of cancer cell lines and tissues.

In the third part of the thesis, this individualized proteogenomics approach was applied to four clinical melanoma samples in response to immunotherapy. Integration of the matching genomics and (phospho)proteomics datasets revealed an extensive number of patient-specific variants and disproportional number of shared variants in immune checkpoint inhibitor (ICi)-treated patient samples compared to untreated (naïve) samples. The proteogenomic signatures of human tissues could be recapitulated in patient-derived xenografts, thus allowing phosphoproteomics analysis. MS-measurements confirmed mutation-driven modification changes of several proteins specific to one sample, most of them were previously not reported in melanoma. Statistical analysis revealed differing mechanisms and associated network-attacking mutations in response to immunotherapy, such as PI3K/AKT signaling or GTPase activation in ICi treated samples. The gain of a new phosphorylation site on the GEF protein DOCK1 was further investigated by interactome studies and the results showed that this mutation has an impact on the interactome of DOCK1. The obtained results have demonstrated

that the developed individualized proteogenomic workflow can be efficiently applied to human melanoma tissue and patient-derived xenografts in response to immunotherapy.

Taken together, this thesis presents a new personalized proteogenomics workflow that can be routinely applied to numerous types of cancer and other diseases involving patient-specific accumulation of mutations in protein-coding genes. Datasets reported in this thesis provide new insights into resistance mechanisms and associated mutations with the potential to rewire signal transduction networks in malignant melanoma. This work can therefore serve as a basis for further improvement of therapeutic treatment of cancer patients.

Zusammenfassung

Das Melanom ist die aggressivste Form von Hautkrebs mit einer schnell ansteigenden Inzidenzrate. Das maligne Melanom ist gekennzeichnet durch Mutationen im Mitogen-aktivierten Kinaseweg (MAPK), die mit einer schlechten Prognose der Krankheit einhergehen. Die Kinase BRAF ist in circa 48% der Melanomfälle mutiert, was vorwiegend zu einer Substitution von V600E und zu einer konstitutiven Aktivierung der Kinase BRAF und der nachgeschalteten Signalwege führt. In den letzten Jahren wurden mehrere therapeutische Behandlungen für Melanome mit verbesserter Effizienz und erhöhter Überlebensrate von betroffenen Patienten entwickelt. Eine gezielte Hemmung der mutierten Kinase BRAF mit selektiven Inhibitoren wie Vemurafenib oder Dabrafenib und eine Immuntherapie mit den Antikörpern gegen die PD-1 und CTLA-4 Rezeptoren führen zu einem Rückgang des Krankheitsverlaufes. Allerdings profitiert nur eine Minderheit der Patienten von den derzeitigen Therapien, während die meisten dieser Patienten schnell eine Resistenz gegen diese Behandlungen entwickeln. Prognostische Biomarker, Resistenzmechanismen und Mutationsprofile wurden vor allem in Genom und- Transkriptomstudien analysiert. Obwohl nur 2% des Genoms für Proteine codieren, haben Mutationen in diesem Bereich des Genoms ein hohes Potential Signaltransduktionsnetzwerke zu verändern. Infolgedessen richten sich die Mehrzahl der derzeitigen Therapien nicht gegen bestimme Gene, sondern gegen Proteine in Krebszellen. Daher ist es sehr wichtig, Proteine und ihre patientenspezifischen Veränderungen im Kontext der personalisierten Medizin zu analysieren. Massenspektrometrie-basierte Proteomik kann proteinspezifische klinische Fragen beantworten, die weit außerhalb der Reichweite der Genomik liegen und molekulare Mechanismen identifizieren, die dem metastasierten und behandlunsgresistenten Melanom zugrunde liegen. Durch die Kombination von personalisierter Genomik und Proteomik, die als Proteogenomik bezeichnet wird, können patientenspezifische Proteinsequenzdatenbank generiert werden, die patientenspezifische Aminosäurevarianten enthalten. Diese können wiederum eine tiefere und umfassendere molekulare Charakterisierung der zellulären Funktionen ermöglichen, die dem Fortschreiten der Krankheit zugrunde liegen. Verschiedene Mechanismen für erworbene Resistenzen, darunter sogar Kreuzresistenzen, wurden im Melanom nachgewiesen, aber Schlüsselproteine, die daran beteiligt sind, sowie damit verbundene Mutationen, die den Modifikationsstatus verändern und daher die Resistenzentwicklung beeinflussen können, sind noch weitgehend unerforscht. Diese Arbeit entwickelt und wendet personalisierte Proteogenomikansätze an, um diese Mechanismen auf der Ebene einzelner Melanomzellen und Patientengewebes zu analysieren.

Im ersten Teil dieser Arbeit wurde ein SILAC-basierter quantitativer Ansatz zur Untersuchung der Proteomveränderungen von Vemurafenib-resistenten und –sensitiven Melanomzellen angewendet, um neue Einblicke in molekulare Prozesse zu erhalten, die die Resistenz steuern. Unter den herunterregulierten Proteinen in Vemurafenib-resistenten Zelllinien befanden sich mehrere Zytoskelettproteine, einschließlich das Intermediärfilament Nestin. Frühere Studien zeigten, dass Nestin in verschiedenen Arten von soliden Tumoren exprimiert wird und seine Abundanz mit einem malignen Phänotyp transformierter Zellen korreliert. Allerdings ist die Rolle von Nestin in Krebszellen im Hinblick auf die erworbene Resistenz noch nicht ausreichend bekannt. CRISPR/Cas9 generierte Knockouts des Nestin-Gens zeigten, dass der Verlust von Nestin bei Behandlung mit Kinase-Inhibitoren zu einer erhöhten Zellproliferation und Koloniebildung führt. Darüber hinaus korreliert die Expression von Nestin mit einem invasiven Phänotyp und einer erworbenen Resistenz gegenüber MEK- und BRAF-Inhibitoren. Schließlich ergab die Phosphoproteomanalyse, dass der Knockout von Nestin die Signalübertragung durch Integrin und PI3K/AKT/mTOR-Signalwege ähnlich wie bei resistenten Zellen beeinflusst. In diesem Teil der Arbeit wurden proteomische und phosphoproteomische Veränderungen von BRAF-Inhibitor-resistenten und sensitiven Zellen bestimmt.

Im zweiten Teil der Arbeit wurde ein individualisierter Proteogenomikansatz auf zwei Melanomzelllinien (A375 und SkMel28) angewendet, um nicht synonyme Mutationen und deren Auswirkungen auf Signaltransduktionsnetzwerke im Zusammenhang mit der erworbenen Resistenz gegenüber Kinaseinhibitoren zu analysieren. Die Integration von Genomik- und Proteomikdaten hob die unterschiedlichen Mutationsprofile beider Zelllinien hervor und zeigte, dass Krebsmutationen in MAPK- und ErbB-Signalwegen in resistenten Zellen akkumulieren. Mehrere mutierte Peptide, die den Modifikationsstatus von Proteinen beeinflussen können und somit ein hohes Potenzial haben Signaltransduktionswege zu beeinflussen, wurden durch hochauflösende Massenspektrometrie bestätigt. Unter diesen befand sich der Transkriptionsfaktor RUNX1, der zuvor mit myeloischer Leukämie und Brustkrebs in Verbindung gebracht wurde. Die Validierung des Verlustes einer bekannten Phosphorylierungsstelle von RUNX1 unter Verwendung von SILAC-basierten Proteininteraktionsstudien legte nahe, dass diese Mutation einen Einfluss auf das Interaktom von RUNX1 hat und dessen Transkriptionsaktivität verändern kann. Zusammenfassend etablierte dieser Teil der Arbeit einen individualisierten Proteogenomikansatz zur Analyse von Mutationsprofilen von Krebszelllinien und Geweben.

Im dritten Teil der Arbeit wurde dieser individualisierte Proteogenomikansatz auf vier klinische Melanomenproben unter Immuntherapie angewendet. Die Integration der Genomik- und dazugehörigen Phospho-Proteomikdaten führte zur Identifizierung einer Vielzahl von patientenspezifischen Mutationen und einer überproportionalen Anzahl derselben Genvarianten in Immun-Checkpoint-Inhibitoren behandelten Proben gegenüber unbehandelten Proben. Die in menschlichem Gewebe identifizierten proteogenomischen Signaturen konnten in vom Patienten stammenden Xenotransplantaten rekonstruiert werden, wodurch Phosphoproteomeanalysen ermöglicht wurden. MS-Messungen bestätigten mutationsbedingte Modifikationsveränderungen an mehreren Proteinen, von denen die meisten bisher nicht im Melanom bekannt sind. Die statistische Analyse ergab unterschiedliche Mechanismen und damit verbundene netzwerkangreifende Mutationen als Reaktion auf eine Immuntherapie, wie z.B. PI3K/AKT-Signalisierung oder GTPase-Aktivierung in Immuntherapie-behandelten Proben. Der Gewinn einer neuen Phosphorylierungsstelle am GEF-Protein DOCK1 wurde durch Interaktomstudien weiter untersucht, und die Ergebnisse zeigten, dass diese Mutation einen Einfluss auf das Interaktom von DOCK1 hat. Die erhaltenen Ergebnisse zeigten, dass der entwickelte individualisierte proteogenomische Ansatz auf menschliches Melanomgewebe und von Patienten stammende Xenotransplantate im Zusammenhang mit Immuntherapie angewendet werden kann.

Zusammenfassend präsentiert diese Arbeit einen personalisierten Proteogenomikansatz, der routinemäßig auf weitere Krebsarten und andere Krankheiten angewendet werden kann, bei denen patientenspezifische Mutationen in proteinkodierenden Genen akkumulieren. In dieser Arbeit beschriebene Datensätze bieten neue Einblicke in Resistenzmechanismen und damit verbundene Mutationen mit dem Potenzial Signaltransduktionsnetzwerke bei malignen Melanomen zu beeinflussen. Diese Arbeit kann daher als Grundlage für eine weitere Verbesserung von therapeutischen Behandlungen von Krebspatienten sein.

1 Introduction

1.1 Malignant melanoma

Melanoma is the 19th most common cancer worldwide, accounting for approximately 232,000 new cases in Germany in 2012 [1]. Although malignant melanoma accounts for less than five percent of all skin cancer cases it is the most serious form of the disease, causing up to 75% skin cancer related deaths. Melanoma predominantly affects Caucasian population of both genders, and once it becomes metastatic, the prognosis is very poor [2]. Ultraviolet radiation from sunlight is one of the main risk factors for the development of melanoma and is directly associated with the UV-B spectrum [3,4]. It is also related to sun patterns and timing in particular intense and intermittent sun exposure. Individuals with a history of sunburn in their childhood or adolescence are at highest risk of developing melanoma [5]. In addition, UV-A exposure from artificial sources have been associated in a number of studies to an increased risk of developing melanoma [6,7]. Besides environmental risk factors, host factors play an important role including family history, genetic susceptibility and the number of acquired melanocytic nevi [8]. Approximately 7-15% of melanoma cases occur in patients with a family history for melanoma and around 25% arises from pre-existing nevus [9,10]. Cutaneous melanoma is the most common form of melanoma accounting for 90% of all incidences and arises from melanocytes [11]. Melanocytes, the pigment producing cells can be found in skin, eye, inner ear, and leptomeninges [12,13]. In skin, melanocytes are located and consistently distributed in the basal layer of epidermis and represents the second largest cell population within the skin [14]. Epidermal melanocyte proliferation and pigment production is stimulated by UV-radiation induced DNA damage in the neighboring keratinocytes, which secretes α-melanocyte stimulating hormone (αMSH) [15]. Binding of αMSH to the melanocortin 1 receptor (MC1R) on melanocytes induces the synthesis of the macromolecule melanin. Melanin is transferred to adjacent keratinocytes, where it accumulates around the nucleus and building a photo-protective barrier against UV-radiation [16,17]. Melanin production is initiated and regulated by the Wnt signaling pathway, c-KIT receptor tyrosine kinase, and downstream transcription factors like melanocyte inducing transcription factor (MITF) [18,19]. Several studies have shown that stimulation of the MC1R receptor dictates differentiation and migration of melanocytes [20]. Melanoma is caused by uncontrolled growth of melanocytes forming a tumor which can become malignant and spread to other organs. Melanoma progression begins from a localized lesion undergoing radial growth, then reaching lymph node by undergoing vertical growth and finally metastasizing to distant organs via the lymphatic and circulatory systems (Figure 1). In

2009, melanoma was classified by Balch *et al.* into four stages (I-IV) based on the thickness, ulceration and metastatic status [21]. Localized melanoma is divided into two stages: stage I and stage II melanoma. This includes patients with no evidence for distant metastases and low risk for metastases and melanoma-specific mortality (stage I) or intermediate risk and melanoma-specific mortality (stage II). The ten-year survival rate after surgical removal is 95% of stage I and 39% of stage II. Regional metastases, stage III melanoma, shows a high in-transit metastasis, absence or presence of ulceration of the primary melanoma and metastatic lymph nodes. The five-year survival rate for patients with stage III melanoma is 60% in absence of ulceration and 36% in presence of ulceration. Contrastingly, patients with distant metastases (stage IV) have a poor prognosis with a median survival time of six to nine months and a one-year survival rate of only 25% [22]. The aggressiveness of this cancer type highlights the need for improvement of existing therapies [23].

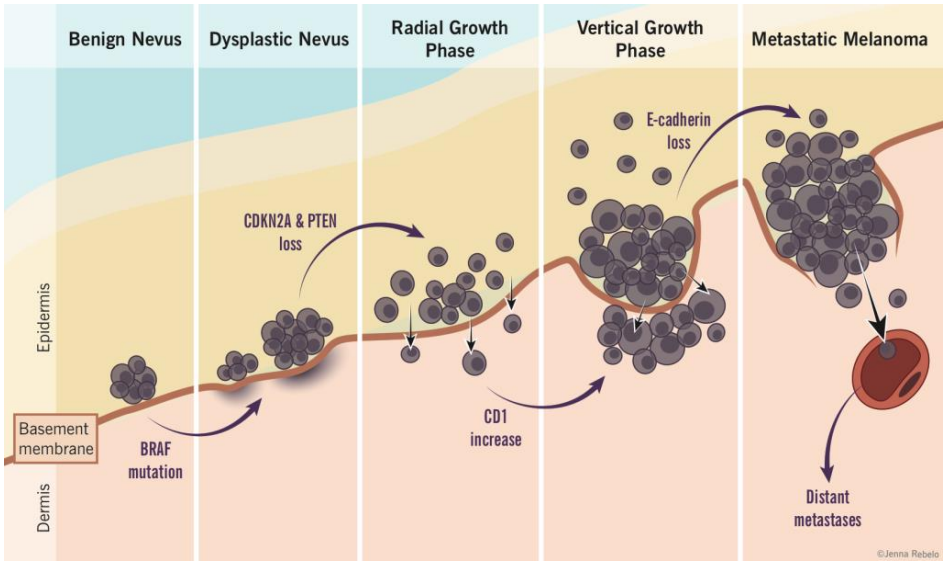


Figure 1: Pathological staging of melanoma progression based on Clark model. Progression can start from uncontrolled growth of normal melanocytes to form a benign nevus. The benign nevus can proliferate to a dysplastic nevus with irregular borders and variable pigmentation. In the radial growth phase, the melanocytes proliferate horizontally into the epidermis, followed by vertical proliferation and invasion of the basement membrane in the vertical growth phase. Malignant melanocytes spread to distinct areas of the body, most of the time first to lymph nodes, then skin, soft tissues, lung and brain. (adapted from Seuradge and Wong, 2006 [24])

1.1.1 Mutational landscape of melanoma

Two main pathways are known to be deregulated in melanoma development, the mitogen-activated protein kinase (MAPK) pathway and phosphatidylinositol-3-kinase (PI3K) pathway (Figure 2). In the MAPK pathway, receptor-tyrosine-kinases (RTKs) are activated by a variety

of growth factors, i.e. epidermal growth factor (EGF), platelet-derived growth factor (PDGF) and nerve growth factor (NGF), as well as G-protein-coupled receptors (GPCRs) and cytokines [25]. Binding of ligands to the receptors activate the tyrosine kinase activity of the cytoplasmic domains, which leads to autophosphorylation of several tyrosine residues [26]. The phosphorylated tyrosine residues form binding sites for Src homology 2 (SH2), Src homology 3 (SH3) or phosphotyrosine binding (PTB) domain-containing proteins like growth factor receptor-bound protein 2 (GRB2) [27]. GRB2 binds the guanine nucleotide exchange factor SOS, which induces the exchange of guanosine diphosphate (GDP) to guanosine triphosphate (GTP) in the G-protein RAS. Activated RAS dissociates from the complex and initiates the phosphorylation of the serine/threonine protein kinase RAF [28]. Phosphorylation-activated RAF kinase can induce downstream signaling by triggering phosphorylation of the serine/threonine/tyrosine kinase MEK and leads to the activation of the serine/threonine protein kinase ERK [29]. ERK phosphorylation induces expression of genes to cell survival, differentiation and proliferation.

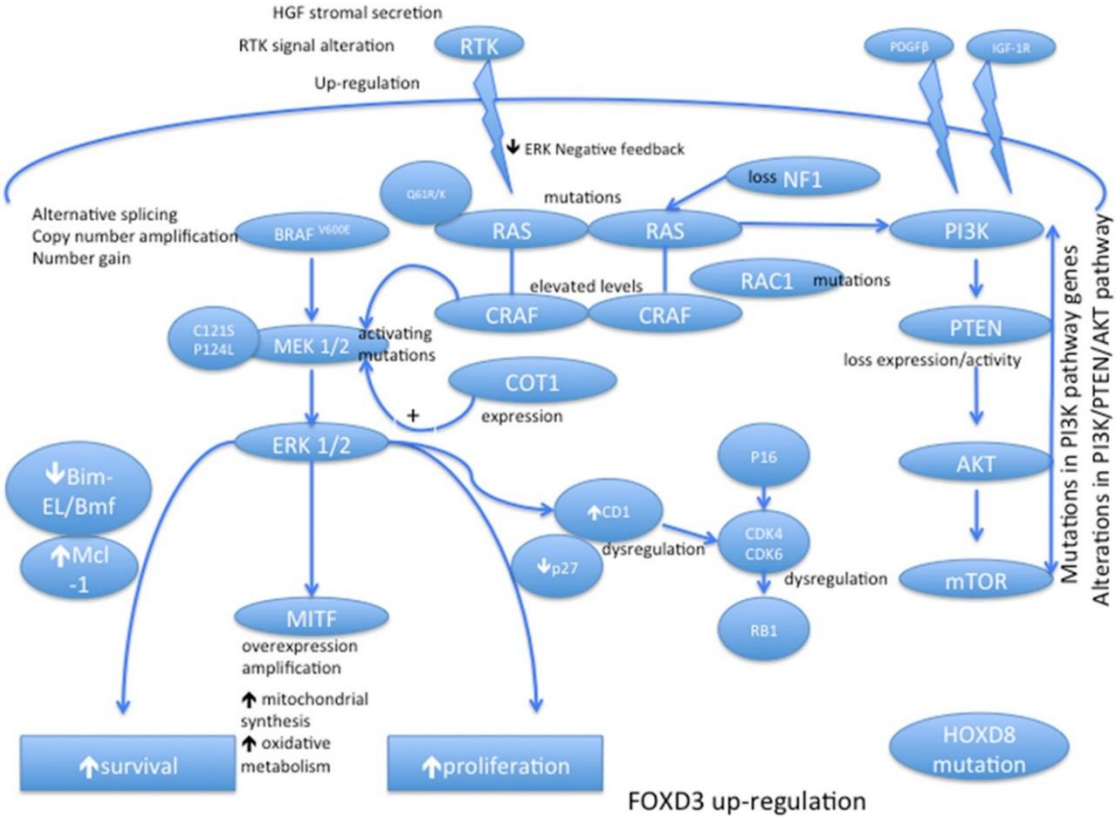


Figure 2: MAPK and PI3K/AKT/mTOR pathway and common mutations and alterations in melanoma. (adapted from Amaral *et al.*, 2017 [30])

The MAPK pathway is frequently mutated in malignant melanoma and other cancer types and gained clinical importance due to the high occurrence of RAF- activating mutations [31]. Around 48% of metastatic melanoma patients harbor non-synonymous somatic missense mutations in the RAF kinase BRAF (v-raf murine sarcoma viral oncogene homolog B) [29]. The predominant mutation, accounting for 90% of BRAF mutations, is within the kinase domain and results in the substitution of valine to glutamic acid at amino acid 600 (V600E) [31]. This mutation can result in a 500-fold increased activation of BRAF, and thus leads to a constitutive activation of downstream signaling in cancer cells [32,33]. In addition, the substitution of valine by lysine (V600K) at the same position is observed in 20% of melanoma patients. Another 6% of patients have other genotypes including alternative splicing or amplification of BRAF and mutations in ARAF [34]. Besides RAF alterations, mutations in the upstream regulator RAS (rat sarcoma) occur in approximately 18% of melanoma patients including NRAS. The dominant mutations are Q61R and Q61K substitutions resulting in inability to hydrolyze GTP to GDP. Activating mutations in KRAS or HRAS occur at a very low frequency in melanoma compared to other cancer types. BRAF and NRAS alterations are, so far, the most common point mutations detected in proto-oncogenes in melanoma and mutations rarely overlap, likely due to redundant pathway reactivation [35]. This supports the hypothesis that both kinase families are major signaling checkpoints in the MAPK pathways promoting malignant progression.

Besides mutations in the MAPK pathway, oncogenic alterations in the PI3K/ protein kinase B (AKT) pathway occur in approximately 50% of melanomas [36]. Similar to the MAPK pathway, the PI3K/AKT pathway is activated by extracellular ligand binding to RTKs, which causes receptor dimerization and autophosphorylation of tyrosine residues in the intracellular domain [37]. P85, the regulatory subunit of PI3K, binds to the receptor via its SH2 domain and recruits p110, the catalytic subunit, to the membrane. Both subunits form the activated PI3K enzyme. Independently of the receptor, p85 and p110 can be recruited and activated by GRB2 and GTP-bound RAS, respectively. Activated PI3K phosphorylates phosphatidylinositol to phosphatidylinositol (4,5)-biphosphate (PIP2) and subsequently to phosphatidylinositol-(3,4,5)-triphosphate (PIP3). These phosphorylated lipids are anchored in the plasma membrane and activate the serine/threonine kinase AKT through binding of the pleckstrin homology (PH) domain. The interaction with PIP3 causes a conformational change of AKT and thus leads to the exposure of the phosphorylation sites in the kinase domain and the C-terminal region of the protein. Phosphorylated S473 in the C-terminal region of AKT by mechanistic target of rapamycin complex (mTORC) or phosphoinositide-dependent kinase 2 (PDK2) stimulates AKT activation. Fully activated AKT leads to substrate-specific phosphorylation events in the

cytoplasm and the nucleus, where it regulates transcription of genes involved in cell proliferation, cell cycle progression and apoptosis, through a variety of downstream proteins like glycogen synthase kinase 3-beta (GSK-3β) [38]. Phosphatase and tensin homologue (PTEN) negatively regulates PI3K/AKT signaling pathway by dephosphorylating phosphatidylinositol at the 3'position [39]. A variety of alterations and copy number changes within significant components of this pathway have been identified and are now the focus of pharmacological development. Loss of expression or inactivating mutations of the tumor suppressor PTEN have been identified in approximately 20-30% of patients with malignant melanoma [40]. PTEN loss commonly occur in the presence of BRAF mutations leading to a hyperactivation of (PKB)/AKT signaling and activation of MAPK pathway whereby induction of apoptosis is abolished [41,42]. Other driver mutations are in the PD domain of AKT1 (at position E17), which were detected in 1-2% of melanoma cases [43]. Interestingly, an analogous mutation on AKT3 has been identified exclusively in melanoma and leads to increased expression and activation of AKT3 [44,45]. Apart from AKT mutations, 2-6% of melanoma contain a point mutations in PIK3CA encoding for the catalytic subunit p110 of PI3K [46,47]. This may result in a constitutive activation of AKT/mTOR. Mutations in the kinase domain of mTOR occur in approximately 10.5% of melanoma patients resulting in gain-of function and activation of a mTORC1-mediated feedback loop within the PI3K pathway. The pathway is also activated by amplifications of RTKs like EGFR or c-KIT. Mutations in c-KIT have been identified in 2-8% of melanoma cases [48]. Several studies have demonstrated a high prevalence of activating mutations in the PI3K/AKT pathway highlighting the critical role in melanoma progression.

In addition to mutations occurring directly within the MAPK and PI3K pathway, other genes are mutated in melanoma including NF1 (neurofibromatose type 1) and CDKN2A (cyclin-dependent kinase inhibitor 2A). NF1, a tumor suppressor gene, is mutated in approximately 10-15% of patients and is the third most frequently mutated gene in melanoma [49,50]. NF1 regulates RAS signaling by converting active RAS-guanosine triphosphate (RAS-GTP) to inactive RAS-guanosine diphosphate (RAS-GDP) [51]. Several studies showed that loss-of-function mutations diminished inhibition of RAS signaling and thus leads to increased MAPK and PI3K signaling [49,50,52]. Patients harboring NF1 mutations are often frequently mutated in various other genes including BRAF and NRAS [48,52]. The G-protein signaling proteins GNAQ (guanine nucleotide-binding protein G(q) alpha subunit) and GNA11 (guanine nucleotide-binding protein alpha 11) are altered in 9.5% of mucosal melanoma and up to 90% of uveal melanoma cases [53,54]. The most common alterations are mutations, amplifications and deletions resulting in an oncogenic GTP signaling and hyper proliferation of melanoma cells. Several studies showed that

mutations in GNAQ and QNA11 have led to MAPK pathway activation through RAS activation and phosphorylation of MEK and ERK [55,56]. In addition, the tumor suppressor gene CDKN2A is frequently mutated in melanoma patients with loss-of-function mutations or deletions. This gene encodes for several proteins including p16(INK4A) and p14(ARF) proteins, who induce cell growth and senescence [57]. The three-year overall survival rate decreases from 55% in patients with wild-type CDKN2A to 24% for patients with mutations in CDKN2A [32].

1.1.2 Standard treatment options for melanoma

Malignant melanoma is treated by combinations of surgical resection, immunotherapy, chemotherapy or targeted therapy. These treatment options are selected depending on the features of the tumor including location, stage and genetic profiles. For patients with stage I melanoma, surgery is the primary treatment; while, for patients with metastatic melanoma, chemotherapy, targeted therapy and immunotherapy are recommended. There are two main limitations in melanoma therapy: 1) reduced drug efficiency due to development of resistance to immune-, chemo- and targeted therapy [58,59] and 2) numerous side effects, which can lead to skin toxicity, associated to immune reactions, and lack of specificity for tumor cells [60]. In recent years, improved knowledge of the genetic profiles of melanocytes and a better understanding of the molecular factors involved in malignant transformation have led to the development and approval of several new therapeutic strategies. Chemotherapy using the standard medication dacarbazine (DTIC), an alkylating agent approved in 1974 by the U.S. Food and Drug Administration (FDA), achieved a complete response of <5% and a five-year survival in 2-6% of patients with metastatic melanoma [61]. DTIC methylates nucleic acids, which results in DNA damage causing growth arrest and cell death. Unfortunately, DTIC shows several gastrointestinal side effects including nausea and vomiting, as well as suppression of blood cell production in the bone marrow leading to anemia and neutropenia [62]. Development of chemoresistance and rapid metastasis are one of the main reasons for the low long-term survival rate of patients with metastatic melanoma. Chemoresistance is probably caused by resistance to drug-induced apoptosis or repair of drug-induced DNA damage [59]. Several studies also showed that resistant melanoma cells showed a reduced drug uptake, increased drug efflux and an intracellular drug inactivation [63]. Replacement with other chemotherapy reagents or combinations did not improve the overall survival of melanoma patients [61]. Therefore, new therapeutic treatments have been explored with improved effectiveness against malignant melanoma and reduced side effects.

1.1.3 Targeted therapy using immune checkpoint and kinase inhibitors

Over the last decades, the role of the immune system controlling tumor progression has been established and new immunotherapeutic targets showed remarkable clinical activity (Figure 3). Under normal physiological conditions, T cell antigen receptors (TCR) on lymphocytes respond to antigenic peptides presented on the cell surface on major histocompatibility complexes (MHCs) type I. The TCR is a disulfide-linked membrane-anchored heterodimeric protein complex with highly variable alpha and beta chains, which leads to a diverse repertoire of TCRs on T cells [64]. Due to infections or cancer, the TCR on T cells recognizes the myriad of possible foreign antigens presented in the surface of cells. The TCR recognition of the antigen leads to the activation of T cells, clonal selection and an activation of the process of immunity [65]. After activation and proliferation of T cells, effector T cells (Teffs) releases cytokines like interferon gamma (IFN-γ) which leads to the destruction of tumor cell by high cytotoxicity. This complex process comprises a constant interplay between inhibitory and stimulatory signals leading to the amplification of antigen-specific immune responses to cancer self-antigens while preventing autoimmunity. Dysfunctional immune reactions against tumor cells are key events in tumorigenesis and tumor progression and may be due to diminished antigen recognition and a highly immunorepressive tumor microenvironment [66]. Several studies demonstrate that impaired antigen recognition can be due to epigenetic and post-transcriptional silencing or alterations in the antigen-processing machineries. The tumor microenvironment can be influenced by a diverse variety of factors including enrichment of regulatory cells such as Tregs or the upregulation of co-inhibitory signals in lymphocytes [67].

Immunotherapy has recently become a valuable option for melanoma treatment. Anticancer immunotherapies could be broadly categorized into two groups: (1) drugs targeting immune tolerance via blockade of negative regulatory signals like co-inhibitory checkpoints and (2) drugs enhancing endogenous antitumor immune response via stimulation of immunogenic pathways. Therapeutic cancer vaccines, exogenous recombinant cytokines and oncolytic viruses are also used as immunostimulatory strategies to enhance antitumor immune response. Several drugs targeting immune tolerance are now in clinical development, and some were approved by the FDA for use in various cancer types. The reagents nivolumab and ipilimumab are immune checkpoint antibodies targeting PD-1 (the programmed cell death-1) and CTLA-4 (cytotoxic T lymphocyte-associated antigen-4) receptors. PD-1 and CTLA-4 are co-inhibitory T cell receptors and acts as negative regulatory receptors that block T cell activation and induce immune tolerance. The CTLA-4 antibody ipilimumab (approved in 2011 by the FDA) was the first agent which showed improved overall survival of melanoma patients. CTLA-4 is mainly

expressed on T cells and is a homolog of CD28, a costimulatory receptor on T cells [68]. Under normal conditions, CTLA-4 is expressed at very low level and is rapidly up-regulated upon antigen binding [69,70]. Both receptors bind the same ligands, CD80 and CD86 on antigen-presenting cells, however the binding affinity is much higher for CTLA-4 than CD28 [71]. The binding of CTLA-4 prevents the co-stimulation and activation of T cells. Several studies showed that anti-CTLA-4 antibodies also promote depletion of Tregs in the tumor microenvironment by expressing high levels of FcγRIV [72]. Subsequently, inhibition of CTLA-4 receptors with therapeutic anti-CTLA4 antibodies demonstrated tumor rejection and significant prolongation in patient survival. The overall survival rate was improved to 9.1 months with higher three-year survival rate of 20.8% compared to placebo [73]. However, only a minority of patients respond to ipilimumab and most patients display immune-related toxicities.

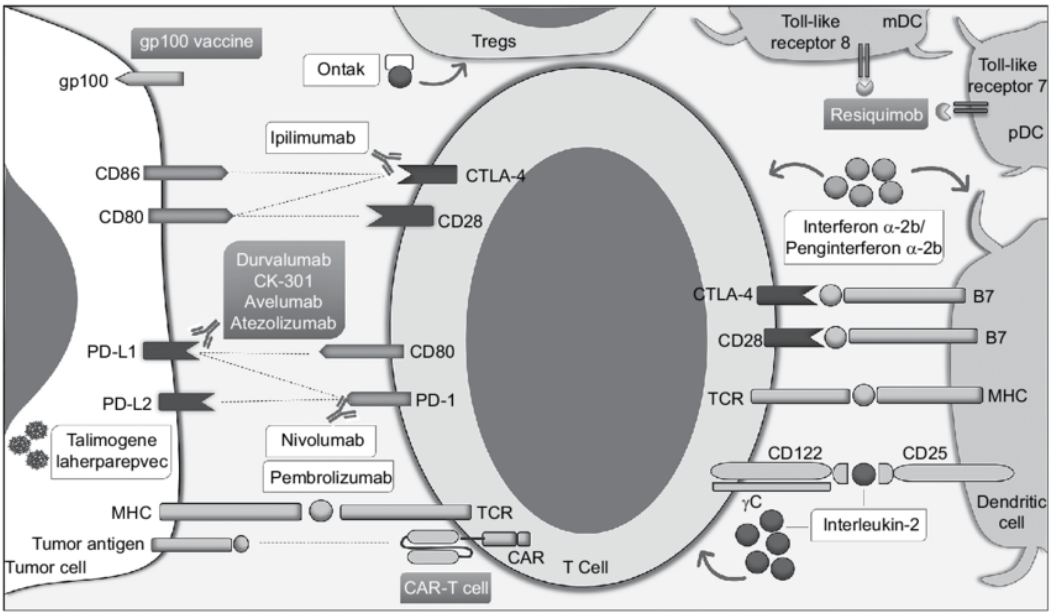


Figure 3: Immunotherapies for cutaneous melanoma treatment. Approved immunotherapy reagents are depicted in a white box, trials in a grey box. Ipilimumab, an anti-CTLA-4 antibody, induces antitumor immune response by binding to the CTLA-4 receptor, and increases clonal T cell expansion and infiltration. The anti-PD-1 antibody nivolumab block the interaction between the PD-1 receptor and the ligands PD-L1 and PD-L2. (adapted from Domingues *et al.*, 2018 [2]).

Clinical trials using another checkpoint receptor, anti-PD-1 antibodies reported higher response rates and fewer immune side effects compared to CTLA-4. PD-1 is expressed on several immune cells including activated T cells, B lymphocytes and natural killer (NK) cells [74,75]. Expression of PD-1 is induced by binding of its ligands PD-L1 and PD-L2 and by cytokines such as interleukin-2. Nivolumab [76] was approved in 2014 by the FDA and blocks the interaction between PD-1 and its ligands, which results in antitumor activity and reduces tumor

progression. Monotherapy with nivolumab showed a median progression-free survival (PSF) of 6.9 months and an improved overall survival compared with monotherapies of ipilimumab (median PSF of 2.2 months) in metastatic melanoma patients that are naïve to both reagents [77]. Combined therapies have been studied and achieved a median PSF of 11.5 months, especially in patients with PD-L1 negative tumors. Combinations of chemotherapy with immunotherapy also showed increased response rates and an improved median PSF compared to monotherapies [78]. However, no increase in the overall survival of patients could be observed and it was associated with severe toxicity [79]. Checkpoint inhibitor immunotherapy showed an increased progression free survival for melanoma patients, however only a small percentage of patients responded to these inhibitors. The complexity and multiplicity of involved mechanisms, heterogeneity in the immune response across tumors, the tumor microenvironment and the varying tumor immunogenicity affect the response and resistance to immune checkpoint blockade. Resistance can be divided into two groups according to the timing of occurrence that is primary resistance, for patients never-responding to the treatments, and acquired resistance developing following a period of response to the drugs. It can be also classified into intrinsic, altered processes due to gene expression, signal transduction, DNA damage and immune recognition in cancer cells, and extrinsic due to external T cell activation. The clinical response to immune checkpoint inhibitors and acquired resistance is often associated with a high mutational load and expression of tumor neoantigens leading to antitumor immunity [80,81]. Several studies showed that non-synonymous mutations were identified to generate tumor neoantigens that drive cytotoxic response against cancer cells [67,82,83]. A high mutational and neoantigen load was also found to be associated with responses to anti-CTLA-4 treatment in patients with metastatic melanoma [84]. The loss of neoantigen expression may lead to poor immunogenicity and acquired resistance to checkpoint inhibitor blockade [67]. A recent study demonstrated a loss of seven to 18 putative neoantigens in resistant NSCLC tumors after treatment with the checkpoint inhibitors PD-1 and CTLA-4 [85]. Several other studies showed that deficiencies in antigen presentation and down-regulation of MHC class I (MHC-I) play a role in immune checkpoint resistance [86-88]. Loss-of-function mutations in β2-microtubulin leading to loss of expression of MHC-I, thus allowing immune evasion of tumor cells [89,90]. Besides mutation in β2-microtubulin, loss of JAK/STAT pathway results in acquired resistance due to downregulation of MHC-1 [91,92]. Additionally, classic oncogenic pathways like MAPK or PI3K pathways can regulate immune response by influencing the tumor microenvironment. Alterations in the MAPK pathway may lead to increased expression of VEGF, a vascular endothelial growth factor, and other inhibitory cytokines, thus mediating immune evasion of

tumor cells [93,94]. Treatment of resistant cells with combination of MAPK inhibitors in mouse models resulted in the overcoming of immune checkpoint blockade resistance [94]. Constitutive activation of PI3K pathway due to loss of PTEN was associated with resistance to PD-1 therapy and decreased overall survival of patients with leiomyosarcoma [95,96]. Combined therapy against PD-1 and PI3K in a mouse model with head and neck squamous carcinoma demonstrated a decrease in inhibitory cytokine production and a modification of the tumor microenvironment, which ultimately lead to tumor regression [97,98]. The combination of different immune checkpoint and targeted therapy may improve response to treatments and better patient outcomes.

Melanoma is characterized by mutations in genes of key signaling pathways that result in cell proliferation and malignant phenotype. Approximately 70% of patients display one or more of these mutations. In the past few years, small molecules inhibitors or antibodies targeting mutated proteins focusing on BRAF and MEK have been developed and showed a rapid antitumor response and regression of this disease. Targeted inhibition of the mutated BRAF protein with the selective inhibitor vemurafenib was approved by the FDA in 2011. Vemurafenib reversibly bind to the kinase domain of BRAF competing with adenosine-triphosphate (ATP) ultimately inhibiting BRAF-induced MEK activation [99]. Melanoma patients positive for $BRAF^{V600E}$ mutations showed an improved clinical response rate, overall survival and progression-free period after treatment with vemurafenib compared to chemotherapy. In 2013, the FDA approved a second BRAF inhibitor (BRAFi) dabrafenib, which also results in a reduction of MAPK pathway signaling [33]. These drugs have a response rates of approximately 50% and result in an average free survival benefit of four months [100,101]. However, almost all patients rapidly develop resistance to RAF inhibitors after a period of approximately five months [100,102,103]. Resistance mechanisms to BRAF inhibition is often associated with intercellular reactivation of the MAPK pathway including expression of the kinases CRAF and COT1 or activating mutations in NRAS, MEK1 or AKT1 (Figure 3). In addition, aberrant splicing of BRAF, activation of PI3K via the loss of PTEN and activation of tyrosine kinases have been identified in patients with acquired resistance [104]. Several studies showed that targeting signaling effectors downstream of driver oncogenes are valuable strategy to overcome resistance. Trametinib a small-molecule inhibitor of MEK1/2, a downstream target of BRAF, was approved by the FDA in 2013 for treatment of BRAF-mutated malignant melanoma patients [105]. Monotherapies with trametinib showed an increased overall survival of six months compared with chemotherapy [106]. Trametinib showed activity in patients with NRAS mutations in contrast to the BRAF inhibitor vemurafenib. However, patients receiving trametinib suffered

from severe side effects including heart toxicity [107]. Cobimetinib, an oral selective MEK inhibitor, is another targeted drug for metastatic melanoma and is often used in combination with vemurafenib [108,109]. PSF and response rates were prolonged in patients with combined therapy compared to vemurafenib alone. However, adverse effects were observed in approximately 71% of all patients and melanoma patients pre-treated with BRAFi were less responsive to the combined targeted therapy [110].

Alterations in RTK signaling play an important role in primary resistance to BRAF treatment (Figure 4). Overexpression or activation of RTK may results in the activation of several parallel signaling pathways and thus lead to resistance. Several studies showed that RTK like EGFR are overexpressed in patients resistant to BRAF and MEK therapy [111]. PDGFR and IGF-1R are also frequently mutated in melanoma cells treated with both inhibitors, leading to the activation of PI3K pathway [42]. The PI3K pathway was also found to be activated by mutations or deletions of PTEN. When treated with BRAFi, patients with mutated PTEN showed a shorter progression-free survival compared to patients with wild-type PTEN [112,113]. Alterations in NF1 were identified in BRAF mutated cells resulting in RAS activation and increased RAS phosphorylation, thus mediating resistance to BRAF inhibitor but not MEK inhibitor therapy [48]. Hepatocyte growth factor (HGF) secreted by stromal cells activates HGF receptor (MET) leading to MAPK and PI3K activation, which translates into primary resistance to BRAF inhibition [114]. Patients with primary resistance may not benefit from targeted therapy, highlighting the need of identifying these patients before starting therapy and developing new therapeutic strategies. One of the most studied mutations related to acquired resistance to BRAF treatment are alterations in RAS. RAS mutant melanoma cells showed activated CRAF, MEK and ERK expression after BRAF inhibition, leading to BRAF/CRAF interaction followed by activated MAPK signaling [115,116]. Activating mutations in NRAS also lead to significant MAPK pathway reactivation after BRAF inhibition [115,117]. Secondary resistance to BRAF inhibition is also characterized by $BRAF^{V600E}$ alternative splicing and copy number amplifications. Several studies demonstrated that approximately 16% of patients with acquired resistance showed BRAF splice variants lacking exon 4-8, thus enhancing BRAF dimerization and ERK activation [118,119]. In approximately 12% of these patients BRAF amplifications were detected and could be linked to associated resistance [120]. Besides BRAF dimerization, BRAF and CRAF heterodimerization and ectopic expression of CRAF were also associated with ERK activating BRAFi resistance [121]. The switching between RAF isoforms as adaptors for MAPK pathway, can counteract and overcome BRAF inhibition in melanoma cells [122]. Consequently, inhibition of all three isoforms may be required to prevent ERK activation for melanoma patients with

acquired resistance. Downstream of the driver oncogenes BRAF and NRAS, activating MEK mutation (Q56P and E203K) were identified in patients with BRAFi resistance, resulting in a shorter PSF and poorer response (33%) compared to patients with MEK wild-type (72%) [123,124]. In another study, patients acquired resistance was associated with the presence of NRAS and MEK mutations, suggesting that mutated cells have a proliferation benefit under selective pressure to BRAFi [101]. Several RTK proteins, specifically PDGFR, IGFR and EGFR, are upstream of the driver oncogenes and their upregulation and overexpression result in activation of signaling pathways other than MAPK [115,118]. PDGFR and EGFR can be activated by MITF, a melanocyte-specific modulator. Loss of MITF inversely correlated with RTKs gain of expression in patients with secondary resistance [30,125]. Levels of MITF could help to predict early resistance to targeted therapy. Independently from MAPK and PI3K pathway signaling, Notch1 signaling was identified to be highly up-regulated in cell culture harboring $BRAF^{V600E}$ mutation and acquired resistance [126].

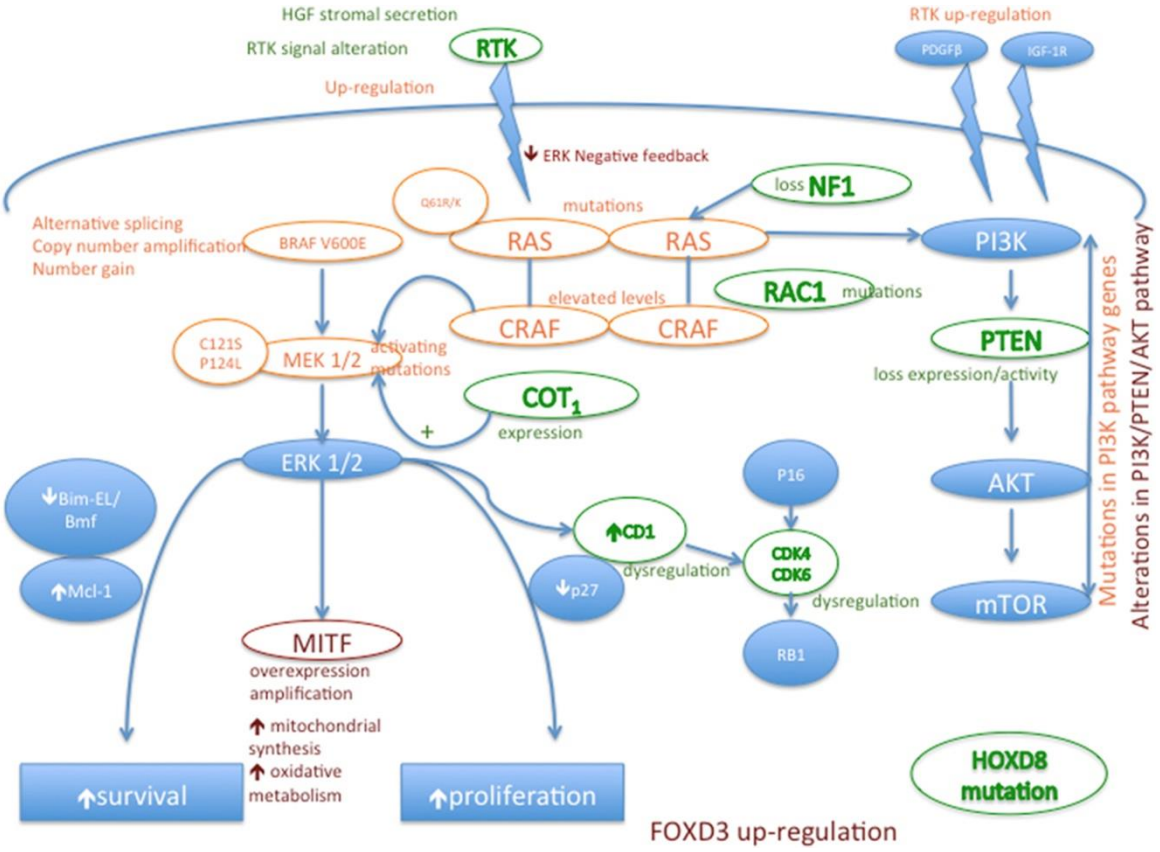


Figure 4: Mechanisms to targeted therapy resistance in melanoma. Resistance mechanisms are divided into primary resistance in green, secondary resistance in orange and adaptive resistance in dark red. (adapted from Amaral *et al.*, 2017 [30])

In order to overcome mechanisms of resistance, combined therapies have been established for melanoma treatment. The combined BRAF and MEK inhibition with dabrafenib and trametinib was approved by the FDA in 2014 [127]. Combined therapy showed an improved PFS of 9.4 months compared to monotherapies with dabrafenib (5.8 months) [128]. Several other studies confirmed the advantages of combining BRAF and MEK inhibitors using vemurafenib and cobimetinib and also demonstrated prolonged overall survival of patients with metastatic melanoma [128]. Despite the great advantages including delay of acquired monotherapy resistance, the combined treatment often leads to serve adverse events and a longer hospitalization is needed [129]. As in monotherapies, acquired resistance of combined therapies have been identified in patients with metastatic melanoma. In a recent study, ERK mutations were identified in resistant melanoma cells treated with BRAFi and MEKi, mediating a reactivation of the MAPK pathway [130]. Interestingly, patients resistant to BRAFi and MEKi were not resistant towards ERKi and vice versa. This suggest a new therapeutic strategy, whereby switching inhibitors may delay resistance development. Heterodimerization of $BRAF^{V600E}$ with either CRAF or MEK were also observed in cells with combined drug resistance, leading to MAPK pathway activation [131]. Combined targeted therapy of key molecules of the MAPK pathway may be insufficient to avoid development of resistance. Several groups are currently investigating targeting the oncogenic PI3K pathway with the combination of PI3K/AKT inhibitors together with BRAFi and MEKi [132,133]. In summary, targeted therapy and immunotherapy offer great advantages over conventional chemotherapy for melanoma treatment. However, drug resistance hampered the prolongation of progression-free survival and response rates. Resistance mechanisms and even cross-resistance between combined targeted therapy is still not fully understood and need further investigation.

1.1.4 Personalized medicine in melanoma

Over the last decade, several therapeutic treatments for melanoma have been developed with improved efficiency and overall survival rates. Only a minority of patients benefits from the current treatment therapies and most of these patients develop resistance [59,82,134]. Malignant melanoma is known to be highly heterogeneous and characterized by its high frequency of somatic mutations compared to other cancer types [135]. Some molecular events are more frequent in some patients and provide an opportunity to adjust the treatment for individual patients. Classical systems are limited by the prediction and prognosis of treatment response and thus the development of more effective targeted therapies for each individual are needed.

Personalized medicine will hopefully lead to an improvement of preventive measures, early diagnosis and the identifications of drug targets and biomarkers. The past decade has seen a revolution in the field of nucleotide sequencing, which now allows for routine nucleotide-based investigations of tumor samples. Different next generation sequencing (NGS) strategies are used in precision medicine allowing for sequencing of just a few hotspots, sequencing of large gene panels or even whole genome sequencing (WGS). Smaller panels like hotspot sequencing panels or actionable sequencing panels are a collection of frequently mutated genes that are either clinically actionable or have diagnostic significance [136]. Larger panels focus either one the entire region of a target gene or disease-associated regions to screen for risks of inherited diseases (germline mutations) as well as to identify tumor driver mutations (somatic mutations). Whole exome sequencing (WES) or WGS provide comprehensive tools to study the complete coding region (WES), as well as intronic and other non-coding DNA regions (WGS). Only 1-2% of the genome codes for proteins. In addition, the majority of targeted therapies do not target the genome, but rather the protein in cancer cells. Thus, it is highly important to quantify and identify proteins through all areas of personalized medicine, including biomarker detection and response prediction. Mass-spectrometry based proteomics allows the identification and quantification of thousands of proteins in complex samples. Several proteomic studies identified proteins that were mis-regulated in melanoma cell lines and patients and some of them showed a positive correlation with tumor progression and patient survival rates. For example, Welinder *et al.* reported 288 proteins that showed a positive correlation between protein expression and disease outcome for six patients with stage III metastatic melanoma [137]. Four of these are currently used as melanoma markers including melanoma-associated antigen D2 (MAGD2) and melanoma cell adhesion molecule (MUC18). In addition, high levels of lactate dehydrogenase and S100 calcium-binding protein B (A100B) were found to be correlated with melanoma stage and progression [138]. Several biomarkers were identified in serum samples of patients with melanoma and were associated with melanoma progression, recurrence and survival [138]. Diagnostic biomarker identifiably in blood of patients would make them clinical relevant for the prediction of responses to both immunotherapy and targeted therapy. Personalized medicine incorporates the molecular and the biological aspects of each patient's disease, thus allowing the selection of optimal treatment on the basis of improved efficiency, overall survival and reduced adverse events.

1.2 Key proteins relevant in this study and their link to melanoma

1.2.1 Intermediate filament nestin

Nestin is an intermediate filament VI, a component of the cytoskeleton and was originally described as a stem cell marker. It was first identified in neuroepithelial stem cells and other progenitor cells during migration and proliferation phases in early embryonic development [139,140]. In adult tissue it is involved in regeneration processes remodeling the cytoskeletal actin network [141]. The human *NES* gene is located on chromosome 1 at position 23 and consists of four exons, separated by three introns. The protein comprises 1621 amino acids (aa) with a highly conserved α-helical rod domain (306 aa) flanked by a small N-terminal head (7 aa) and a long C-terminal tail domain (1308 aa) [142] (Figure 5). The rod domain is composed of four α-helical coils (1A, 1B, 2A and 2B) and is essential for dimerization.

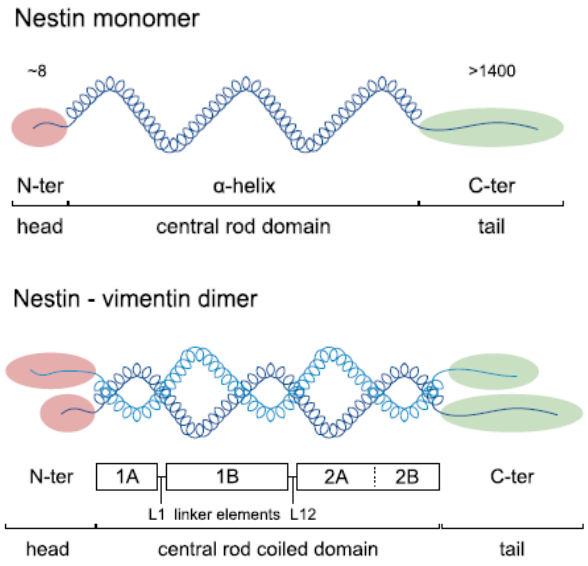


Figure 5: Nestin protein structure with functional domains and heterodimer formation with vimentin. The nestin protein (1621 Aa) is composed of a N-terminal (N-ter) head domain, an α-helical rod domain and a C-terminal (C-ter) tail domain. Heterodimer formation of nestin and its interaction partner vimentin through the central rod coiled domain. (adapted from Bernal *et al.*, 2018 [143])

Nestin forms heteropolymers with other filaments, mostly vimentin and keratin, mediated by the rod domain [144,145] (Figure 5). The formation of heteropolymers plays an important role in the organization of the cytoskeletal network and is regulated by PTMs mostly phosphorylation [143]. Phosphorylated nestin triggers the disassembly of other bound cytoskeletal filaments like vimentin, keratins and nuclear lamin through the disruption of cross-links [143]. The dynamic regulations of the cytoskeletal network mediated by nestin and other filament proteins are

important in many cellular processes such as cell survival, proliferation and cell motility [144]. Several studies described the role of nestin in a variety of tumors like pancreatic cancer, prostate cancer, glioblastomas, breast cancer and especially in malignant melanomas [146-150]. Piras *et al.* showed that nestin was highly expressed in cancer stem cells, which are involved in tumorigenesis and malignancy [151]. In addition, the expression of nestin correlated with a poor prognosis for patients with malignant melanoma. This was also observed with increased metastasis, invasion and migration in pancreatic cancer and neurogenic and mesenchymal tumors [142,145,150]. To counteract theses effects, several studies used shRNA and siRNA to inhibit nestin expression in melanoma and liver cancer mouse models and lung adenocarcinoma [152,153]. The inhibition of nestin resulted in less metastases in the mouse models and showed reduced cell proliferation, migration and invasion of cancer cells. In addition, nestin was also described to be involved in angiogenesis [154]. In endothelial cells of blood vessels, high expression of nestin was observed and inhibition of nestin-positive mouse blood vessels led to a decrease in blood vessels close to the tumor and consequently to tumor regression [154]. Quendro *et al.* identified in a proteomic screen an increased expression of nestin in late stages III and IV of metastatic melanoma and a correlation with the aggressiveness of the subtypes, suggesting nestin as marker of melanoma staging [141]. In cutaneous melanomas of the nodular type, nestin expression was observed in 92% of cases and associated with increased tumor thickness, ulceration and increased proliferation [149]. In addition, nestin expression was also reported to be up-regulated in melanoma cell lines and depletion of nestin revealed an activation of matrix metalloproteinases suggesting a stronger invasive phenotype of the cancer cells [155]. In summary, nestin expression indicates a poor prognosis and a low survival rate in many different types of cancer [151,152]. Nestin over-expression in malignant melanoma could be a potential prognostic biomarker for the aggressiveness and invasiveness of the tumor [141,147].

1.2.2 Transcription factor RUNX1

The transcription factor RUNX1 belongs to the family of the runt related transcription factors (RUNX), consisting of the members RUNX1, RUNX2, RUNX3. The RUNX family was reported in several key processes like cell proliferation, differentiation, senescence and apoptosis [156]. All members of the family share an evolutionary conserved runt domain of 128 aa at the N-terminus mediating DNA binding [157]. The three RUNX proteins are encoded on different genes and show distinct tissue-specific expression pattern [156]. The structures of the three RUNX proteins are similar, however, the function differs including roles in neurogenesis,

gastric epithelial cell proliferation (RUNX3), osteogenesis (RUNX2) and hematopoiesis (RUNX1) [14,158]. RUNX1 was originally described in chromosome aberrations related to human acute myeloid leukemia. Alternative splicing of the RUNX1 gene leads to eleven isoforms expressed in different cell types and development pathways like TGFβ, WNT or NOTCH signaling [159]. Binding to the core-binding factor subunit beta (CBFB) enhances the transcriptional activity and DNA binding of RUNX1 [160]. DNA binding is mediated by the runt domain and leads to a conformational change and a specific spatial reorganization of the amino acids that are involved in DNA recognition. RUNX1 recognizes the core sequence 5'-YGYGGT-3' in the DNA sequences [161]. RUNX1 also mediates CBFB independent functions by binding to other transcriptional cofactors and chromatin modifiers like histone deactelylase HDAC1 or acetyltransferase p300 [162]. Binding to transcriptional activators and suppressors regulates the transcription of hematopoietic genes especially during differentiation of B cells, T cells and myeloid cells. Several studies suggest that disrupted or mutated RUNX1 gene can lead to various malignancies of the hematological system like platelet disorder, myelomonocytic leukemia (MLL), lymphocytic leukemia (ALL), or acute myeloid leukemia (AML) [163,164]. RUNX1 was found to be mutated in around 2.04% of all cancer types and so far 43 mutations are described in the literature [165]. In around 20% of adult AML patients therapy-related myeloid neoplasms and fusion genes through chromosomal aberration were identified may influence the protein stability and function [166]. The most prominent gene fusion is RUNX1-ETO resulting from the runt domain of RUNX1 and the four nervy homology regions of (NHR1-4) of ETO. The fusion complex enhances the interaction with transcriptional regulators leading to self-renewal of hematopoietic stem cells (HSCs) and leukaemogenesis [167]. RUNX1 expression and mutation was not only reported in leukemia, but also in various solid tumors, like breast cancer, oesophageal adenocarcinoma and epithelial cancers including skin squamous cell carcinomas [168-170]. In some cancer types activity of RUNX1 was associated with a strong tumor suppression, but oncogenic in others. Recently, RUNX1 was found to be involved in resistance mechanisms to $BRAF^{V600E}$ inhibition in malignant melanoma [171]. In this study, RUNX1 mediated up-regulation of the receptor tyrosine protein kinase CSF1R was identified leading to growth and invasion via activation of the ERK and PI3K/AKT pathways. In resistant melanoma cell lines, increased expression of CSF1R was observed with concomitant up-regulation of the CSF1R ligand IL34, which were shown to correlate with tumor progression, invasion and acquired resistance to BRAF inhibitors. Thereby, ERK pathway activity stimulates RUNX1 to activate transcription of CSF1R and its ligand IL34. Co-expression of both the receptor and the ligand causes high oncogenic potential by enabling para- and autocrine

activation. RUNX1 is regulated by several regulatory mechanisms including splice variants, transcriptional control by two different promoters and post-translational modifications [172]. Methylation, acetylation and phosphorylation regulate RUNX1 transcriptional activity, whereas the interplay between phosphorylation and ubiquitylation regulate the stability of the protein. During cell cycle progression, RUNX1 gets phosphorylated by the cyclin dependent kinases (CDKs) on S276, S293, T300 and S303 in the M phase of the cycle, which leads to the anaphase-promoting complex mediated degradation of RUNX1 [173,174]. Upon cytokine stimulation, RUNX1 can be also activated through phosphorylation by ERK1/2 kinases at positions S249, S266 and S276 enhancing the interaction with p300 and transcriptional activity [175]. Phosphorylation of RUNX1 can also disrupt the interaction with the transcriptional suppressor SIN3A leading to carcinogenic processes in some types of cancer [176]. The interaction with SIN3A can be also abolished by PRMT1-mediated arginine methylation resulting in a higher activity of RUNX1 [177]. In leukemia, phosphorylation by the homeodomain-interacting protein kinase 2 (HIPK2) was reported [178]. Triple-phosphorylation of RUNX1 by HIPK2 on S249, T276 and S273 mediates the binding and activation of the histone acetyltransferase p300, which results in up-regulation of target gene expression. In turn, p300 acetylates RUNX1 on K24 and K43, which enhances its DNA binding capacity and thereby influence the transcriptional activity [179]. P300-mediated acetylation of RUNX1 can be also stimulated by TGFβ signaling. Deregulation of PTMs or mutations may disrupt RUNX1-mediated differentiation and direct cells towards cancer fate.

1.2.3 Guanine nucleotide exchange factor DOCK1

The guanine nucleotide exchange factor (GEF) DOCK1 belongs to the eleven-member family DOCK mediating direct GDP/GTP exchange to promote cytoskeletal reorganization in cell proliferation, differentiation and migration [180,181]. DOCK1 contains two evolutionarily conserved domains, the DOCK homology region-1 (DHR-1) and the DHR-2 domain [182]. DOCK1 interacts with specific members of Rho family such as RAC or Cdc42 but not with the Rho protein RhoA. The DHR-1 domain mediates the recruitment of DOCK1 to the membrane by direct binding to PIP3 after PI3-kinase activation. Recent investigations suggest that DOCK1 binds the signaling lipid phosphatidic acid (PA) instead of PIP3 [183-185]. Sanematsu and colleagues showed that PDGF treatment in fibroblasts promotes the DHR-1-dependent and PIP3-dependent recruitment of DOCK1 and DOCK5 to the membrane, resulting in RAC1-dependent peripheral membrane ruffle formation [185]. The binding of DOCK1 to PA via the

DHR-1 domain is important for the localization of DOCK1 at the membrane, resulting in the activation of RAC1 to form dorsal ruffles. Both RAC-dependent membrane ruffles are involved in specific biological processes such as cell migration for peripheral ruffles and cell invasion linked to dorsal ruffles. DOCK1-mediated cytoskeletal remodeling is also regulated by the formation of the complex between DOCK1 and ELMO scaffolding proteins (ELMO1 and ELMO2), binding to the N-terminal SH3 domain of the protein [186]. The ELMO scaffolds directs DOCK1 to distinct areas of the cell including the cell membrane to allow activation of RAC1 and their downstream signaling pathways. In the basal state, both proteins are in the closed confirmation and extracellular stimulation release ELMO autoinhibition followed by activation of DOCK1 and optimal RAC activation [187-189]. The C-terminus of the DOCK1 GEF protein, including a PBR domain and a 'PXXP' region, second mediates the interaction with SH3-containing adaptor proteins, such as CRK and GRB2. DOCK1 activity is regulated by several post-translational modifications including phosphorylation, acetylation and ubiquitylation. Phosphorylation on Y722, Y1811 or S1250 by SRC and PKA kinases increases its GEF activity towards RAC activation, while ubiquitylation of DOCK1 was enhanced after EGF stimulation and binding to CRK [190,191]. Binding of ELMO1 inhibited the ubiquitylation of DOCK1 and thus stabilizing the protein [190]. This regulation of GEF activity by PTMs and binding of interaction partners might contribute to the distinct activation of RAC at the plasma membrane. Several studies described the roles of DOCK1 in a variety of tumors such as glioblastoma, thyroid cancer, breast cancer and malignant melanoma [192-194]. In glioblastoma, DOCK1 and its complex partner ELMO1 were observed to be highly expressed in the invasive areas of the cancer tissue sections and could be linked to the PDGFR- induced downstream signaling in glioblastoma cell lines [192,193]. In addition, they could show that suppression of DOCK1 expression prevents the cell migration and activation of downstream targets such as RAC, ERK1/2 and AKT1. Activation of PDGFR promotes the phosphorylation of DOCK1 at Y1811 through SRC kinase, which leads to the interaction with CRK and BCAR1 followed by activation of RAC and downstream targets like AKT and ERK1/2 [193]. Phosphorylation of DOCK1 at Y722 by SRC kinase and S1250 by PKA kinase can also be induced after activation of oncogenic EGFRvIII to increase its affinity to RAC and promote GTP load [191,195]. Mutations of these different phosphorylation sites of DOCK1 prevents growth and invasion of oncogenic RTK overexpression in cancer cells. Notably, the overall survival of patients with glioblastoma was significantly decreased in tumors with DOCK1^{Y1811F} and PDGFRA mutations [193]. In breast cancer, high levels of DOCK1 mRNA expression was associated with a poor prognosis for patients with either HER2-positive or basal breast cancer, suggesting that DOCK1 is an

important downstream effector of HER2 [194]. Inhibition or knockout of DOCK1 *in vivo* in breast cancer cells showed that DOCK1 expression is essential for the activation of RAC though HER2-dependent activation of DOCK1 and thus regulates cancer metastasis and migration [194]. In human melanoma cells, DOCK1 expression was linked to invasion through complex formation with the focal adhesion kinase FAK and CRK [196]. This complex formation leads to RAC activation and induced JNK activity directly regulating actin cytoskeletal reorganization and formation and secretion of matrix-metalloproteases MMPs. In addition, Tomino *et al.* showed that DOCK1 promotes the GDP/GTP exchange of oncogenic RAC^{P29S} resulting in increased invasion and metastasis in melanoma and breast cancer cells [197]. This somatic mutation on RAC1 was reported in 9% of sun-exposed malignant melanomas and leads to self-activation due to increased inherent GDP/GTP exchange [29]. In another study, they reported a DOCK1-mediated RAC activation in RAS-driven cancer cells promoting cellular invasion and macropinocytosis [198]. Similar as in RAC mutated cells, inhibition of DOCK1 suppresses cell growth and metastasis of RAS-transformed cancer cells *in vivo*. The RAC^{P29S} mutation was also reported to be associated with resistance to kinase inhibitors and immunotherapy and could serve as a predictive biomarker for therapy resistance in melanoma [199,200]. Inhibition of DOCK1 suppressed the $RAC1^{P29S}$-induced invasion in cancer cells and might be a potential treatment option in melanoma patients associated with RAC mutations [197,198].

1.3 CRISPR/Cas9 system – a new tool for genome editing

Prokaryotes have developed innate and adaptive immune systems to cope with the constant threats of phage infections and plasmid transfer. The adaptive immune system clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) genes was identified in 40% of bacterial and 90% archaeal genomes [201]. The CRISPR immunity is based on small fragments of foreign DNA of previous infections (so called spacers) integrated in the host cells CRISPR array [202,203]. The spacers are of 26-47 base pairs (bp) long and separated by short repetitive sequences of 2-48 bp. The CRISPR system has been classified into three classes based on the core element content and sequences. Type I and III systems require large multisubunit ribonucleoprotein complexes for the silencing of foreign nucleic acids in subsequent rounds of infections. In contrast, in the type II system the DNA cleavage of foreign viruses and plasmids is mediated by only a single Cas protein bound to a dual RNA molecule. The CRISPR/Cas9 system from *Streptococcus pyogenes* (*S. pyogenes*) is the most commonly used system for genome editing using the well characterized Cas9 endonuclease [204,205] (Figure

6). The CRISPR associated (Cas) genes are adjacent to the CRISPR locus. During the immunization process, the foreign DNA is cleaved by the Cas complex and the short fragments are integrated into the host CRISPR repeat-spacer locus as spacers between CRISPR RNA (crRNA) repeats [206,207]. In the second stage, the CRISPR array containing acquired spacers is transcribed into precursor crRNA and enzymatically processed and cleaved into mature crRNA. The crRNA consists a spacer sequence at the 5'end, which is responsible for targeting it to the foreign genomic element, as well as the part of the crRNA repeat sequence allowing the recognition by the Cas protein. In the type II CRISPR/Cas9 system, hybridization between the crRNA repeat sequence and a noncoding trans-activating CRISPR RNA (tracrRNA) is critical for the processing of the crRNA and the target-mediated cleavage by Cas9 [208]. In the third stage of the immune response, the crRNA spacer hybridizes with complementary foreign DNA and the Cas9 protein mediates the cleavage of the invading genome upon a second infection [206].

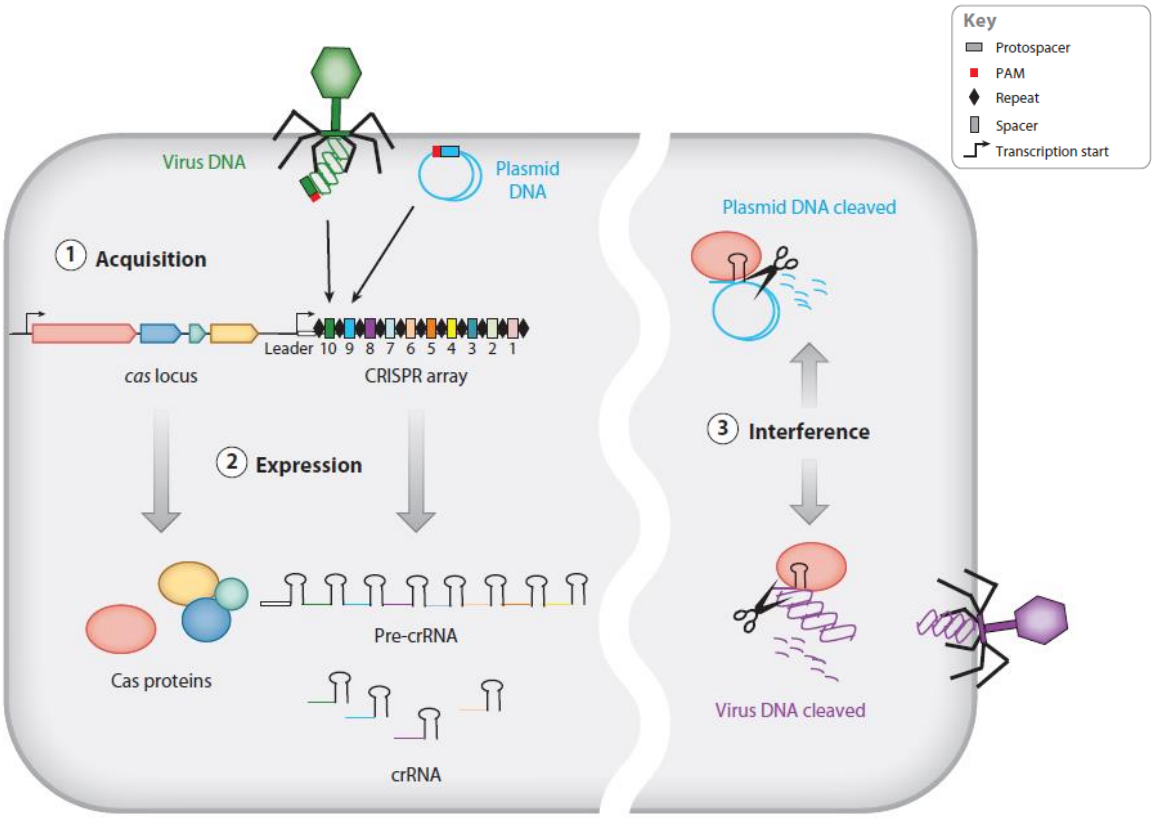


Figure 6: Different phases of CRISPR/Cas9-mediated immunity in type II system in bacteria. In the CRISPR spacer acquisition phase (1), invading foreign DNA is cleaved by the Cas complex into short sequences (spacers, colored boxes). These spacers are inserted into the CRISPR locus between short palindromic repeat sequences (repeats, black diamonds). In CRISPR expression phase (2), the precursor CRISPR RNA (pre-crRNA), and Cas9 proteins are transcribed and subsequently processed. The foreign DNA sequence is recognized, in phase 3, by the crRNA-Cas9 complex via complementary base pairing and subsequently cleaved by the Cas9 protein. It is essential that the target sequence is in close proximity to a protospacer adjacent motif (PAM; red). (adapted from Bhaya *et al.*, 2011 [206])

The action of CRISPR/Cas9 depends on the assembly of the crRNA and the Cas9 protein to interrogate DNA targets. The complex recognizes a sequence-specific PAM motif located in close proximity of the crRNA target site in the invading genome [209]. This motif is absent in the host genome, which protects it from self-cleavage. In the class II CRISPR/Cas9 system, the cleavage of the DNA is mediated by the nuclease domain of the Cas9 protein, a HNH domain. This domain cleaves the complementary target strand to the gRNA and the RuvC-like nuclease domain of the Cas9 protein cleaves the non-target strand [210,211]. Cas9 is guided to the DNA by the crRNA-tracrRNA duplex, which is unique for the class II system. The nuclease Cas9 from *S. pyogenes* recognizes the PAM motif 5'NGG (N represents any nucleotide) in the foreign DNA.

Recent achievements in genome editing techniques allow the precise manipulation of any gene at its genomic locus in a variety of experimental models including cell lines, laboratory animals, plants and was even tested human clinical trials [211]. The specificity of genome editing is based on 'programmable' nucleases producing specific double strand breaks (DSBs), which are repaired by endogenous cellular repair mechanisms. Zinc-finger nucleases (ZFNs) and transcription activator-like (TAL) effector nucleases (TALENs) were the first engineered systems [212-214]. The DNA-binding domains of transcription factors have been fused with the nuclease domain of the restriction enzyme FOKI [211]. When targeted to specific site in the genome, the nuclease domain of FOKI forms a dimer that activates the nuclease activity mediating a DSB near the target site. These systems function through DNA-protein interactions and require engineering and cloning of proteins for each target site. The CRISPR/Cas9 system is a powerful RNA guided DNA platform, which can be used for high-throughput applications (Figure 7). The crRNA-tracrRNA complex can be fused into a chimeric single guide RNA (sgRNA), the basis of the flexible genomic engineering toolkit [209,215,216]. The sgRNA-Cas9 complex induces the cleavage of a specific target site adjacent to a PAM sequence. Essentially any genomic locus containing a PAM motif can be targeted with a customized sgRNA.

The DSB created by Cas9 triggers endogenous DNA repair mechanisms, such as non-homologous end joining (NHEJ) or homology-directed repair (HDR) pathway [218,219] (Figure 7). The NHEJ pathway causes random insertions or deletions of nucleotides at the DSB site, so called InDel mutations. These may lead to gene knockouts by causing a shift in the target gene reading frame, premature stop codons or mutations of a critical region in the encoded protein. Error-prone NHEJ is also used for loss of function screening, genomic rearrangements and NHEJ-mediated homology independent knockin of genes [220,221]. HDR can be exploited to generate precise modifications in a defined locus at the DSB site through homologous

recombination guided by exogenous donor repair templates. The HDR pathways provides a powerful tool for sequence-specific gene editing including gene knockout, gene knockin, mutagenesis and gene corrections [222]. In addition, it is also often used to introduce a tag or a reporter gene to study protein localization or protein-protein interactions [223].

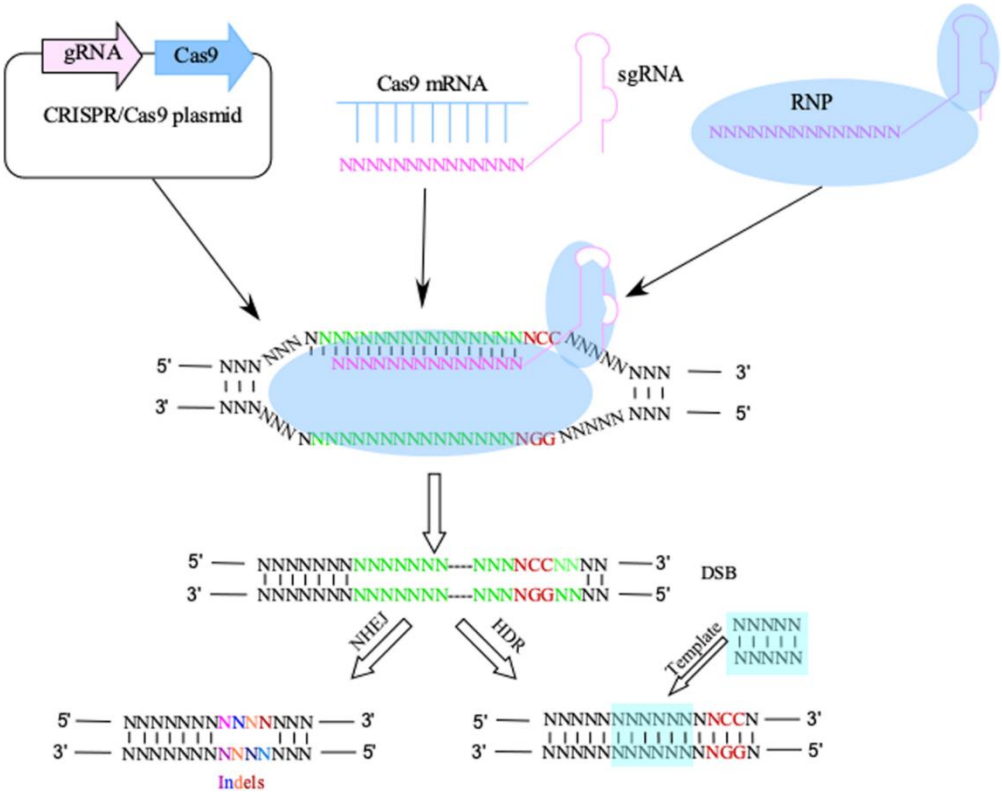


Figure 7: CRISPR/Cas9 based gene modification by commonly used delivering methods and DNA repair mechanisms. Plasmid containing the mRNA sequence of the Cas9 protein as well as the sgRNA, Cas9 mRNA in complex with sgRNA or a ribonucleoprotein (RNP) complex methods can be used. The sgRNA (pink) bind to the target site (green) of the genomic DNA via complementary base pairing and the Cas9 protein induces a double strand break (DSB) three base pairs upstream of the PAM sequence (red). The double strand break can be repaired by two endogenous cellular repair mechanisms. The non-homologous end joining (NHEJ) pathway results in random insertions or deletions (InDels) of nucleotides. The homology-directed repair (HDR) pathway requires a donor template with homologous sites. Thus, precise genome editing is enabled by a specific donor sequence. (adapted from Tian *et al.*, 2018 [217])

Despite the great potential of CRISPR/Cas9, several limitations and challenges exist and can still be improved concerning specificity, efficiency and control. The PAM sequence motif of Cas9 limits the availability of genomic targets to an average of one target site per eight bp. Several Cas9 protein variants with altered PAM sequences have been developed to improve Cas9 specificity and availability of target sites in the human genome [224]. For example, xCas9 recognizes several PAM sequences including 5'NG, 5'GAA and 5'GAT [225]. Another challenges are possible off-target effects as the 20-bp targeting sequence of the crRNA and the 3-bp PAM

sequence may exist elsewhere is the genome. To reduce off-target effects several Cas9 variants have been generated to improve Cas9 sensitivity [226]. Shen *et al.* (2014) engineered a Cas9 protein with an inactive nuclease domain, which functions as a nickase, cleaving only one DNA strand [227]. This strategy was applied to cancer cell lines and mouse embryos and showed reduced off-target effects and improved specificity. In addition, a catalytically deactivated Cas9 variant (dCas9) was developed and can be used for sequence-specific gene regulation [228]. Transcriptional activators and repressors can be fused to the dCas9 and this complex can be directed to any genomic locus by specific sgRNAs and serve as a RNA-guided DNA-binding platform. This technique is termed CRISPR interference (CRISPRi) and widely used to repress or activate DNA transcription by blocking RNA polymerase binding, elongation or the binding of transcription factors [229]. Multiple targeting is also possible and the effects of dCas9 binding are reversible. In addition, by tagging dCas9 with an enhanced green fluorescent protein, it can be also used as an imaging tool for example for specific sequences or repetitive elements [230].

1.3.1 CRISPR/Cas9-mediated genome engineering in cancer biology

In the recent years, CRISPR/Cas9 has revolutionized the field of genomic engineering and helped to understand and manipulate biology in a variety of studies ranging from basic research to clinical applications [231,232]. The CRISPR/Cas9 system has been used in several studies to investigate rearrangements of the genome in cancers and other diseases. For example, chromosomal translocations could be precisely reproduced by joining two chromosomes via the NHEJ pathway in lung cancer and acute myeloid leukemia cell lines. This has been even further developed to generate cancer mice models by CRISPR/Cas9 to study cancer progression and malignancy. In addition, the CRISPR/Cas9 system was also successfully applied to high-throughput screenings to identify genes essential for cell viability in cancers, which enables the identification of potential drug targets. In a breast cancer study, CRISPR/Cas9 was used for the diagnosis, treatment and drug resistance research. Using the dCas9 system, the CDKN2A gene was identified as a diagnostic marker indicating abnormal cell division in breast cancer cells [233]. In another study, the HER2 gene was targeted by CRISPR/Cas9, revealing a correlation between HER2 expression and inhibition of cell growth and attenuated tumorigenicity [234]. For leukemia models, lentivirus-delivered Cas9-sgRNA systems in primary hematopoietic cells has been developed and was used to target inactivated genes including TET2 and RUNX1 to study development of myeloid malignancy [235]. The targeted studies enhance the development of precision cancer medicine and provide a powerful tool to study functional cancer genomics.

Acquired resistance to drugs is one of the major challenges in cancer therapy. Resistance in leukemia is often mediated by synergistic gene target interactions, which can be identified by sgRNA library screenings for combinatorial genes in CRISPR/Cas9 based knockout cell lines [236]. The depletion of these combinatorial genes showed great promise in leukemia cells and may allow the development of personalized genotype-based therapies. In precision cancer medicine also screens for cancer metastasis-related genes using CRISPR/Cas9 loss of function libraries are often used to identify new therapeutic targets and potential biomarkers [237-239]. Gene therapy by the CRISPR/Cas9 system could be consequently a powerful tool to treat genetic disorders in the future. For example, in mice, researchers corrected the *CFTR* locus which is responsible for cystic fibrosis in intestinal stem cells and the mutated β-globin locus in sickle cells disease in hematopoietic stem cells by genome editing. The CRISPR/Cas9-based diagnosis panels SHERLOCK and DETECTR rely on the Cas13a or Cas12a RNA-guided RNase which mediates a non-specific single strand cleavage [240-242]. For SHERLOCK; a reporter signal released after RNA cleavage is used to detect common driver mutations in cancer like $BRAF^{V600E}$ or $EGFR^{L858R}$ in highly sensitive diagnostic approach [242]. Recombinase polymerase amplifications (RPA) are used in the DETECTR system to amplify micro samples and detect infections in cancer like HPV types in lung carcinomas. Promising results from CRISPR/Cas9 studies have been achieved, however more work is needed to develop a safe and effective tool for diagnosing and treating cancers.

1.4 Mass spectrometry based proteomics and proteogenomics

1.4.1 Proteome and post-translational modifications

Biological processes are often carried out by proteins and their characterization is of fundamental interest in cell and cancer biology. The proteome comprises all expressed proteins within a cell, tissue or organism at a given time under defined conditions. The human genome consists of approximately 20,000 protein-coding genes, however complexity increases from the transcriptional to the translational level due to alternative variants and post-translational modifications (PTMs) [243-245] (Figure 8). Not only the complexity of the proteome also the abundance, dynamic range and the localization of proteins plays also an important role in biological functions [246,247].

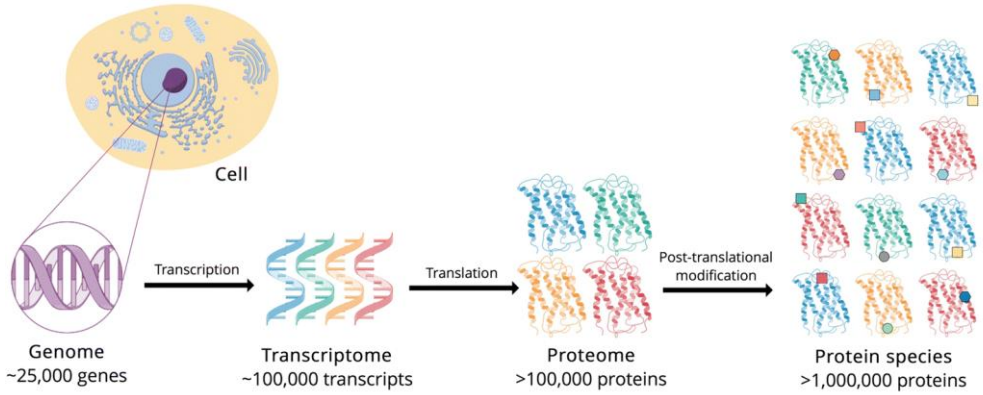


Figure 8: The increase in complexity from genome to proteome. The genome comprises 20,000 to 25,000 genes, the proteome is estimated to encompass over 1 million proteins. Transcriptional and mRNA level changes increase the size of the transcriptome relative to the genome, and post-translational modifications of proteins exponentially increases the complexity of the proteome. (adapted from Virag *et al.*, 2019 [245])

Proteomics, the large scale study of proteins, is widely used for the analysis of cellular compositions under changing conditions. Mass spectrometry (MS)-based proteomics enables the global analysis of the protein composition, protein abundance, PTMs state of proteins and their dynamic processes in the cell [248]. Post-translational modifications of proteins regulate a broad variety of biological processes including cell growth, proliferation and apoptosis. Over 300 PTMs have been identified ranging from small chemical moieties such as acetylation or phosphorylation to more complex structures like glycosylation or ubiquitylation [249-251]. Phosphorylation of proteins is the most prominent PTM and can influence signal transduction networks, protein activity, protein-protein interaction and sub-cellular localization [252]. Protein phosphorylation, mediated by protein kinases, is a reversible process and the removal of phosphorylation is performed by protein phosphatases. The human proteome encodes for more than 500 protein kinases and roughly 150 protein phosphatases accounting for up to 3.5% of the proteome. Most kinases are specific for phosphorylating one residues. Several amino acids can be phosphorylated and divided into four groups: (a) O-phosphorylation at serine, threonine and tyrosine, (b) N-phosphorylation at arginine, histidine and lysine, (c) S-phosphorylation at cysteine and (d) acyl-phosphorylation at aspartic acid and glutamic acid. O-phosphorylation is the best studied type of phosphorylation due the chemical stability and compatibility with proteomic workflows. The stoichiometry of phosphorylation is generally relatively low at a given time point and usually not all copies of a protein are phosphorylated or multiple sites may be variably regulated during different processes. Recent phosphorylation studies have revealed that the majority of proteins in a mammalian cell are phosphorylated at one or more residues [253]. Many other types of PTM having become more prominent and are now extensively studied

by proteomics like acetylation and methylation of non-histone proteins [254]. Acetylation on lysine residues has been associated for a long time with the regulation of chromatin structures and was described in the 1964 for the first time [255,256]. In the recent years, several proteomic studies identified over 5,000 acetylation sites on over 2,000 proteins localized in the cytoplasm, mitochondria and plasma membrane [257,258]. In addition to acetylation, ubiquitylation and SUMOylation are also lysine modifications and play an important role in protein degradation as well as other cellular regulatory functions such as apoptosis and DNA repair [259]. Many proteins are sequentially modified at multiple residues as one PTM can serve as a positive or negative signal for the addition or removal of a second PTM or for the recognition of other binding partners that may modify the protein further [260]. Bioinformatic tools and databases can be used to predict hotspots for different types of PTMs on the same protein and revealed that these modifications are often in close proximity to each other, about 15 amino acids apart [261]. Several proteomic studies obtained evidence about the crosstalk of PTMs in eukaryotic and prokaryotic cells including phosphorylation and ubiquitylation during the process of protein degradation or SUMOylation and phosphorylation for the activation of kinases like CDKII [262,263]. PTMs and even the crosstalk between PTMs expand the landscape of the proteome and functions as a fine-tuning mechanism regulating the function, localization and interaction of proteins.

1.4.2 LC-MS/MS instrumentation

The field of mass spectrometry based proteomics is rising since the last decades due to technical investigations. The work of Thomson in 1912 and his student Aston in the next few years led to the development of the first mass spectrometer and the detection of isotopes of elements [264]. Both scientists were awarded with the Nobel prize in 1906 and 1922 for their pioneering work in the field of physics and chemistry, respectively. Developments in instrumentation including soft ionization techniques, increased sensitivity and resolving power moved mass spectrometry from chemistry into the field of biology, where it becomes a powerful tool to study and characterize thousands of proteins. Mass spectrometry is the measurement of the mass-to-charge (*m/z*) ratio of an ionized molecules [265]. Mass spectrometer consist of an ion source that converts the analyte of interest from liquid to gas phase, a mass analyzer that separates charged ions based on their *m/z* ratio and a detector that records the numbers of ions at each *m/z* value [266]. Two soft ionization techniques, called matrix-associated laser desorption ionization (MALDI) and electrospray ionization (ESI) are used for many proteomic configurations and

the efficiency of ionization can differ by order of magnitudes for different peptides [267]. For MALDI, the analyte is embedded in a crystallized matrix and will be vaporized and ionized by laser impulses, which initiates an energy transfer from the matrix to the molecule [268,269]. ESI takes place at atmospheric pressure and the analyte molecules are ionized and transferred from liquid to gas phase [270] (Figure 9). The analyte is transported through a needle at high electrical potential (2-6 kV). An electric field between the end of the needle and the entry capillary of the mass spectrometer is applied and the charged analyte will form a Tylor cone and droplets will be released. The solvent of the formed droplets will evaporate and the charged analyte enters the gas phase (Figure 9). Most of the ions introduced by ESI will be multiply charged in contrast to MALDI.

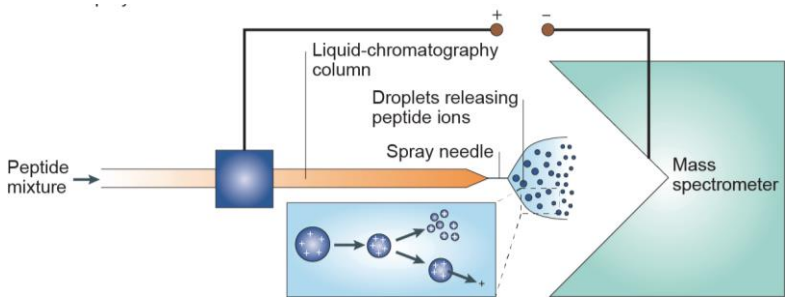


Figure 9: Electrospray ionization for biomolecules. An electric field between the end of the needle and the entry capillary of the mass spectrometer is applied and the charged droplets will be released. The solvent of the formed droplets will evaporate and the charged analyte enters the gas phase. (adapted from Steen and Mann, 2004 [271])

The charged ions are separated by the mass analyzer in the MS instrument based on their *m/z* ratio. In addition, ions can be also temporarily trapped and even detected in some types of analyzers [272]. Mass analyzers varying in their resolving power, mass accuracy and sensitivity based on their ability to separate ion with similar or near-identical *m/z* values [273]. The resolution (R) refers to the ability to separate two narrow mass spectral peaks and the mass accuracy also depends on the signal-to-noise ratio for each mass spectral peak [265]. A quadrupole mass analyzer consist of four parallel round metal rods, connected as opposite pairs [274,275]. A direct current voltage and a radio frequency are applied to the rods. Under a fixed electric field, ions will oscillate between the rods and ions with a narrow window of *m/z* will have stable trajectories through the quadrupole. In contrast, ions with *m/z* values outside of the *m/z* window will have an unstable path and will collide with the rods. By varying the voltages applied to the quadrupole, different *m/z* ions can be selected and reach the detector. A quadrupole ion trap (IT) is also composed of four linear rods in which the ions are trapped in the center of the device

and then scanned from the trap to a detector [276]. The sequential ejection of trapped ions using ramped voltages allows the recording of *m/z* ratio for a wide mass range. The IT is limited by the number of ions trapped in the device, which results in poor resolution and mass accuracy. Linear quadrupole ion traps have an increased trapping space and consist of two end cap electrodes [276]. The end cap electrodes have small holes in their centers that allows ions to either enter or exit the trap. Different voltages will trap and eject the ions equally through the sides of the trap and the *m/z* ratio will be recorded. In 1999, a new type of mass analyzer, the Orbitrap was invented by Makarov and was used for the first time in proteomics in 2005 by Hu *et al.*[277]. A Orbitrap consist of a central inner axial electrode, and outer electrode and two end cap electrodes [278]. Ions are trapped in an electric field, orbit around the central axial electrode and oscillate harmonically along its axis with a frequency characteristic for each *m/z* value [279]. This introduces an image current in the outer electrodes, which is transformed by Fourier-transformation into a frequency spectrum and then converted into a mass spectrum. The Orbitrap is coupled to a nitrogen-filled C-trap (curved RF-only quadrupole), where ions are trapped and pulsed into the Orbitrap by high voltages. The Orbitrap shows a very high mass accuracy and high resolution capabilities [273]. Different types of mass analyzers can be combined within an instrument, called hybrid mass spectrometer [274]. It combines the robustness and sensitivity of one mass analyzer with the high mass accuracy and resolution of another mass analyzer.

Tandem mass spectrometry (MS/MS) is a key technique to achieve the amino acid sequences from peptides (Figure 10). A MS1 spectrum shows intact peptide ions eluting at a given time with a specific *m/z* value and the heights of the signal reflects the numbers of detected ions [267]. For MS/MS, a precursor peptide with a specific *m/z* is selected and fragmented to generate product ions for detections. The precursor peptides break at the weakest bond, usually the peptide bond between amino acids forming b and y ions by colliding with an inert gas [271]. The resulting fragment ions are analyzed in a MS2 or MS/MS spectrum [267]. Several different fragmentation methods are utilized in mass spectrometry including collision induced dissociation (CID), higher energy collisional dissociation (HCD) and electron transfer dissociation (ETD) [280]. CID and HCD causes fragmentation primarily at amide bonds by collision of the precursor peptide with inert gas molecules. CID is performed in a linear ion trap filled with helium, whereas HCD takes place in a separate collision cell. The fragment ions produced by HCD are recorded by an Orbitrap. Due to slow-heating, internal fragmentation and neutral-losses of H_2O, NH_3 and labile PTMS are common [281]. ETD uses radical anions to rapidly transfer electrons with low electron affinity to multiply protonated peptides resulting in

a backbone fragmentation with c and z ions [282]. The electron transfer is highly efficient and fast and is therefore suited for peptides with higher charge state and labile PTMs like phosphorylation. Fragment ions containing the N-terminus of the peptide are referred as a,b,c, and the fragments containing C-terminus are described as x,y,z ions (Figure 10) [267,271].

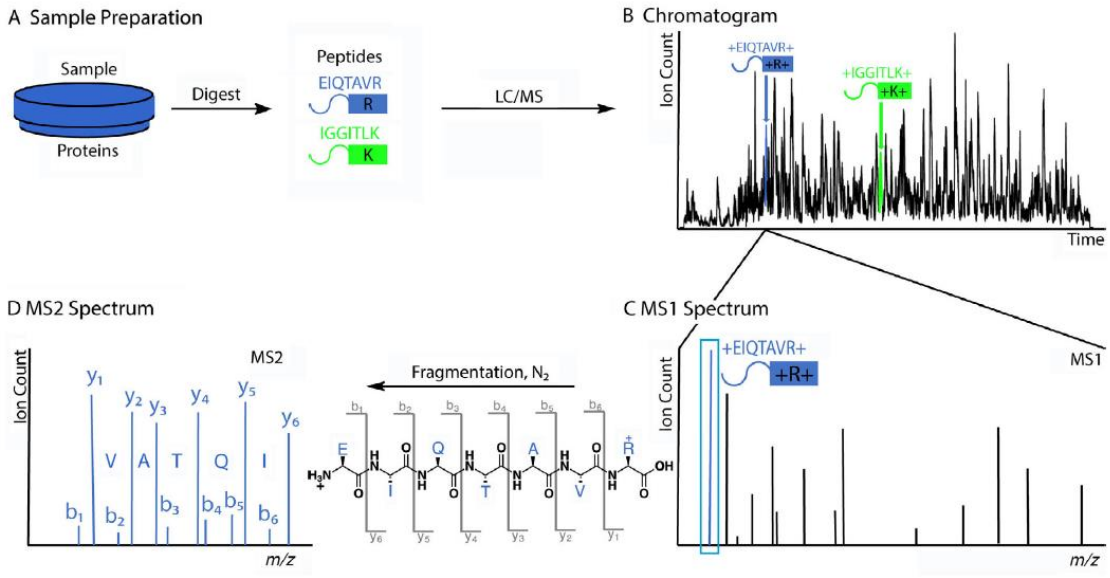


Figure 10: Peptide fragmentation with shotgun proteomics. Proteins extracted from different types of sample are digested into peptides using trypsin, which cleaves peptide bonds at the C terminus of lysine (K) and arginine (R) residues. To reduce complexity, the peptides are separated by liquid chromatography (LC), which shows the most abundant signal at each retention time and injected into the mass spectrometer. At any given time, multiple peptides co-elute at the same time and can be distinguish by their mass to charge ratio (*m/z*). The mass spectrum of the intact peptides is called the MS1 spectrum. The signal corresponding to the peptide EIQTAVR is highlighted in blue. The top intensive precursor ions are isolated and fragmented by collision with an inert gas. The *m/z* values of the fragment ions, derived from the blue peptide, are recorded in the MS2 spectrum. By convention, peptide fragments containing the N terminus are called b ions, whereas fragments with the C terminus are called y ions. (adapted from Pappireddi *et al.*, 2019 [267])

A mass spectrometer can be operated in different acquisition modes including data-dependent acquisition (DDA) and data-independent acquisition (DIA). In DDA, the most intensive precursor ions are selected for fragmentation and analyzed in the mass spectrometer [267]. The selected precursors are excluded for a certain time and the MS/MS spectra which are acquired will differ from run to run. However, dependent on the sample complexity, more signals are available than the instrument chooses for the isolation for fragmentation. The DIA approach could help to overcome this problem, MS2 spectra are continuously collected over the entire MS1 spectrum, resulting theoretically in a MS and MS/MS spectra for each *m/z* value. However, current instruments are not fast enough and isolation windows are selected to reduce the number of MS/MS spectra for each *m/z* value by still covering the entire *m/z* range. The

resulting MS/MS spectra typically contain fragment ions of multiple precursor ions and thus the data is more complex compared to DDA data. The obtained mass from the MS1 and the MS/MS spectra are matched against theoretical spectra generated from a predefined database [283] (Figure 11). The databases are specific for each organism and are usually protein databases translated from genomic data, although databases from spectral libraries or mRNA data were successfully applied. The databases are generated by *in silico* protein digestion for the specific protease using search engines, such as Mascot [283], Sequest [284] or Andromeda implemented into MaxQuant [285,286]. The best matches to a known amino acid sequence are scored and reported as a peptide identification. Each genetically encoded amino acid except for leucine and isoleucine has a different molecular weight, which allows the identification of the amino acid sequence by comparing the differences in the *m/z* values. A final step is the assembly of identified peptides into proteins. In order to control the extent of false and positive hits, a decoy database containing reverse peptide sequences together with the protein database is used [285,287] (Figure 11). Identifications generated from the decoy database can be considered as true negative hits and can be used to calculate the false discovery rate (FDR) [288]. The FDR threshold is typically 1% at the peptide and 5% at the protein level.

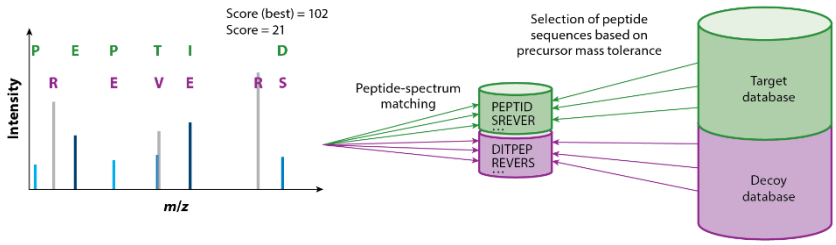


Figure 11: Data analysis of proteomic data and false discovery rate estimation. The data-dependent acquisition (DDA) data is analyzed by matching each acquired spectrum individually against a protein sequence database. A decoy database is used to estimate the false positive hits. Only the best scoring hit/spectrum is reported as peptide identification. (adapted from Gillet *et al.*, 2016 [288])

1.4.3 Discovery proteomics

Bottom-up or shotgun proteomics is the most widely used approach in MS-based proteomics to identify and characterize proteins across a broad dynamic range. This approach involves the proteolytic digestion of proteins prior analysis by liquid chromatography coupled to mass spectrometry (LC-MS) (Figure 12). In a typical workflow, proteins are extracted from different biological samples such as cultured cell or cancer tissues. Proteins are digested into peptides using specific proteases like trypsin, which cleaves the peptide bond at the carboxyl terminus of arginine (R) and lysine (K) [289]. The resulting oligopeptides are of optimal mass range for

chromatographic separation and due to the basic amino acid content these peptides can be ionized properly in electrospray ionization (ESI) [290]. For specific applications or to increase the sequence coverage, several other proteases like chymotrypsin, AspN or GluC can be used alone or in combinations [291]. Proteomic samples often contain a number of buffer components that may interfere with the MS/MS analysis, therefore peptides are usually concentrated and desalted on reverse-phase C18 columns. The human genome consists of around 20,000 proteins, resulting in approximately 106 million possible tryptic peptides [247]. To reduce the complexity of the analyte entering the mass spectrometer, high pressure liquid chromatography (HPLC) or ultra-high liquid chromatography (UHPLC) can be coupled to a mass spectrometer (LC-MS). Reverse phase (RP) chromatography is widely preferred for chromatographic separation [292,293]. The separation is based on the hydrophobicity of peptides determined by the amino acid composition. The peptides form hydrophobic interactions to alkyl chains coupled to silica beads and are eluted with increasing organic solvent over time. At the beginning of the gradient, hydrophilic peptides will elute, whereas hydrophobic peptides will stay longer on the column. The stepwise elution of peptides from the column allows also the detection of low abundant peptides and the separation of peptides with different hydrophobic properties. Eluting peptides are ionized by nano-ESI and thereby positively charged. The reproducibility of retention time between different runs is improved with a stabilized column temperature. The extensive co-elution of peptides is a still a major problem for on-line coupled LC-MS/MS analysis [294]. A pre-fractionation at the protein level or peptide level can be introduced to further reduce sample complexity and several techniques have been developed [295,296]. Proteins can be separated by 1D gel electrophoreses based on the molecular weight of the protein or different chromatography methods including off-line RP chromatography, size exclusion chromatography (SEC) or strong anion or cation exchange chromatography (SAX, SCX) [293,297]. For 1D gel electrophoreses, proteins are denatured in SDS and separated in a polyacrylamide matrix based on the molecular weight of the protein. Separated proteins can be digested with proteases directly in the gel and peptide extraction will be carried out with organic solvents such as acetonitrile prior LC-MS analysis [298,299]. Using SEC, proteins are separated according to their size and this technique is often used in top-down proteomics, the analysis of intact proteins by mass spectrometry. SAX and SCX are based on the interaction of positively (SCX) or negatively (SAX) charged peptides with opposite charged groups of the stationary phase [295,300]. The proteins are separated according to their electric charge using a salt or pH gradient. Most tryptic peptides carry two positive charges, due to C-terminal R or K residues and N-terminal

amines, which is preferred in proteomics analysis. In addition, proteins can be resolved into fractions using organelle separation to reduce sample complexity.

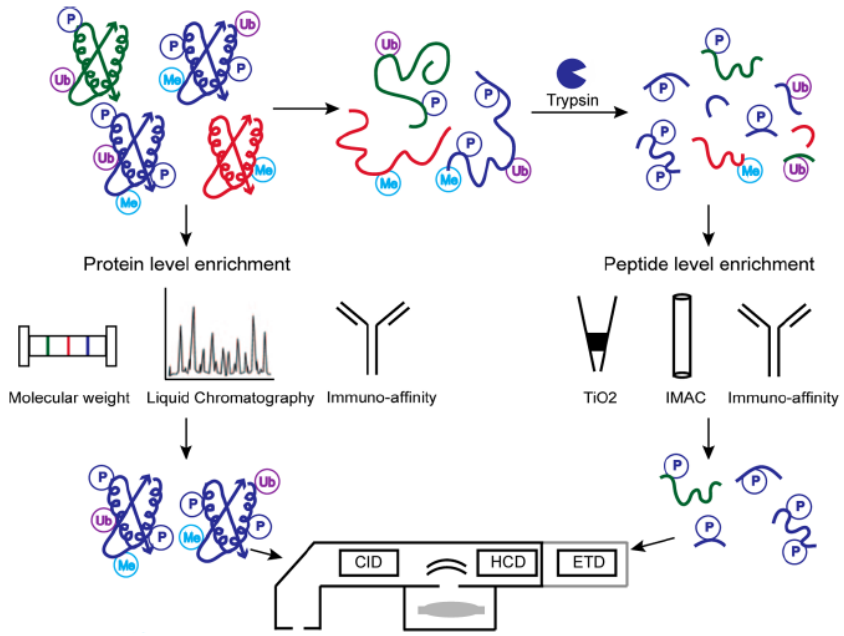


Figure 12: General workflow of mass spectrometry based proteomics. The proteins are extracted from cells or tissues and then enzymatically digested by proteases (trypsin) into peptides. The proteins can be multiply modified including phosphorylation (P), methylation (Me) or ubiquitylation (Ub). The complexity of the sample can be reduced at the protein and at the peptide level. Proteins can be fractionated by their molecular weight, liquid chromatography or immune affinity enrichment. Peptides can be pre-fractionated by cation exchange, reverse phase chromatography, and isoelectric focusing. Phosphorylated peptides, can be enriched by immobilized metal affinity chromatography (IMAC), titanium dioxide (TiO_2) or immune affinity enrichment. Each fraction is analyzed by liquid chromatography coupled to a mass spectrometer (LC-MS/MS). For fragmentation of precursor ion collision-induced dissociation (CID), higher-energy collisional dissociation (HCD) or electron-transfer dissociation (ETD) can be used. (adapted from Doll *et al.*, 2015 [301])

Proteins or peptides carrying a PTM are generally difficult to analyze by mass spectrometry and several strategies for the analysis of PTMs have been developed [297] (Figure 12). Phosphorylation introduces a negative charge to the peptides, whereas ESI is generally performed in the positive mode. Peptide separation and sample preparation prior MS analysis may be insufficient due to the hydrophilic properties of phosphopeptides. The low stoichiometry of phosphopeptides compared to non-phosphorylated peptides often results in ionic suppression and can be observed in low peaks [302]. To overcome these limitations several modification-specific enrichment techniques combined with advanced MS/MS methods were successfully applied [303,304]. The main goal of these approaches is the separation of modified peptides from unmodified proteolytic peptides due to the low stoichiometry of the modified peptides in a cell. The sensitivity of the technique is based on the yield of the enriched peptides,

extent of unmodified peptides contamination, complexity of the sample and the sensitivity of the MS/MS system [305]. Antibody-based enrichment of modified peptides is widely used method for a wide range of modifications including lysine acetylation, arginine methylation and tyrosine phosphorylation. It is based on the isolation of modified peptides by immunoaffinity purification using a PTM specific antibody [306]. In addition, to site-specific antibodies, antibodies recognizing a motif like kinase motifs are available. Ionic interaction-based enrichment strategies are mostly used for the enrichment of phosphopeptides from complex peptides mixtures. Ion metal affinity chromatography (IMAC) is based on the interaction of negatively charged phosphopeptides with positively charged ions (mainly Fe^{3+}) coupled to a matrix or beads. Peptides are eluted with increasing pH, which change the charge state of the peptide. SCX and SAX can be also applied to phosphopeptide enrichment. In the recent years, titanium dioxide (TiO_2)- based matrix enrichment has been prevalently used to enrich phosphopeptides [307,308]. The high affinity of TiO_2 to negatively charged phospho groups makes this method highly efficient and specific for the enrichment of phosphorylated peptides. To enrich for mono- and polyphosphorylated peptides, serval methods were combined in a wide range of studies and showed great results in sensitivity [293]. In addition, modified peptides can be also enriched by tagging PTMs by chemical derivation including *in vitro* chemical labelling and *in vivo* metabolic labelling. Besides the sample preparation also the data analysis including the precise localization of PTMs to a specific site can be challenging.

1.4.4 Quantitative proteomics

Quantification of thousands of proteins by mass spectrometry based proteomics is often used to study biological processes. Different techniques can be used to analyze protein-protein interactions, protein abundance or post-translational modifications between two or more physiological states [309-311]. These approaches can be mainly divided into relative quantification, the relative amount of a high number of proteins or absolute quantification to determine the absolute amount of distinct proteins [312]. For both, quantification is either performed without the introduction of stable isotopes (label-free) or with chemical or metabolic stable isotope labelling (label-based) (Figure 13). The most widely used form of quantitative proteomics is label-free and is based on the quantification of MS1 signals of peptides [313]. Two strategies can be applied: (1) spectral counting or (2) spectrometric signal intensity to measure the protein expression. In spectral counting, the number of measured spectra is determined as a proxy of the abundance peptide concentration in a sample [314]. A high sequence coverage enables the comparison over

multiple data samples. MS2 spectra are not used for the quantification, but required for the identification of the peptide sequence. Label-free quantification can be used for hundreds or thousands of samples and avoids additional costs and sample preparation for labelling or tagging of proteins and peptides. Each samples needs to be measured separately and the multiple runs results in a reduced throughput. Another limitation is the poor reproducibility between samples, which comes from each sample runs separately and variations in the MS1 and MS2 spectra acquisition. Another problem are missing values, as a significant fraction of peptides will not be detected in each samples. Imputation of missing values due to better data analysis technique help to overcome this problem [315]. The software imputes values by matching retention time and *m/z* values between samples. Improvement of mass spectrometry technologies may help to overcome these limitations. Label-based approaches allow the multiplexing of samples, which results in high accuracy and precision. Stable isotope labelling in cell culture present a pre-harvest labelling methods and is based on the incorporation of stable isotopes into proteins during their synthesis in active cells [316,317]. Metabolic labelling of proteins is often carried out by using arginine and lysine with stable isotopes of 2H, 13C and 15N. Methionine, histidine and leucine can also be used for SILAC as well [312]. In a typical SILAC experiment, cells are either grown in medium with unlabeled arginine/lysine (light) or in medium with labelled amino acids (heavy). After incorporation of these amino acids into newly synthesized proteins, cells are mixed, proteins are extracted and digested using trypsin protease prior LC-MS/MS analysis. The combination of samples in early step of the workflow minimizes variances between replicates and results in high reproducibility and accuracy compared to label-free approaches [318]. The tryptic digestion with trypsin results in peptides with C-terminal arginine and lysine and ensures a labelling of nearly all peptides. The heavy-labelled peptides produce a distinct mass shift in the MS1 spectra compared to its light-labelled counterpart. However, the chemical properties of the peptides are nearly identical results in the same chromatographically behavior and ionization efficiency [319]. The *m/z* difference between the heavy and light peptide will be resolved and the ratio can be calculated, which shows the relative changes in protein abundance between the different samples. SILAC has been successfully used for mammalian cell culture, yeast, mouse and bacteria [320-324].

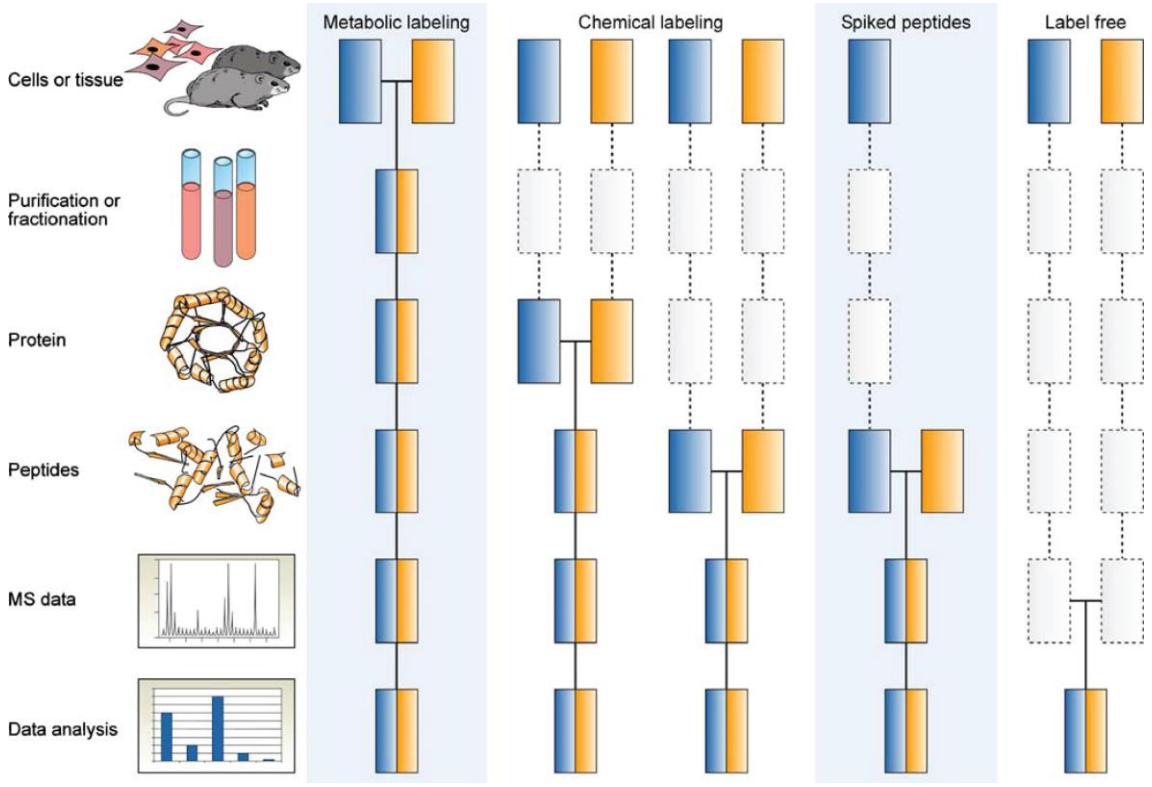


Figure 13: Overview of labelling techniques used in proteomics approaches. Quantification is either performed without the introduction of stable isotopes (label-free) or with chemical or metabolic stable isotope labelling (label-based) ate the protein or peptide level. Labels can be also introduced as spiked in as synthesized peptides. After labelling, samples can be combined for further sample preparation and processing. (adapted from Bantscheff *et al.*, 2012 [325])

Post-harvest labelling techniques like isotope-coded affinity tag (ICAT), isobaric tags for relative and absolute quantification (iTRAQ) or tandem mass tag (TMT) introduce labels at the protein or peptide level after cells have been harvested [326-328]. The mass of the isobaric tags is all the same and can be distinguished after fragmentation of the precursor ion. ICAT have been developed and often used for clinical samples such as tissues of cancer patients, which cannot be labelled with SILAC. The ICAT reagents consist of a reactive group (i.e. thiol), a linker group and an affinity handle like biotin [329]. The reactive group targets cysteine residues of peptides, which are labelled with linker consisting of light and heavy labelled stable isotopes. The affinity handle is used for the purification of tagged peptides and reduces the complexity of peptides mixtures from biological samples. One of the major limitations of ICAT is the labelling of only one amino acid, which will come in line with the loss of peptide analysis lacking cysteine. Therefore, the isobaric multiplexing tagging reagent iTRAQ has been developed and can be used to label up to four or eight different biological samples simultaneously [330]. The reactive group of iTRAQ targets primary amines at the N-terminus and side chain amino groups of peptides. iTRAQ can also be applied to absolute quantification using

synthetic peptides with a known sequence and mass. In recent years, isobaric labelling techniques were further developed and TMT allows multiplexing of up to eleven samples. Similar to iTRAQ, the tag contains a site that fragments in the MS2 spectrum and the mass of the reporter ion can be used to distinguish the peptide originated from different samples. Isobaric labelling techniques can be only applied to non-living samples and the labelling is applied at the peptide level, which does not control for variances introduced in the early step of the workflow. In addition, co-elution of multiple peptides is one of the major limitations as all peptides carrying the same reporter ion. Another chemical labelling technique at the peptide level is dimethyl-labelling [331]. Primary amines of lysine residues and N-termini of peptides are modified and the comparison is based on the mass difference of the label. Three different samples can be studied simultaneously. Quantitative proteomics is a powerful tool study dynamic changes of proteins and PTMs between different types of samples and allows insights into a variety of biological processes.

1.4.5 Proteogenomics in cancer biology

Proteogenomics, the combination of genomics and proteomics is a new research field and is often use to identify peptides containing mutations, novel protein-coding loci or alternative splicing forms [332-335]. The term proteogenomics was used for the first time in 2004 by Church group, performing a genomic reannotation of *mycoplasma pneumonia* using proteomic data [336]. Since then, the field of proteogenomics is emerging based on technical investigations enabling high throughput genomics and deep MS-based proteomics [337,338]. In addition to the data integration for genome reannotation, proteogenomics was widely used to study multilevel gene expression, signaling networks, disease subtypes and clinical prediction of patient's outcome [339,340]. Most of the proteogenomic workflows involve several main steps (Figure 14). First, nucleotide data can be used to encode the sequence of the proteins expressed in a sample and this data can be retrieved from the genome, exome, transcriptome or translatome level [332]. The genome contains sequence information of protein-coding and non-coding regions but also the backbone of all protein sequences. Exome sequencing of enriched exonic sequences through hybridization capture covers only 1% of the genome that codes for proteins [341]. Transcriptome represents the output of the gene transcription and translatome comprises the portions of transcriptome bound to the ribosome. These different types of sequence data can be used to generate a proteogenomic database, which ultimately include possible single amino acid variants (SAVs), insertions and deletions (InDels) and splice junctions for each individual

sample [332]. The NGS data needs be analyzed before by either *de novo* assembly or alignment of genomic sequences to a pre-existing database- such as Ensembl [342], RefSeq [343] or UniProt [344]. *De novo* assembly is used to build a reference genome from one or multiple sequences of these genomes [345]. The newly-synthesized and assembled genomes, the protein-coding regions and other functional elements within an assembled genome must be identified [346]. A majority of proteogenomic databases were created by six frame translation of the genome of an organism. This includes the prediction of all possible protein-coding sequences expressed in a sample under analysis. For the alignment (reads mapping) the re-sequenced organism is mapped to a pre-existing reference genome and *in silico* translated variants are incorporated into a protein database [347]. This enables the identification of a variety of genomic variants, which are absent in the reference genome. These sample-specific variants are of fundamental biological interest including novel or unannotated proteins, variants specific for an individual, mutations underlying a disease. Similar to proteomics, MS/MS spectra are matched to peptide sequences, known as peptide spectrum matches (PSM) [346,348]. Peptide sequences obtained from PSM are then matched against a customized protein sequence database. By incorporating sample specific sequences like SAVs, InDels, alternative splice junctions and novel gene fusions the size of the database increases. However, the likelihood of experimentally observing a peptide decreases and the matching of more theoretical candidates against experimental MS/MS spectrum will be likelihood incorrect [349]. To identify high confidence PSMs, a quality spectra and fragmentation will be required. In addition, the computational time increases with larger databases and the FDR must be controlled at the protein and peptide level. Separate FDR estimation should be applied to known and novel peptides [350]. For proteogenomic approaches it will always be a compromise between completeness of the database, size and a well-controlled FDR.

A number of bioinformatical tools have been developed for proteogenomic approaches. For instance, costumProDB [351], Galaxy-P [352], PPline [353] or PGA [354] software can be used to generate customized protein databases from NGS data, including SNVs, InDels ad novel splice junctions [355]. In addition, also specific software are available with a special focus for example on novel splice isoforms (QUILTS [356] or SpliceDB [357] software). Not only sample-specific NGS data can be incorporated, several public SNV repositories like COSMIC [53] or dbSNP [358] database can be used. Besides the database construction tools, visualization of peptides on the genome and proteome level have been developed. However, there is a processing need for efficient and easy-to use tool for the bioinformatic analysis and visualization of proteogenomic data [332].

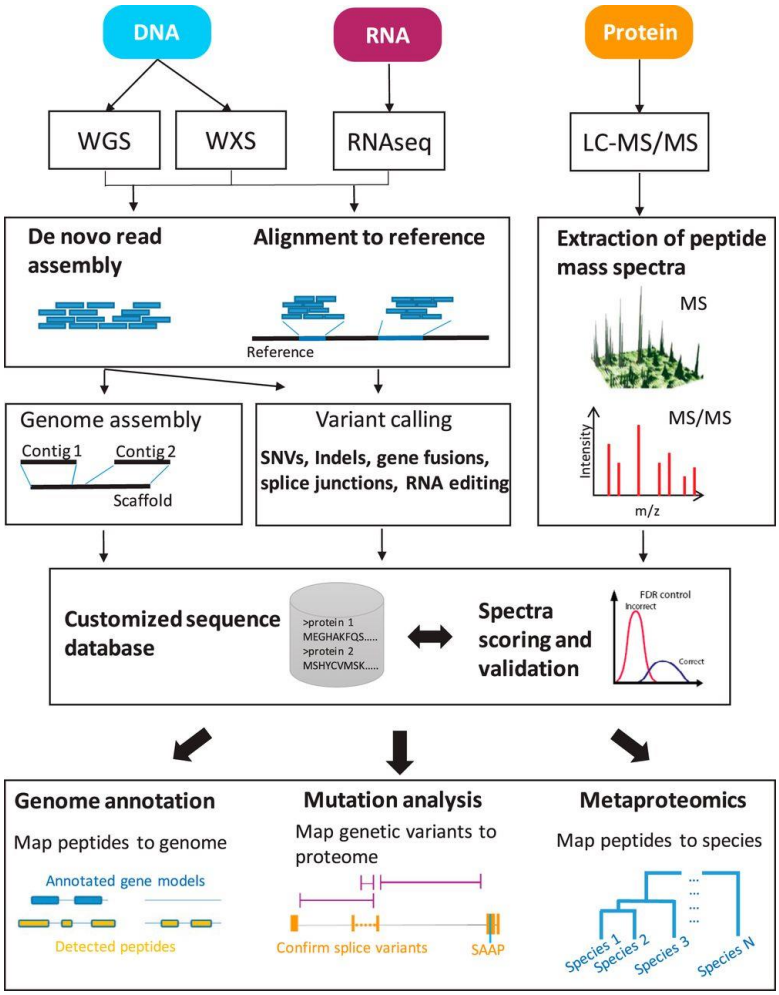


Figure 14: Sequence-centric proteogenomics workflow. Nucleotide data that can be retrieved from the genome (whole genome sequencing (WGS)), exome (whole exome sequencing (WXS)), transcriptome (RNAseq) level by next generation sequencing (NGS). The sequences need to be assembled into the DNA or RNA sequence by either *de novo* assembly or alignment of genomic sequences to a pre-existing reference database. Sample specific alterations are determined and incorporated into customized sequence databases. Peptide identifications from corresponding LC-MS/MS runs of the same sample are used to identify sample specific sequences and alterations by matching the acquired MS/MS spectra against the spectra derived from the customized database. Proteogenomic approaches can be used for genome annotation by mapping of peptides to unannotated genome regions; to identify tumor-specific variants on the proteome level as well as novel protein splice variants; and detect species-specific peptides in microbial communities. (adapted from Ruggles *et al.*, 2015 [355])

Cancer and especially melanoma is characterized by an accumulation of mutations such as SNVs, InDels, frameshifts or copy number variants (CNVs) rewiring cellular networks [332,340]. The identifications of cancer specific mutations in cancer proteomics is referred as onco-proteogenomics and has been studied intensively in the last decade [359]. A proteogenomics approach and a cancer-specific database allows the detection of variants on the proteome level, which has the potential to yield novel insights into cancer biology [355]. The proteomic detection of mutated proteins can help to identify clinical biomarkers or actionable drug targets [357,359,360]. Two consortium namely The Clinical Tumor Analysis Consortium (CPTAC) and the Cancer

Genome Atlas (TCGA) were launched simultaneously and showed great contribution and effort for the field of onco-proteogenomics. Colorectal, breast and ovarian cancer types has been studied by CPTAC so far and several proteogenomic studies described the classification of cancer subtypes, correlation of copy number alterations with protein expression and the identification of mutated proteins and PTMs and associated signaling pathways [360-362]. The TCGA created a large database containing novel, cancer-specific peptides including splice junctions, InDels and substitutions. Using this TCGA database, 524 novel variant peptides in a single ovarian cancer sample were identified in 2014 by Woo *et al.* [363]. In a large scale study of 49 NCI60 cancer cell lines, 7.3 million novel peptides and 4,771 mutations were reported by combining dbSNP, COSMIC, UniprotKB and sample-specific genomic and transcriptomic data [364]. In addition to the identification of novel variant peptides, Sun and colleagues identified several unique fusion peptides in non-small cell lung cancer (NSCLC) [365]. These studies provide valuable insights into cancer biology and how the proteome is regulated by genetic effects. The repertoire of reported somatic mutations and the cellular responses within a cell will help to identify driver and cancer-specific mutations and disturbed signaling networks in precision oncology. The altered signaling networks in cancer cells can be studied by an integrative proteogenomics analysis of genetic variants and PTMs, which can identify consequences of genetic variants on the molecular level [366]. It is very likelihood that a change of an amino acid residue in a PTM target directly rewiring cellular networks. Approximately 22.2% of all amino acids in the human proteome are serine (S), threonine (T), tyrosine (Y) and lysine (K) [344]. These four amino acids are predominantly modified by phosphorylation at S/T/Y or acetylation and ubiquitylation at K. Several studies have been reported that these four amino acids are disproportionally affected by missense mutations [366-368]. In addition, not only the amino acid itself can be affected also the recognition motif in a flanking region of the mutation site for corresponding transferase like protein kinases or ligases might be altered [369,370]. Yang *et al.* showed that 64 phosphorylation sites potentially change the phenotype compared to nonphosphorylated amino acid by using SwissProt and dbSNPs annotated in the NCBI sbSNP database [371]. In a CPTAC breast cancer study, Mertins and colleagues performed at integrative analysis of proteomic and phosphoproteomic study of 105 genomically annotated samples [362]. This allowed the classification of the breast cancer subtypes on the genome and proteome level and the characterization of the somatic cancer genome including the chromosomal loss of 5q. In addition, the phosphoproteomic dataset revealed a new G-protein receptor cluster and activated kinases specific for breast cancer subtypes, which was not identified by mRNA analysis alone. As for the breast cancer study, Zhang *et al.* identified specific acetylation sites

that could be associated with homologous recombination in ovarian cancer [372]. These newly identified modification sites may help to find the best treatment strategy for each individual patient. Proteogenomics is often used to classify cancer subtypes and to identify actionable mutated proteins for the best course of treatment for each individual. However, onco-proteogenomics was also recently applied to study acquired resistance to treatments and drug toxicity. In a lung cancer study, several mutations were reported to be linked to the efficiency of tyrosine kinase inhibitors and showed a potential to predict patient's survival rates [373]. In addition, acquired resistance and driver mechanisms of resistance against tyrosine kinase inhibitors were also studied in gastrointestinal stromal tumors using a proteogenomic phosphoproteomic approach [374].

In summary, the application of proteogenomics in clinical research is a powerful tool study the mode of action of disease-associated mutations on the genome, proteome and PTM level. The mutation can directly or indirectly change PTM sites and several studied showed that this may change the protein stability, protein-protein interactions and signal networks. It also highlights the value of state of the art MS-based proteomics in the era of precision medicine. In addition, proteogenomics will may improve the treatment of patients in personalized manner and may be implemented as a routine clinical lab test in the future.

2 Aims and objectives

Malignant melanoma is characterized by a high frequency of somatic mutations in key signaling pathways and development of resistance mechanisms against treatments. The overreaching aim of this thesis was to characterize proteomic changes and the mutational landscape of melanoma cells in response to kinase inhibitors and immunotherapy. To achieve this, I first compared drug-sensitive and drug-resistant melanoma cells in order to gain new insights into the general molecular mechanisms underlying resistance to kinase inhibition. I next applied a proteogenomic approach to drug-sensitive and -resistant melanoma cell lines in response to kinase inhibitors to assess the influence of non-synonymous mutations on signal transduction networks. Finally, I applied a proteogenomics approach to melanoma patient tumor material to investigate the molecular mechanisms underlying immunotherapy response in individual patient samples.

Specific aims and objectives were:

1. Comprehensive (phospho)proteomics analysis of vemurafenib-sensitive and resistant melanoma cells
   a. Comparison of drug-sensitive and drug-resistant A375 melanoma cells using quantitative (phospho)proteomics
   b. Identification of significantly regulated proteins and phosphosites between both phenotypes
   c. CRISPR/Cas9-mediated knockout of the interesting candidates identified at the proteome level
   d. Functional validation of candidates using cell proliferation, migration and invasion assays, as well as proteomic analyses

2. Individualized proteogenomic characterization of melanoma cell lines A375 and SkMel28 in response to vemurafenib resistance
   a. Identification of mutations at the genome and (phospho)proteome level
   b. Investigation of over-represented pathways due to the accumulation of alternate proteins between phenotypes and cell lines
   c. Identification of phosphoproteins harboring mutations with a potential to rewire signal transduction
   d. Define significantly regulated proteins and phosphosites between phenotypes

e. Validation of mutated protein isoforms by interactome studies

3. Individualized proteogenomic characterization of human melanoma cells in response to immunotherapy
   a. Optimization of the protein extraction protocols for patient tissue samples
   b. Identification of patient-specific variants at the genome, proteome and phosphoproteome level in the context of immunotherapy
   c. Investigation of pathway over-representation due to accumulation of mutated proteins between naïve and immune checkpoint inhibitor treated samples
   d. Comparison of proteogenomic profiles of human patient material and patient-derived xenografts generated in NSC mouse
   e. Prioritization of patient-specific mutations with a high impact to rewire signal transduction
   f. Validation of mutated protein isoforms by interactome studies

3 Results

3.1 Manuscript I

Schmitt, M., Sinnberg, T., Nalpas, N. C., 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 18, 1096-1109

doi.org/10.1074/mcp.RA119.001302

Author contributions:

The study was designed by Tobias Sinnberg, Birgit Schittek, Boris Macek and me and all experiments were performed at the Proteome Center Tuebingen, University of Tuebingen. Cell culture, sample preparation, quantitative proteomics and phosphoproteomics analysis, biological and biochemical assays were performed by me under the supervision of Boris Macek. Tobias Sinnberg performed immunohistochemistry staining of clinical FFPE samples. Nicolas C. Nalpas performed analysis of microarray data from pre and post treated patient tumors. Annika Maass generated the CRISPR/Cas9-mediated knockout of nestin in melanoma cells under my supervisor. The data was analyzed and interpreted by Tobias Sinnberg, Nicolas C. Nalpas and me under the supervision of Boris Macek. Birgit Schittek provided the drug-sensitive and –resistant melanoma cells and reagents and contributed in the discussion of the manuscript.

A list of all used materials and extended method description can be found as Supplementary Information on the provided CD.

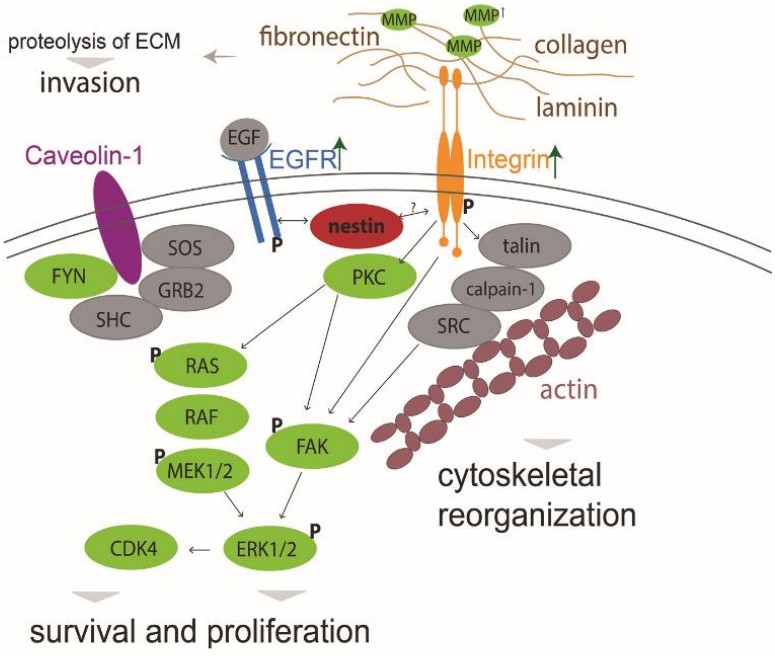
Quantitative Proteomics Links the Intermediate Filament Nestin to Resistance to Targeted BRAF Inhibition in Melanoma Cells

Authors
Marisa Schmitt, Tobias Sinnberg, Nicolas C. Nalpas, Annika Maass, Birgit Schittek, and Boris Macek

Correspondence
boris.macek@uni-tuebingen.de

In Brief
The proteome and phosphoproteome have been determined for BRAF inhibitor resistant and sensitive cells by applying quantitative MS-based proteomics. We identified the intermediate filament protein nestin as one of the highest downregulated proteins in melanoma cells and tumors. The results reported reveal a link between loss of nestin and an invasive phenotype and provide a quantitative view of PI3K/Akt and integrin pathways involved in resistant and nestin knockout cells.

Graphical Abstract



Highlights

- Quantitative phosphoproteome of BRAF drug-resistance in melanoma cells.
- Cytoskeletal proteins are downregulated in resistant vs. sensitive cells.
- Nestin is associated with an invasive phenotype and resistance to MEK and BRAF inhibitors.
- Nestin depletion affects PI3K/AKT and integrin signaling similar to resistant cells.

Schmitt et al., 2019, Molecular & Cellular Proteomics 18, 1096–1109
June 2019
https://doi.org/10.1074/mcp.RA119.001302

Research

Quantitative Proteomics Links the Intermediate Filament Nestin to Resistance to Targeted BRAF Inhibition in Melanoma Cells*

Marisa Schmitt‡, Tobias Sinnberg§, Nicolas C. Nalpas‡, Annika Maass‡, Birgit Schittek§, and Boris Macek‡¶

Targeted inhibition of mutated kinases using selective MAP kinase inhibitors in malignant melanoma often results in temporary improvement of clinical symptoms followed by rapid development of resistance. To gain insights in molecular processes that govern resistance, we performed SILAC-based quantitative proteomics profiling of vemurafenib-resistant and -sensitive melanoma cells. Among downregulated proteins in vemurafenib-resistant cell lines we detected multiple proteins involved in cytoskeletal organization and signaling, including the intermediate filament nestin, which was one of the most downregulated proteins. Previous studies showed that nestin is expressed in various types of solid tumors and its abundance correlates with malignant phenotype of transformed cells. However, the role of nestin in cancer cells regarding acquired resistance is still poorly understood. We performed CRISPR/Cas9 knockout of the nestin gene (NES) in vemurafenib-sensitive cells and showed that loss of nestin leads to increased cellular proliferation and colony formation upon treatment with BRAF ^{V600E} and MEK inhibitors. Moreover, nestin depletion led to increased invasiveness and metalloproteinase activity like the phenotype of melanoma cells with acquired resistance to the BRAF inhibitor. Finally, phosphoproteome analysis revealed that nestin depletion influenced signaling through integrin and PI3K/AKT/mTOR pathways and led to increased focal adhesion kinase abundance and phosphorylation. Taken together, our results reveal that nestin is associated with acquired vemurafenib resistance in melanoma cells. Molecular & Cellular Proteomics 18: 1096–1109, 2019. DOI: 10.1074/mcp.RA119.001302.

Malignant melanoma is the 19 th most common cancer worldwide, accounting for ~232,000 new cases in 2012 (1). Although melanoma accounts for less than one percent of skin cancer cases, it is responsible for 79% of skin cancer-related deaths (2). Therapies for advanced melanoma have changed greatly in recent years with the US Food and Drug Administration approval of several immunotherapy and targeted drugs, including cobimetinib, an inhibitor of MEK kinases (3), and vemurafenib, a BRAF inhibitor, for patients carrying the BRAF ^{V600E} mutation (4). These drugs have a response rate of ~50% as monotherapies and result in an average survival benefit offour to six months (5–7). However, almost all patients ultimately develop resistance to drug treatment (5, 8). Therefore, an understanding of the acquired resistance is essential for the development of effective therapies for malignant melanoma. Multiple cellular pathways have been postulated to explain drug resistance, ranging from signal transduction networks to transcriptional pathways (9). The majority of kinase inhibitor resistance development is caused by molecular or genetic alterations that lead to MAPK pathway reactivation. A variety of genetic causes, including NRAS (10, 11), KRAS (12) and MEK mutations (13), alternative splicing or amplification of BRAF and loss of NF1 (9, 11, 12), have been identified in tumors with acquired resistance. In addition, activation of the PI3K/AKT pathway can be responsible for BRAF inhibitor resistance—because of downregulation of PTEN through loss or mutational inactivation, or somatic mutations in AKT1/3 and PIK3CA (11, 12).

Nestin, a member of the type VI intermediate filament protein family, was originally described as a stem cell/progenitor cell marker, especially during migration and proliferation phases in early embryonic development (14). Expression of nestin is also associated with the regulation of cell death in neural progenitor cells, podocytes of kidneys and neuromuscular junction development in a CDK5-dependent manner (15). In adult tissue, it plays an important role in regeneration processes where it is localized to tissue/organ-specific sites (16). Previous studies have reported that nestin is expressed in various human malignancies, including pancreatic cancer (17, 18), prostate cancer (19), breast cancer (20), glioblastomas (21), gastrointestinal stromal tumors (22), trichoblastoma (23), angiosarcoma (22) and malignant melanoma (24, 25). In some tumor types, nestin expression correlates with aggres-

From the ‡Proteome Center Tuebingen, Interfaculty Institute for Cell Biology, University of Tuebingen, Tuebingen, Germany; §Center for Dermatooncology, Department of Dermatology, University of Tuebingen, Tuebingen, Germany

Received January 7, 2019, and in revised form, February 15, 2019
Published, MCP Papers in Press, March 19, 2019, DOI 10.1074/mcp.RA119.001302

sive growth, metastasis, migration and poor prognosis (18); however, the roles of nestin in cancer cells have not been characterized at a molecular level. In advanced stages of melanoma, nestin- and CD133-positive melanoma cells were detected in the peripheral blood of patients, at the invading front and at sites of melanoma metastases (26–28). These studies indicate that nestin could be significantly involved in the invasion and distant metastasis of melanomas. In a large-scale proteomic approach, nestin was found to be a useful diagnostic and prognostic biomarker that can potentially distinguish melanoma subtypes and can help to predict melanoma aggressiveness in these different subgroups (29). Interestingly, depletion of nestin in melanoma was shown to increase expression of matrix metalloproteinases (MMP) [1] and enhance melanoma invasion (30). Recent evidence indicates that nestin downregulation in prostate cancer cell lines triggers an expression pattern of phosphorylated focal adhesion kinase (FAK). Phosphorylated FAK (pFAK) localizes at the cell membrane and promotes integrin clustering. This results in pFAK- and integrin-dependent matrix degradation and an invasive phenotype (31). In the context of targeted BRAF ^{V600E} and MEK inhibitor therapy in melanoma, a loss of nestin expression in tumor cells was identified immediately after treatment therapy (32). All these findings suggest that nestin is associated with tumorigenesis, however, little is known about the role of nestin in melanoma and the relationship of nestin and acquired resistance.

In this study, we use quantitative proteomics to identify proteome and phosphoproteome alterations in A375 melanoma cells resistant to BRAF ^{V600E} inhibitor vemurafenib. Our analysis identified nestin as one of the most downregulated proteins in resistant cells. Extensive biological follow-up revealed its connection with invasiveness and cell survival of resistant melanoma cells. Finally, phosphoproteome analysis revealed that nestin depletion influenced signaling through integrin and PI3K/AKT/mTOR pathways.

EXPERIMENTAL PROCEDURES

Experimental Design and Statistical Rationale— The (phospho)proteomics data is derived from three sets of samples prepared and analyzed by LC-MS/MS. A total of 143 runs analyses were performed with 230 min gradient for proteome, 42 min gradient for fractionated proteome and 90 min gradient for phosphoproteome measurements on a Q Exactive HF mass spectrometer. In part 1, SILAC labeled A375 S and A375 R cells ("light," "medium," and vice versa) were used in two different screens (123 samples); screen 1, proteome and phosphoproteome measurements (33 samples, three biological replicates (11 per replicate), ten rounds of phosphopeptide enrichment for each replicate), whereas in screen 2, the proteome was functioned into 30 fractions (90 samples, three biological replicates (30 per replicate)). In part 2, SILAC labeled Nes-KO, A375 S and A375 R cells were used ("light," "medium," "heavy") (22 samples, two biological replicates (eleven per replicate), ten rounds of phosphopeptide enrichment per replicate). Raw data was processed by MaxQuant software as described below. Statistical analysis was performed with Perseus (t test, FDR < 0.01, s0 $= 1$) (version 1.5.0.31), STRING: functional protein association networks analysis (STRING Consortium 2018) and GraphPad Prism (version 7.04). For detailed description of statistical analysis of each experiment see MS data analysis and statistical analysis in the section methods.

All biological assays were performed in three biological and technical replicates, so that appropriate statistical analysis could be performed. Statistical analysis was performed with two-tailed unpaired t test in GraphPad Prism. p values < 0.05 were considered statistically significant, with * for $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$ and **** for $p < 0.0001$. In each experiment, separate controls were included. Images of experiments and Western blotting were quantified using ImageJ software.

Chemicals— Stock solutions of the BRAF ^{V600E} inhibitor PLX4720 (vemurafenib analog, Selleckchem, Houston, TX) and the MEK inhibitor cobimetinib (Hycultec, Beutelsbach, Germany) were prepared in dimethyl sulfoxide (DMSO).

Cell Culture and SILAC Labeling— The use of human tissue from an internal biobank was approved by the local ethical committee (781/2018BO2) and experiments were performed in accordance with the declaration of Helsinki Principles. The human metastatic BRAF ^{V600E}-mutated melanoma cell lines A375, Mel1617, 451lu, SKMel28 and SKMel19 were used in this study (33, 34). The generation of the cell lines with acquired resistance to vemurafenib analogue PLX4720 (for simplicity referred to as "vemurafenib" in the Results section) was conducted as described previously (33) (List of used cell lines in supplementary Information S1). Cells were either grown in RPMI 1640 (Sigma-Aldrich, Darmstadt, Germany) containing penicillin/streptomycin (100 U/ml, PAN) and FBS (10%, PAN) or RPMI 1640 SILAC (Sigma-Aldrich) lacking arginine, lysine and leucine. Leucine (12.5 mg/ml, Sigma-Aldrich), penicillin/streptomycin (100 U/ml, PAN, Aidenbach, Germany), dialyzed FBS (10%, PAN) and stable isotope-encoded arginine and lysine were added to the SILAC medium. The "light" SILAC media was further supplemented with L-[$^{12}C_6$,$^{14}N_2$] lysine (Lys0) and L-[$^{12}C_6$,$^{14}N_4$] arginine (Arg0) (Cambridge Isotope Laboratories, Tewksbury, MA), whereas L-[$^{2}H_4$] lysine (Lys4) and L-[$^{13}C_6$] arginine (Arg6) were added to the "medium" SILAC media and L-[$^{13}C_6$,$^{15}N_2$] lysine (Lys8) and L-[$^{13}C_6$,$^{15}N_4$] arginine (Arg10) to "heavy" SILAC media. Cells were grown in a humidified atmosphere, 5% CO_2 at 37 °C in either RPMI 1640 or "light," "medium," or "heavy" RPMI 1640 SILAC media.

Protein Extraction of Cultured Cells— Cell lysis was performed in lysis buffer (6 M urea, 2 M thiourea, 10 m M Tris pH 8.0) supplemented with protease inhibitor (complete Mini EDTA-free tablets, Roche, Basel, Switzerland), phosphatase inhibitor buffers (5 m M glycerol-2-phosphate, 5 m M sodium fluoride, and 1 m M sodium orthovanadate) and 1% N-Octylglucoside (NOG) on ice for 10 min. DNA and RNA in the lysate was removed using benzonase (Merck, Darmstadt, Germany) for 10 min on room temperature (RT) followed by centrifugation at 2800 × g (10 °C, 20 min). Remaining NOG detergent was removed by acetone precipitation. Briefly, four volumes of acetone (−20 °C), one volume of methanol were added, and the proteins were precipitated overnight at −20 °C. After centrifugation (2800 × g, 4 °C,

[1] The abbreviations used are: MMP, matrix metalloproteinase; BRAFi, BRAF inhibitor; CRISPR, clustered regularly interspaced short palindromic repeats; DSB, double strand break; DTT, dithiothreitol; ECM, extracellular matrix; FDR, false discovery rate; FBS, fetal bovine serum; FFPE, formaldehyde fixed paraffin embedded; IAA, iodoacetamide; LC-MS/MS, liquid chromatography combined with tandem mass spectrometry; NOG, N-ocetylglucosamine; NonSil, nonsilencing siRNA; PAM, protospacer adjacent motif; SDS, sodium deoxycholate; SILAC, stable isotope labeling in cell culture; sgRNA, single guided RNA; NonTar, nontargeting control RNA; siRNA, small interfering RNA.

20 min), the detergent-containing supernatant was removed, and the protein pellet was washed with 80% acetone (−20 °C). Protein pellets were then resolved in lysis buffer without NOG and protein concentration was measured using Bradford assay.

Sample Preparation for MS Analysis— Extracted proteins from each cell line were mixed 1:1 ("light" to "medium") or 1:1:1 ("light" to "medium" to "heavy") and reduced with 10 mM DTT for 1 h, alkylated with 55 mM iodoacetamide for an additional hour and digested with Lys-C (Lysyl Endopeptidase, Wako Chemicals, Neuss, Germany) for 3 h at RT. After adding offour volumes of 50 mM of ammonium bicarbonate, trypsin (Promega Corporation, Madison, Wisconsin) was added and tryptic digestion was carried out overnight. To stop the digestion, 1% TFA was added, peptides were purified on Sep-Pak C18 Cartridge (Waters, Milford, MA) and either eluted in 80% acetonitrile (ACN) for high pH reverse phase chromatography or desalted on C18 StageTips (as described previously (36)).

High pH Reverse Phase Chromatography— 1–2 mg of peptides were fractionated on an off-line Ultimate 3000 high-pressure liquid chromatography (HPLC) system (Dionex, Thermo Fischer Scientific, Waltham, MA) equipped with xBridge BEH130 C_{18} 130A, 3.5 μm, 4.6 × 250 mm column (Waters), as described previously (37). The system was operated under high pH conditions using buffer A (5 mM NH_4OH) and buffer B (5 mM NH_4OH in 90% ACN) at pH 10. Peptides were eluted using an 80 min gradient at a flow rate of 1 ml/min. The gradient consisted of 5% to 25% B over 45 min, followed by 40% B during 10 min and finally 70% B for 5 min. The gradient was held at 70% B for 5 min, reduced to 5% B within 5 min and the column equilibrated for 10 min. One-minute fractions were collected from 0 to 60 min. The 60 fractions were concatenated into 30 pools and dried by vacuum centrifugation. Peptide pools were reconstituted in 1 ml of 80% ACN, 10 μg of the pool were concentrated and desalted on StageTips prior LC-MS/MS measurements for proteome analysis.

Phosphopeptide Enrichment— Phosphopeptides were enriched using TiO_2 beads. TiO_2 beads (Titansphere, 10 μm, GL Sciences, Tokyo, Japan) were resuspended in DHB solution (80% ACN, 1% TFA, 3% 2,5-dihydroxybenzoic acid (DHB)) and incubated for 20 min. Purified peptides were added to the TiO_2 beads in DHB solution and incubated for 10 min for each enrichment round. This step was repeated nine to ten times. Phosphopeptide-bound TiO_2 beads were sequentially washed with 30% ACN in 1% TFA, 50% ACN in 1% TFA followed by 80% ACN in 1% TFA. Peptides were eluted with 5% NH_4OH in 60% ACN 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.

Formalin Fixed Paraffin Embedded Tissue Preparation for MS Analysis— FFPE tissue of pre-and posttreated patients with the kinase inhibitor vemurafenib (2 serial sections, 5 μM thick; List of FFPE specimens used in this study in supplementary Information S3) were first deparaffinized by two washes in xylene (5 min, 50 °C) followed by three serial washes in ethanol (100%, 95% to 70%) for 10 min each. Ethanol was removed completely, and sections air-dried. Lysis was carried out in 4% SDS, 50 mM DTT, 100 mM HEPES pH 7.5 supplemented with protease inhibitor at 95 °C for 60 min and by sonication for 15 min. 10 μg of proteins were purified by acetone precipitation and protein pellet was resolved in lysis buffer (6 M urea, 2 M thiourea, 10 mM Tris pH 8.0). Proteins in the cleared lysate (13,000 × g, 10 min) were reduced with 10 mM DTT for 60 min, alkylated with 55 mM iodoacetamide for an additional 60 min and LysC digestion was carried with 1 μg of LysC fo r 3 h at RT. After adding four volumes of 50 mM ammonium bicarbonate, 1 μg of trypsin was added for tryptic digestion overnight. Digestion was stopped by adding 1% TFA and peptides were loaded onto C18 StageTips for desalting and subsequent dimethylation labeling (35). Briefly, peptides were labeled with either 4% CH_2O, 0.6 M $NaBH_3CN$ ("light") or 4% $^{13}CD_2O$, 0.6 M $NaBD_3CN$ ("heavy") in phosphate buffer (50 mM NaH_2PO_4, 50 mM Na_2HPO_4). After washing with solvent A, peptides were eluted with solvent B and analyzed by LC-MS/MS.

Liquid Chromatography-MS Analysis— All samples were analyzed on an Easy-nLC 1200 UHPLC (Thermo Fischer Scientific) coupled to an Q Exactive HF mass spectrometer (Thermo Fischer Scientific) equipped with a nanoelectrospray source. Peptides were separated on a 20 cm analytical column (75 μm ID PicoTip fused silica emitter (New Objective, Woburn, MA)) in-house packed with ReproSil-Pur C18-AQ 1.9 μm resin (Dr Maisch GmbH (Ltd.), Ammerbuch, Germany). Peptides were eluted using a 90 min gradient for phosphoproteome, 230 min gradient for proteome analysis and 42 min gradients for fractionated peptide pools. Gradient was generated by solvent A (0.1% FA) and solvent B (80% ACN in 0.1% acetic acid) at 40 °C and 200 nl/min flow rate. The mass spectrometer was operated in data dependent mode. Full MS scans were acquired with a resolution of 120,000 and within a mass range of m/z 300 to 1650. For higher-energy collisional dissociation (HCD), the 12 most intensive peptides were selected, and MS/MS spectra were recorded with a resolution of 60,000. For 45 min gradients, fast scanning top20 method was used with a resolution of 15,000 for HCD scans and maximum fill time of 30 ms. For the analysis of TiO_2 enriched phosphopeptides, full MS were acquired in the range of 300–1500 m/z at a resolution of 120,000. Seven most abundant precursor ions from a survey scan were selected for HCD fragmentation (isolation width of 1.20 m/z; 27% normalized collision energy and activation time of 0.1 ms were allowed) and MS/MS spectra were acquired at a resolution of 60,000 on the Orbitrap analyzer.

MS Data Analysis and Statistical Analysis— The raw data files were processed with the MaxQuant software suite (version 1.5.2.8) (38). The Andromeda search engine (39) searched MS/MS data against Uniprot Homo sapiens (release 2015_10_23; 91,646 entries) database containing commonly observed contaminants. Carbamidomethylation of cysteine (C) was set as fixed modification and oxidation of methionine, phosphorylation at serine (S), threonine (T) or tyrosine (Y) were defined as variable modifications. Trypsin/P was selected as a protease. For quantification, the amino acids (Lys4)/(Arg6) and (Lys8)/(Arg10) were defined as "medium" and "heavy" labels for the comparison of cell lines. Dimethylation on peptide N termini and lysine residues was defined as "light" (+28.03 Da) and "heavy" (+36.08 Da) labels for the comparison of pre-and posttreated tumors (FFPE specimens). 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 all other parameters, the default settings were used. Only protein groups with at least two peptides were included in the final data sets and all contaminants were removed. Protein groups were kept for further statistical analysis only if quantified in 3 out of 3 replicates (for the first experiment) and 2 out of 2 replicates (for the second experiment). To find significant differences between sensitive and resistant A375 cells, $\log_2$-transformed SILAC ratios were subjected to t test in Perseus, with a permutation-based FDR threshold of 0.01 and s_0 value of 1. The SILAC ratio ofidentified phosphorylation sites were normalized to the ratios of corresponding protein groups. All term enrichment analyses were performed using Perseus, by mapping reference lists or subsets of proteins with annotation terms. The resources used for annotation of proteins were Gene Ontology (GO), Biological Processes (GOBP), GO Cellular Compartment (GOCC), GO Molecular Functions (GOMF) and Kyoto Encyclopedia of Genes and Genomes (KEGG). In addition, significantly changing proteins determined by t test analysis (FDR ≤ 0.1, s_0 value of 1) were mapped to pathways and network associations using STRING (40) against the whole genome as the statistical

background. The top 20 significant pathways were selected. The median of the t test difference of assigned significantly changing proteins was calculated for each pathway. A list of all peptide sequences is provided in the supplementary Peptide Table and a list of all protein identifications and phosphorylation site identifications are provided in supplementary Table S1.

Analysis of Microarray Data from Pre- and Posttreated Patient Tumors— To confirm the relevance of our findings regarding nestin protein abundance profile in A375, we reanalyzed NES gene expression profile in human patients with melanoma pre- and posttreatment. The publically available microarray data with accession number GSE50509 and GSE61992 were retrieved from the NCBI GEO database. These data were acquired on Illumina Sentrix HumanHT12 v.4.0 Expression BeadChip as described in the original publications (41, 42). We reanalyzed the microarray datasets within the R environment (43). Each probe was annotated using the illuminaHumanv4.db package (44). Quality control analysis (probe intensity density plot) was performed on all microarrays to check for normalization bias between samples or batches. The normalized data were filtered based on probe quality using illuminaHumanv4.db package. Differentially expressed genes were identified from the filtered normalized data using the LIMMA package (45). Briefly, this was performed by blocking samples on a per-patient basis, fitting a linear model and generating Empirical Bayes statistics to compare pre- and posttreatment samples. The Benjamini-Hochberg multiple testing correction method was applied to all differentially expressed genes (supplementary Table S2).

Immunohistochemistry— Immunohistochemistry staining of clinical FFPE specimens (List of FFPE specimens used in this study in supplementary Information S3) was performed with a nestin specific polyclonal rabbit antibody (N5413, Sigma-Aldrich) diluted 1:50 in PBS containing 0.3% Triton-X100 and 1% BSA. Briefly, 5 μm FFPE tissue sections were de-paraffinized and antigen retrieval was performed in citrate buffer pH 6 in a pressure cooker for 2 min under pressure before a slow cooling down of the samples. Afterward tissue sections were stained according to the manufacturer's protocol (Thermo Scientific, UltraVision LP Detection System: AP Polymer) using FastRed (Thermo Scientific, Liquid Fast-Red Substrate System) as substrate.

Generation of Genome Edited Melanoma Cell Lines—NES gene knockout was carried out by CRISPR/Cas9-mediated genome editing according to the published protocol (46). The SpCas9 plasmid PX459 (Plasmid 62988) was obtained from a nonprofit plasmid share repository (Addgene, Watertown, MA). Suitable CRISPR target sites within NES Exon 1 positive strand were identified using the "CRISPR Design Tool" (http://crispr.mit.edu/). The respective oligonucleotide inserts (Biomers, Ulm, Germany) (5′-cctcgacggcgcgccggttg-3′ (forward), 5′-caaccggcgcgccgtcgagg-3′ (reverse)) were designed with overhangs compatible for ligation into PX459 linearized by digestion with BbsI (New England BioLabs, Frankfurt, Germany). Oligonucleotides were phosphorylated with polynucleotide kinase T4 PNK (New England BioLabs), annealed and inserted into the plasmid using T4 DNA ligase (New England BioLabs) and transformed into chemocompetent DH5α E. coli cells (New England Biolabs). Oligonucleotide inserts (ITD) (5′-gtattactgatattggtggg-3′ (forward), 5′-cccaccaatatcagtaatac-3′ (reverse)) were designed as CRISPR/Cas9 nontargeting (NonTar) control sgRNA and cloned into the SpCas9 plasmid PX459. Melanoma cells were seeded with low density (100,000 cells per ml), grown for 24 h in RPMI- 1640 medium without FBS. Transfection of the SpCas9/sgRNA plasmid or SpCas9/NonTar plasmid was carried out with Lipofectamine 2000 (Thermo Fischer Scientific) according to the manufacturer's instructions. On the next day, cells were selected using puromycin (2 μg/ml, Invivogen, Toulouse, France) and incubated for 2 days. Once individual colonies formed, single colonies were picked, cultured in separate wells and expanded in 6-well plates until cell number was enough for further analysis. Genomic DNA was isolated using GeneElute Mammalian Genomic DNA MiniPrep Kit (Sigma-Aldrich) and PCR amplification was performed using primers (5′-agatgtggggagctcaatcgg -3′ (forward) and 5′-tccaacctctgttccaacgc -3′ (reverse)) and sequenced by Sanger sequencing. Off-target effects of the guide sequence were predicted using Cas-OFFinder online tool (47). A mismatch of three bases were allowed. (List of predicted off-target sites in supplementary Information S4).

siRNA Knockdown— Nestin knockdown was carried out by transfecting A375 S cells with a pool offour FlexiTube siRNA oligos (100 nM, Qiagen, Hilden, Germany) against human nestin. NonSilencing siRNA (NonSil, Dharmacon) was used as control. siRNA transfections were performed using Lipofectamine RNAiMAX (Invitrogen) according to manufacturer's protocol. The medium was changed the following day. All assays were performed at 48–72 h posttransfection.

Western Blotting— Cells were harvested in lysis buffer and proteins were precipitated overnight with acetone/methanol (−20 °C). Cell lysates from SkMel28 S/R, SkMel19 S/R, 451lu S/R and Mel1617 S/R were used for immunoblotting analysis. Protein extracts were separated on 4–12% NuPAGE Bis-Tris gels (Novex, Life Technologies, Carlsbad, CA), transferred to PVDF membranes (0.2 μm, Sigma-Aldrich). The blot membranes were blocked in 1% Tween-20 and probed with primary antibody followed by horseradish peroxidase-conjugated secondary antibodies. Primary antibodies used were anti-Nestin (N5413, Sigma-Aldrich; sc-23927, Santa-Cruz Technologies), anti-EGFR (#2232, Cell Signaling Technologies, Danvers, MA), anti-Integrin β1 (E11) (sc-374430, Santa-Cruz Technolgies), anti-Integrin β4 (B4) (sc377523, Santa-Cruz Technologies), anti-GAPDH (MA5–15738, Thermo Fischer Scientific), anti-AKT (SAB4500797, Sigma Aldrich), anti-AKT (phospho-S124) (SAB4301497, Sigma Aldrich), anti-ERK1/2 (SAB1305560, Sigma Aldrich), anti-ERK1/2 (phospho-T202/Y204) (SAB1306604, Sigma Aldrich), anti-FAK (D1) (sc-271126, Santa-Cruz Technologies) and anti-Histone H3 antibody (D1H2) (#4499, Cell Signaling Technologies). Secondary antibodies used were anti-rabbit IgG, HRP-conjugated (#7074, Cell Signaling Technologies) and anti-mouse IgG, HRP-conjugated (#7075, Cell Signaling Technologies). ECL was detected by exposure with the Fusion SL instrument (Vilber Lourmat, Eberhardzell, Germany). For quantification, ImageJ software was used.

Clonogenic Assay— Cells were seeded in low density with 200 cells/cavity in 12-well plates and treated with $BRAF^{V600E}$ and MEK inhibitors (1 μM PLX4720, 0.1 μM cobimetinib). After 7–10 days, cells were fixed with 4% formaldehyde and stained with 0.05% Coomassie Brilliant Blue solution (Bio-Rad Laboratories, Hercules, CA) containing 80% methanol and 10% acetic acid. Stained colonies were washed twice with DPBS and counted under the microscope.

Proliferation/Cell Viability Assay— A MTS assay was used to analyze the proliferation and survival of melanoma cells. Cells were seeded (2×10^3) into 96-well plates. On the next day, media was changed, and cells were incubated for 96 h with increasing concentrations of respective inhibitors (0.1 μM to 20 μM of PLX4720 or cobimetinib) or with DMSO. After washing with DPBS, cell viability was assessed by CellTiter 96® Aqueous One Solution Cell Proliferation Assay (Promega Corporation), according to the manufacturer's instructions. The IC_{50} values were determined from the dose-response curves using log inhibitor versus response (three-parameter) test in GraphPad Prism.

Gelatin Zymography— For gelatin zymography, melanoma cells were cultivated to 70–80% confluence, washed twice with DPBS and grown for 24 h in FBS-free RPMI 1640 medium in presence or absence ofinhibitors (1 μM PLX4720, 0.1 μM cobimetinib). Conditioned media was collected, concentrated using SpeedVac and nonreducing sample buffer was added (4% SDS, 20% glycerol, 0.01% bromphenol blue, 125 mM Tris-HCL pH 6.8). Samples were

separated on Novex 10% Zymogram Plus (Gelatin) gels (Thermo Fischer Scientific). Gels were washed twice with washing buffer (2.5% Triton X-100, 50 m M Tris-HCl, 5 m M $CaCl_2$, 1 μM $ZnCl_2$), incubated for 24 h with incubation buffer (1% Triton X-100, 50 m M Tris-HCl, 5 m M $CaCl_2$, 1 μM $ZnCl_2$) at 37 °C, stained with staining solution (40% methanol, 10% acetic acid, 0.5% Coomassie Blue) fo r 1 h at RT and incubated with destaining solution (40% methanol, 10% acetic acid) until white bands appeared. For quantification of bands, ImageJ Software was used.

3D Melanoma Spheroid Culture— Cells were seeded on 1.5% agar noble (VWR) coated 96-well plates and incubated for 3 days to form spheroids. Spheroids were embedded into collagen I rat tail protein (1 mg/ml, Thermo Fischer) diluted in RPMI medium and cultured for four to ten more days. A total of 15 to 25 spheroids from each experiment were analyzed per condition. For quantification of outgrowth of spheroids, images were taken on day 0 and day 10. The outgrowth length for 15 spheroids was quantified using ImageJ software.

RESULTS

We used high-resolution mass spectrometry to characterize the proteome and phosphoproteome changes associated with resistance to BRAF inhibitor therapy in vitro (supplemental Fig. S1 A). To study acquired drug resistance, the BRAF ^{V600E} mutant melanoma cell lines sensitive (A375 S) and resistant to vemurafenib (A375 R) were used (33). For pro - teomic and phosphoproteomic analysis, A375 S and A375 R cells were subjected to stable isotope labeling by amino acids in cell culture (SILAC) using Lys4/Arg6 supplemented in growth medium. The resulting peptide mixture was measured using on-line liquid chromatography coupled to high-resolution mass spectrometry (LC-MS/MS). In addition, phosphopeptide enrichment was performed and analyzed by LC-MS/MS (supplemental Fig. S1A and supplementary Table S1 and supplementary Peptide Table).

Cytoskeletal Proteins Are Downregulated in Vemurafenib-resistant Melanoma Cells— We performed two proteomic screens (each in three biological replicates) that resulted in detection of 129,485 unique peptide sequences corresponding to 9453 distinct protein groups at an estimated FDR of 0.27% at the peptide and 1.28% at the protein level (Fig. 1 A, 1B; supplementary Table S1). Protein differential abundance analysis indicated widespread regulation of protein abundance that occurs during the development of BRAF resistance in vitro. The significantly downregulated proteins were involved in focal adhesion, integrin signaling, and actin cytoskeleton regulation (Fig. 1 C), upregulated proteins in the activation of PI3K/AKT, mTOR, and MAPK/ERK signaling pathway (Fig. 1 C). In the phosphoproteome measurement, we identified 13,354 phosphorylation sites, of which 10,444 were localized to a specific Ser/Thr/Tyr residue (supplemental Fig. S1 B and supplementary Table S1) and 9106 phosphorylation sites were quantified reproducibly in all three replicates. We mapped our dataset to the known resistance mechanisms MAPK/ERK and PI3K/AKT pathway and observed a good coverage of regulated key proteins (supplemental Fig. S1 C). For example, the RTKs EGFR, PDGFR, and IGF1R were identified upregulated in A375 R versus A375 S cells, as well as NRAS, BRAF, ERK1, MITF, AKT2, and mTOR (supplemental Fig. S1 C). In addition, we identified phosphorylation sites on AKT (S124) and ERK1/2 (T202/Y204) significantly upregulated in A375 R cells (supplemental Fig. S1 C).

Among pathways activated in the A375 R cells, multiple proteins involved in cytoskeletal organization were significantly over-represented (Fig. 1 A and 1B). Reorganization of the cytoskeleton is often associated with migratory and invasive phenotype of tumor cells and contribute to cancer's aggressiveness. In this dataset, a remarkable number of differentially regulated proteins were involved in cytoskeleton and adhesion pathways (Fig. 1 A, 1B, and 1 C), and several key proteins, such as nestin, vimentin and gelsolin, showed significant changes in abundance. The type VI intermediate filament protein nestin showed a median $\log_2$ ratio of −2.71 in resistant cells in all replicates and was one of the highest changing proteins in the whole dataset. To validate these findings, we assessed expression of nestin by Western blotting in several widely used melanoma cell lines, such as SkMel28 and Mel1617 (Fig. 1 D). Quantification of the Western blotting by densitometry revealed significant changes of nestin abundance between drug-sensitive and drug-resistant cells (Fig. 1 E).

To further address the clinical relevance of the downregulation of nestin in patients, we used two public transcriptomic datasets available from NCBI GEO repository. We selected all microarray data from patients with matching pre- and post-treatment tumors (vemurafenib, dabrafenib and dabrafenib/trametinib); i.e. 21 and 9 patients from microarray datasets GSE50509 and GSE61992, respectively (41, 42). The bioin - formatics reanalysis of these microarray data revealed 23, 0 and 17 significantly changing probe sets in vemurafenib, dabrafenib and dabrafenib/trametinib treated tumors versus pre-treatment tumors (FDR ≤ 0.1, supplementary Table S2). The mRNA expression of nestin (probeset ID = ILMN_1738147) was not significantly changing across posttreatment tumors versus pretreatment tumors. However, we did observe a tendency toward nestin downregulation between paired tumors (supplemental Fig. S1 D and S1 E). We could confirm these findings in FFPE specimens of two patients pre- and post-therapy with BRAF inhibitor using immunohistochemistry against nestin protein (supplemental Fig. S1 F). The abundance of nestin in pretreated tumors differed between patients; however, there was a clear tendency toward nestin downregulation between post- and pretreated tumors within each individual patient. Proteomic analysis of FFPE specimens of one patient pre- and post-BRAF inhibitor therapy identified nestin as one of the most downregulated proteins in the post-BRAF inhibitor treated tissue (supplemental Fig. S1 G) with a $\log_2$ ratio of −2.5. In total, we identified 14 peptides of nestin in both tissue specimens with a sequence coverage of 36.5%. Taken together, these data show that acquired resistance is associated with alterations in cytoskeletal and adhesion molecules and leads to the downregulation

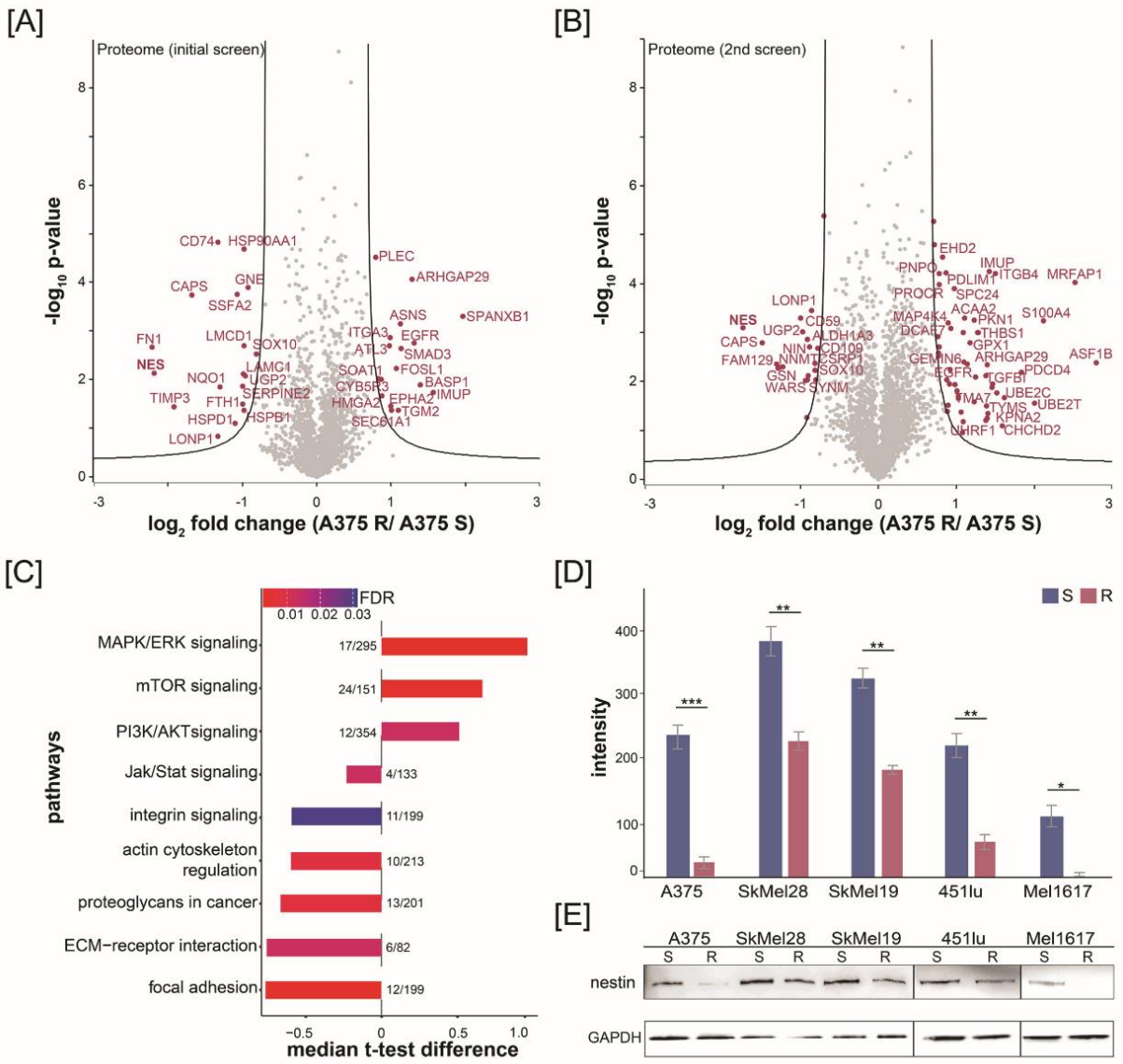


FIG. 1. Quantitative proteomics of melanoma cells identifies downregulation of nestin in drug-resistant cells. Volcano plot of vemurafenib-resistant and -sensitive A375 proteomes for (A) the initial proteomic screen and (B) the second proteomic screen. Fold change of SILAC ratios between A375 R and A375 S (log $_2$) are plotted against p value ($-\log_{10}$) (n = 3). Black lines indicate the significance threshold (FDR < 0.01; s0 = 1). Significantly up- and downregulated proteins are highlighted in magenta. C, Over-representation of selected KEGG signaling pathways of A375 R compared with A375 S cells using String database analysis. The t test difference of SILAC ratios between A375 R and A375 S (log $_2$) were plotted for each pathway (t test, FDR < 0.1; s $_0$ = 1). Numbers represent identified significantly changing proteins mapped to the pathway by the total protein count involved in that pathway. Color of the bars represents the FDR. D, and E, Western blot analysis of different drug-resistant and drug-sensitive melanoma cell lines (A375, SkMel28, SkMel19, 451lu and Mel1617) against nestin and quantification of signal intensities using ImageJ software.

of the intermediate protein nestin in melanoma cell lines and metastases of human patients.

Nestin Expression Correlates with Invasive Properties of A375 Cells— To evaluate the functional role of nestin in resistance toward BRAF inhibition, we induced a NES gene knockout in vemurafenib sensitive A375 S cells using the CRISPR/Cas9 system (supplemental Fig. S2 A). Single clones (A375 Nes-Ko #1–5) were selected for further analysis based on their effective NES knockout (supplemental Fig. S2 B–S2 E). As a control we used a nontargeting (NonTar) control guide sequence

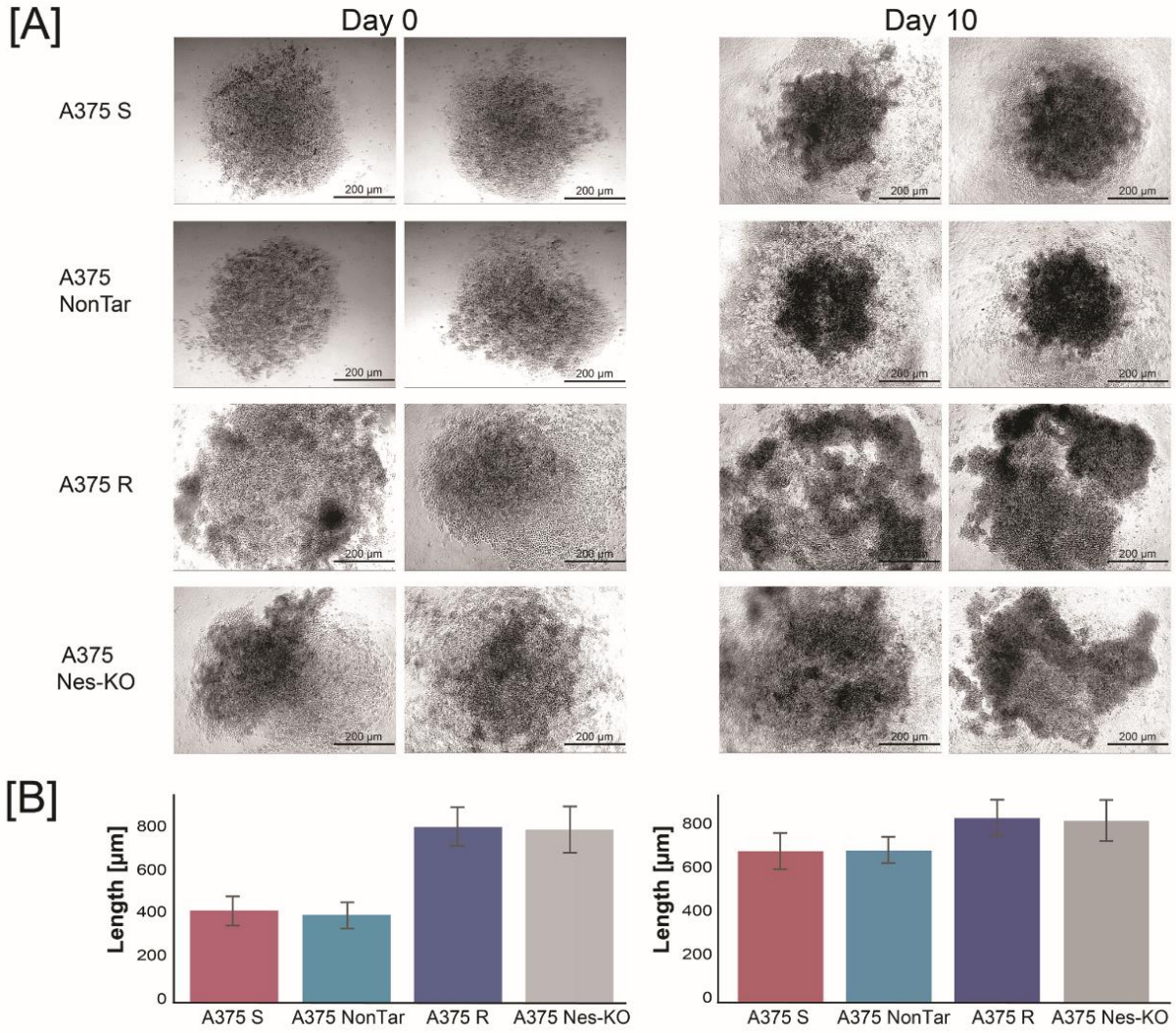


FIG. 2. Nestin expression correlates with invasive properties of melanoma cell lines. A, Anchorage-independent growth assays of A375 S, A375 NonTar, A375 R and A375 Nes-KO clone #1 cells. Cells were seeded on agar, incubated for 3 days and spheroids were embedded into a collagen I matrix and further incubated for 10 days. Images are representative of three independent experiments of day 0 and day 10 (15–20 spheroids per cell line and day). All images were acquired with a light microscope at 5-fold magnification. Scale bar: 200 µm. B, Quantification of spheroids outgrowth. Analysis of outgrowth length was performed by measuring the total length from the center for 15 spheroids each on microscopic images using ImageJ software. Results are displayed as mean values ± S.E.

and transfected A375 S cells with this construct. A375 NonTar cells showed no InDels in the NES gene and no change in nestin protein abundance (supplemental Fig. S2 F and S2 G).

To address the relationship between nestin abundance and the invasive properties of melanoma cells, we performed three-dimensional spherical outgrowth assay. Melanoma spheroids formed from A375 S and A375 NonTar cells increased in size over 10 days (Fig. 2 A and 2B). However, vemurafenib-resistant and A375 Nes-KO spheroids were not compact, showed less cellular adhesion and exhibited cells that progressively infiltrated into the surrounding collagen gel. In contrast, A375 S and A375 NonTar spheroids showed restricted invasive movement of a few cells away from the spheroid edge. Loss of nestin expression reduced the spheroid-forming ability of A375 Nes-KO cells like A375 R cells. Collectively, these data show that expression of nestin correlates with an invasive phenotype of resistant melanoma cells.

Depletion of Nestin Affects Cell Proliferation and Colony Formation On Treatment with BRAF ^{V600E} and MEK Inhibitors—To investigate the effects on cell proliferation in drug-sensitive, drug-resistant and genome edited A375 Nes-KO cell lines, we treated cells with the MAPK signaling pathway

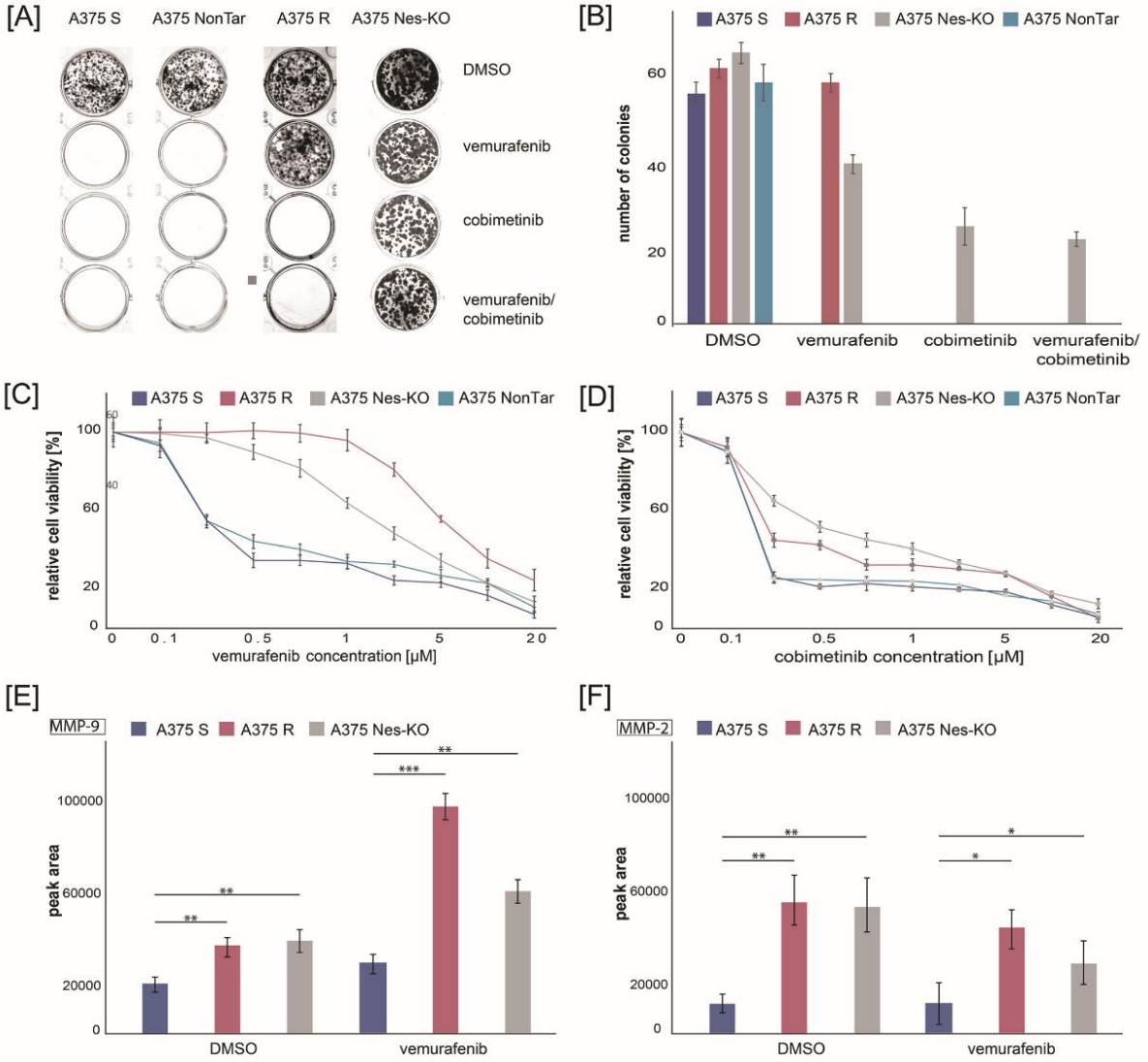


FIG. 3. Depletion of nestin affects cell proliferation and colony formation on treatment with BRAF ^{V600E} and MEK inhibitors. A, Clonogenic assay of A375 S, A375 NonTar, A375 R and NES knockout (A375 Nes-KO) cells after 10 days treatment with signaling pathway inhibitors. Cells were treated with vemurafenib (1 μM) and cobimetinib (0.1 μM), either alone or in combinations. Controls were treated with DMSO. Cultures were fixed with formaldehyde and stained with crystal blue. Images are representative of three biological and three technical replicates. B, Quantification of single colonies per condition. Results represent the mean of three biological experiments and three technical replicates (9 wells per cell line). Error bar represents standard deviations C–D, A375 S, A375 NonTar, A375 R, and A375 Nes-KO were cultured for 24 h, and then treated with vemurafenib (C) or cobimetinib (D) at the indicated concentrations (0, 0.1, 0.25, 0.5, 1, 2.5, 5, 10 and 20 μM) or DMSO as control. Cell viability was determined by MTS assay 96 h later. Results expressed as percentage of control represent the mean of three biological experiments and six technical replicates (n = 24). Error bar represents standard deviations of replicates. E–F, Activity of MMP9 (E) and MMP2 (F) using gelatin zymography. Supernatants of A375 S, A375 R and A375 Nes-KO cell lines were treated with DMSO or vemurafenib for 24 h, analyzed by zymography gelatin plus gels and stained with Coomassie. ImageJ software was used for quantification of signals. Results represent the mean of three biological experiments ± S.E.

inhibitors vemurafenib (BRAF ^{V600E} inhibitor) and cobimetinib (MEK inhibitor) and performed in vitro colony formation and cell proliferation (MTS) assays. First, we examined the colony formation ability of A375 Nes-KO compared with A375 S and A375 R controls. A375 NonTar cells were used as CRISPR/Cas9 control cell line. In absence ofinhibitors, cell lines showed similar colony formation ability in terms of colony number and size (Fig. 3 A). A375 S and A375 NonTar single

cells were not able to grow into a colony after treatment with the BRAF ^{V600E} inhibitor vemurafenib. In contrast, A375 R showed the same colony formation ability with vemurafenib. For the genome edited A375 Nes-KO cell lines, we observed a 2-fold decrease in colony number (Fig. 3 A and 3B). However, the size of the BRAFi treated nestin knockout cell line colonies were smaller (data not shown). When cells were treated with the MEK inhibitor cobimetinib or in combination with BRAFi, the A375 S and the A375 R cell lines showed no ability to grow into colonies (Fig. 3 A and 3B). Interestingly, after both treatments, the A375 Nes-KO cell lines colonies were observed in similar numbers and sizes (Fig. 3 A and 3B). Next, we examined whether different concentrations of BRAF ^{V600E} and MEK inhibitors impair cell survival and proliferation of melanoma cells. We calculated the IC_{50} value for each cell line and treatment using a three-parameter dose-response test. When treated with vemurafenib, the IC_{50} of A375 R, A375 S, A375 NonTar, and A375 Nes-KO was 2.46 μM, 0.268 μM, 0.259 μM and 2.091 μM, respectively (Fig. 3 C). This confirmed that A375 Nes-KO cells tolerate vemurafenib treatment better than the A375 S cells. In addition, the observed effect was constant over different concentrations of BRAF and MEK inhibitors. No difference was observed between A375 S and A375 NonTar cells. For the cobimetinib treatment, the IC_{50} of A375 Nes-KO cells (IC_{50} 0.31 μM) was 7.7-times higher than for A375 R cells (IC_{50} 0.04 μM) (Fig. 3 D). A375 S and A375 NonTar cells were not able to grow under these conditions.

To confirm this finding, we used a siRNA loss-of-function model system and investigated the phenotype of a conditional nestin knockdown (A375 Nes- K_d) with regards to vemurafenib sensitivity. Downregulation of nestin expression was effectively achieved in A375 S cells as opposed to cells harboring nonsilencing siRNA (NonSil) (supplemental Fig. S3 A). In agreement to NES gene knockout (Nes-KO), nestin RNA downregulation resulted in increased cell proliferation on BRAF ^{V600E} kinase inhibition compared with A375 S cells as seen in the MTS assay (supplemental Fig. S3 B). The IC_{50} of vemurafenib in A375 Nes- K_d cells was similar (1.454 μM) to the NES knockout cells; pointing to the fact that nestin affects cell survival on stimulation with different BRAF ^{V600E} and MEK inhibitors.

Depletion of Nestin Induces Matrix Metalloproteinase Activity—Downregulation of nestin was previously connected with increased activity of matrix metalloproteinases (MMP) MMP-9 and MMP-2, which are involved in migratory potential and invasiveness of cancer cells (30). We hypothesized that A375 R with a significantly downregulated expression of nestin will induce melanoma matrix metalloproteinases and that this effect will be mimicked by our functional knockout cell line A375 Nes-KO. To maintain enzymatic activity of MMP-9 and MMP-2, we treated A375 S, A375 R and A375 Nes-KO cells with vemurafenib or DMSO as a control for 24 h and analyzed the supernatants using gelatin zymography (supplemental Fig. S3 C). All cell lines secreted MMP-2 and MMP-9; however, A375 R and A375 Nes-KO cells showed a significantly stronger signal for both metalloproteinases compared with the supernatant of A375 S cells (supplemental Fig. S3 C). This pattern was like that observed in cells cultured with the BRAF ^{V600E} inhibitor. A 2-fold difference in MMP-9 activity and 3-fold difference in MMP-2 activity was observed between untreated supernatants of drug-sensitive and drug-resistant lines (Fig. 3 E and 3F). Interestingly, quantification of signals resulted in a similar peak area for bands of MMP-9 and MMP-2 in A375 Nes-KO and A375 R supernatants (Fig. 3 E and 3F), highlighting the similarity of their phenotypes. Treatment of cells with the BRAF ^{V600E} inhibitor led to significant increased activation of MMP-9 in A375 R and A375 Nes-KO supernatants and decreased activation of MMP-2 in genome edited A375 Nes-KO lines (Fig. 3 E and 3F). These results suggest a relationship between expression of nestin and activity of certain MMPs known to enhance tumor invasion.

Loss of Nestin Expression Is Associated with PI3K/AKT and Integrin Signaling— To understand how the loss of nestin protein abundance may alter cellular protein homeostasis, we performed a quantitative (phospho)proteome analysis of A375 Nes-KO versus A375 S versus A375 R cells using a SILAC approach and LC-MS/MS. In two biological replicates, we identified 5965 protein groups and 7524 phosphorylation sites, of which 91 showed significant changes in abundance at a FDR < 0.01 ($s_0 = 1$) in the A375 Nes-KO compared with A375 S (Fig. 4 A and 4B, supplementary Table S1 and supplementary Peptide Table). The comparison of A375 R to A375 Nes-KO cells showed significant differences between these two cell lines (supplemental Fig. S4 A). Biological replicates showed a good correlation of the proteome and phosphoproteome (supplemental Fig. S4 B and S4 C). Interestingly, ECM interacting proteins, such as vinculin, fibronectin, integrin β4 and integrin α6, as well as Proteinkinase C, focal adhesion kinase FAK and other downstream signaling proteins were significantly upregulated in the genome-edited cells (Fig. 4 A). Enrichment analysis of the significantly regulated proteins for GO cellular component indicated an involvement in adhesion junctions, extracellular region and focal adhesion. (Supplemental Fig. S4 D). The KEGG pathways PI3K/AKT signaling, remodeling offocal adhesions, actin cytoskeleton signaling and integrin signaling were significantly over-represented in A375 Nes-KO cells compared with A375 S cells (supplemental Fig. S4 D). Furthermore, the phosphoproteome analysis revealed differentially regulated phosphorylation sites on key members of the integrin signaling pathway and downstream proteins. Phosphorylation sites on FAK, ERK1/2 and Integrin β4 were significantly upregulated in the genome edited cell line compared with A375 S cells (Fig. 4 B). To identify the (phospho)proteomic overlap between A375 Nes-KO and A375 R cells, we correlated the ratios of A375 Nes-KO versus A375 S against ratios of A375 R versus A375 S (Fig. 4 C and 4D). We identified differentially regulated proteins and phos-

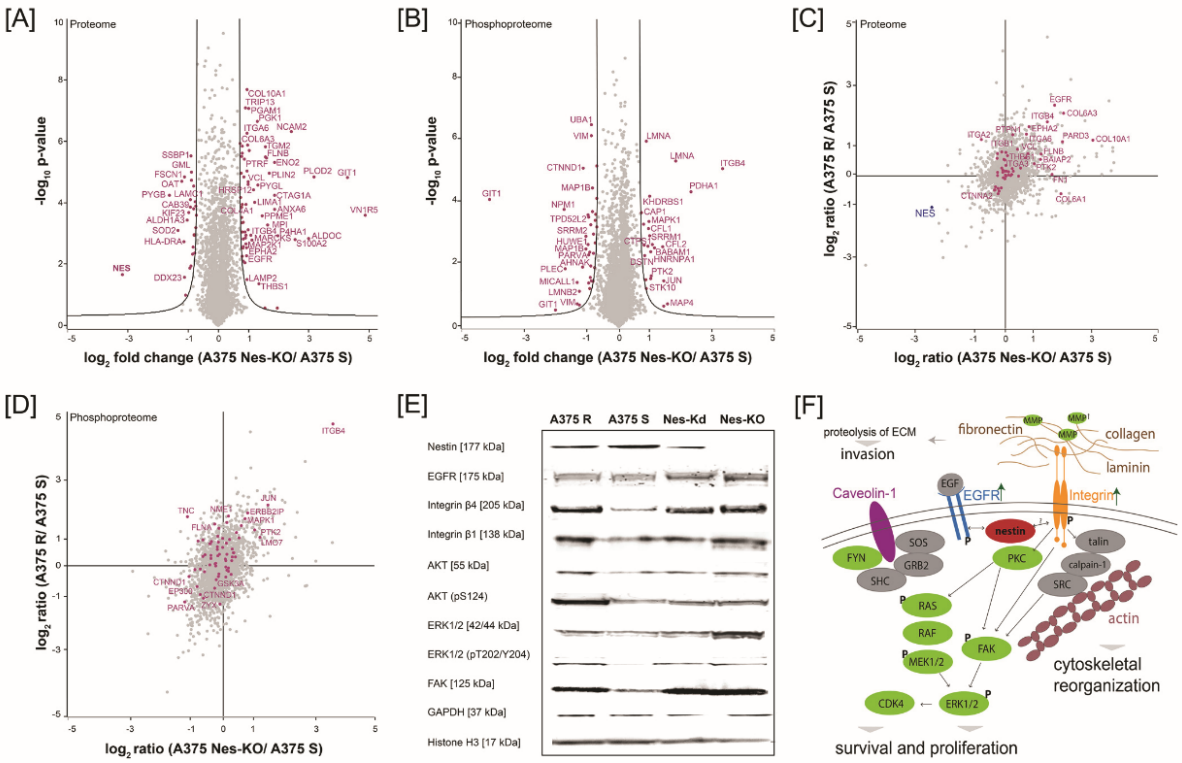


FIG. 4. Loss of nestin is associated with to PI3K/AKT signaling. Volcano plot of nestin CRISPR/Cas9 knockout (Nes-KO) versus vemurafenib-sensitive (A375 S) cells (A) proteome and (B) phosphoproteome. Fold change of SILAC ratios between A375 Nes-KO and A375 S ($\log_2$) are plotted against p value ($-\log_{10}$) ($n = 3$). Black lines indicate the significance threshold (FDR < 0.01; $s_0 = 1$). Significantly up- and downregulated proteins are highlighted in magenta. C, Proteome correlation of Nes-KO relative to A375 S against A375 R relative to A375 S. Magenta: proteins involved in integrin signaling using GO annotation. D, Correlation of phosphorylation sites of Nes-KO relative to A375 S against A375 R relative to A375 S. Magenta: proteins involved in integrin signaling using GO annotation. E, Western blot analysis of A375 R, A375 S, siRNA knockdown cells Nes- K_d and CRISPR/Cas9 knockout cells Nes-KO against key molecules of the integrin signaling, MAPK/ERK and PI3K/AKT signaling pathways. F, Schematic overview of proposed interaction of nestin with key molecules ofintegrin, MAPK/ERK signaling pathway. Green: upregulated in A375 Nes-KO versus A375 S cells, red: downregulated in A375 Nes-KO versus A375 S cells, gray: not quantified; P: identified upregulated phosphorylation site in A375 Nes-KO versus A375 S cells; arrows: indicate upregulated receptors in A375 Nes-KO versus A375 S cells.

phorylation sites on key members of the integrin signaling pathway in A375 Nes-KO and A375 R cells with similar ratios (Fig. 4 C and 4 D). For example, abundance of EGFR, Filamin-B and Collagen 6A3 and 10A1, as well as phosphorylation sites on Integrin $\beta 4$, FAK and ERK1. Focal adhesion kinase FAK plays a central role in cancer cell motility, adhesion, and invasion (48, 49). In melanoma cells, we found Nes-KO to be associated with increased FAK phosphorylation and protein levels (Figs. 4 B). Phosphorylation of FAK at S910 was identified with high quality by LC-MS/MS (supplemental Figs. S4 E). We confirmed these results by immunoblot using antibodies against key signaling molecules in NES knockout (A375 Nes-KO), nestin knockdown (A375 Nes- K_d), A375 S and A375 R cells (Fig. 4 E). Immunoblot revealed a higher expression of EGFR, Integrin $\beta 4$ and FAK in A375 R, A375 Nes-KO and Nes- K_d cells. Signals for phosphorylation sites of ERK1/2 (T202/Y204) and AKT (S124) were also increased in A375 R, knockout and knockdown cells compared with A375 S. Taken together, these results link nestin with integrin and PI3K/AKT signaling pathways in melanoma; while also revealing new molecular events in context of acquired resistance to vemurafenib (Fig. 4 F).

DISCUSSION

To identify additional resistance mechanisms and reveal new molecular targets to overcome resistance, we investigated two melanoma cell lines (A375 S and A375 R) with differing phenotypes of acquired resistance to the BRAF ^{V600E} inhibitor vemurafenib. In this context, we utilized SILAC coupled to mass spectrometry to characterize the global proteomic and phosphoproteomic changes in melanoma cell lines. This study is, to our knowledge, one of the largest global

(phospho)proteomic analyses assessing the differentially expressed proteins in drug-sensitive and drug-resistant melanoma cells. We identified several pathways to be over-represented in resistant cells including PI3K/AKT/mTOR signaling, integrin signaling and MAPK/ERK signaling pathways. The MAPK/ERK and PI3K/AKT/mTOR signaling pathways are known resistance mechanisms and constitutively activated in malignant melanoma. Here, several key proteins and phosphorylation sites within these pathways were identified with high confidence like EGFR or phosphorylated ERK1/2. In addition to signaling pathways, the tumor microenvironment and remodeling of the cytoskeletal organization have been reported to play an important role in the development of acquired resistance. For example, Kim and colleagues showed that actin signaling through YAP/TAZ activation confers BRAF inhibitor resistance in melanoma (50). Similarly, we detected several cytoskeletal proteins such as nestin, vimentin and gelsolin to be downregulated in resistant cells. In the present study, we investigated whether the intermediate filament nestin may contribute to resistance in melanoma cells. Expression of nestin in various cancer cell types has been studied, however the mechanistic basis of the function of nestin is still unknown. Nestin was reported to be involved in cancer cell migration, invasion, and metastasis (18, 31, 51). Quendro and colleagues showed in a large-scale proteomic study that nestin and vimentin are both upregulated in melanoma cells and tissue material compared with control melanocytes (29). We could confirm this in A375 melanoma cells and further show that nestin and vimentin are downregulated in resistant cells compared with sensitive cells. Nestin and vimentin are interaction partners with important functions in cell migration, cytoskeletal reorganization and apoptosis (52). Doxie and colleagues showed that nestin expression was completely depleted in nestin-expressing cells in human tumors after BRAF and MEK inhibitor therapy, which highlights the loss of nestin expression in human tumors. In previous reports, NES expression has been reported to be regulated by the transcription factors SOX9 and SOX10 and nestin and SOX9 may be negative prognostic markers in melanoma (53). In agreement with this, both transcription factors were also identified in our data set with the same abundance trend as nestin. Because nestin is a known stemness marker (27) we investigated the presence of other stemness markers in our data set. We could identify the stemness marker ABCG2, which is known to enhance tumorigenic potential of melanoma cells (54). Although ABCG2 was not significantly regulated, we identified the potential cancer stem cell (CSC) marker of ALDH1 to be upregulated in drug-resistant and in nestin CRISPR/Cas9 knockout cells compared with sensitive cells. ALDH1 is associated with multidrug resistance in different types of human melanoma tumors (55) and therefore may influence the stemness of melanoma cells. Taken together, this highlights the good coverage of the dataset and utility as a resource for the melanoma community.

To confirm the clinical significance of our findings, we reanalyzed public microarray datasets for matching pre- and posttreatment tumors. We could only observe a tendency toward NES downregulation, however in none of the treatments did it reach significance. Several hypothetical reasons could explain this. The first reason may be that nestin clinical relevance is patient specific and depends on the patient genetic background or tumor microenvironment. The second contributing factor may be that, contrary to A375 cell line where we have a clear BRAFi sensitive/resistant phenotype, in pre- and posttreatment tumors the response to BRAFi is not well defined (i.e. nonresponder, partial responder, full responder). In addition, we could confirm downregulation of nestin in FFPE specimens using immunohistochemistry; however, expression of nestin differed between tumor specimens. Quantitative proteomics of one pair of pre- and post-treated tumors identified nestin as one of the most downregulated proteins in the dataset with a good sequence coverage. These results highlight the significance of nestin expression in human tumors.

To study the effect of nestin, we used CRISPR/Cas9 approach to generate a NES gene knockout in drug-sensitive melanoma cells. We identified two peptides at the N terminus of the protein by high-resolution mass spectrometry because the knockout occurred at the end of Exon 1 of the genomic sequence. However, the knockout cell line was considered as an effective functional knockout of nestin because both interaction and functional active domains were absent from the resulting protein. To confirm that our results are not because of cell manipulation in terms of Cas9 expression, guide RNA transfection or single cell picking, we generated a CRISPR/Cas9 control cell line (A375 NonTar) using a nontargeting guide sequence. We could not observe differences in the cell proliferation and colony formation between A375 NonTar and drug-sensitive cell lines. This study provides novel data showing that nestin expression significantly correlates with cell survival and colony formation on MAPK signaling pathway inhibitor treatment. Indeed, we are not only describing an increased cell survival and colony formation ability in knockout cells under BRAF inhibitor treatment, but we also show a direct effect of NES expression on the growth of melanoma cells on inhibition with the MEK inhibitor cobimetinib. Several studies have suggested that combined therapy with BRAF and MEK inhibitors are promising trials to delay MAPK-driven acquired resistance and may activate other resistance mechanisms like PI3K/AKT/mTOR signaling pathways (56). Depletion of nestin may activate these resistance mechanisms and increase cell survival on mono- or combined therapy.

Our results indicate a phenotypic difference in invasion and proliferation of drug-sensitive and drug-resistant cells. Interestingly, the genome edited A375 Nes-KO cell lines derived from drug-sensitive cells showed a similarly invasive phenotype to drug-resistant cells. However, expression of nestin has been reported to mediate both, three-dimensional tumor-

igenesis and cell invasiveness (19, 25). It is also reported that depletion of nestin using shRNA results in an invasive phenotype of melanoma cell lines, which is mediated through up-regulation of specific matrix metalloproteinases (MMPs) (30). In this study, we could also show that nestin gene knockout induces activity of MMP2 and MMP9 like resistant phenotype. This falls in line with the hypothesis proposed by Lee and colleagues, who identified nestin depletion to be associated with the activation of MMP-2 (gelatinase A, type IV collagenase) and MMP-9 (gelatinase B, type IV collagenase) (30). The invasion and metastasis of tumor cells have been shown to require proteolytic activity to degrade components of the extracellular matrix (57) and to involve FAK/integrin-mediated cell/matrix adhesion pathways (58). We provide evi dence that nestin depletion is associated with signaling through focal adhesion, integrin and PI3K/AKT/mTOR pathways. Interestingly, ECM interacting proteins, like Laminin-B or Filamin-B, the integrins β1 and β4, Proteinkinase C, FAK and other downstream signaling proteins were significantly upregulated in the genome edited cells compared with drug sensitive cells at the proteome level. FAK activated by integrins plays a central role in cell invasion and adhesion by triggering several signaling pathways. Furthermore, the phosphoproteome analysis revealed differentially regulated phosphorylation sites on the key players of the integrin signaling pathway and downstream proteins. In this regard, recent evidence in prostate cancer research indicate that nestin depletion is associated with an expression pattern of phosphorylated FAK (pFAK) at the cell membrane. Phosphorylated FAK promotes integrin clustering, which results in pFAK- and integrin-dependent matrix degradation and an invasive phenotype (31). In melanoma cells, Hyder and colleagues could also observe increased levels of phosphorylated FAK and protein levels in nestin knockdown cells and a localization of phosphorylated FAK at the cell membrane like the studies in prostate cancer (31). In conclusion, we could link nestin protein levels, not only with an invasive phenotype, but also with acquired drug resistance in melanoma.

Acknowledgment— We thank Bianca Kuhn for her help in the initial stages of the project.

DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org/) via the PRIDE partner repository (59) with the data set identifier PXD010683.

Excel files containing the analyzed data are provided in Supplementary Materials.

* This work was supported by a grant from DFG (INST 37/741-1 FUGG) to B.M.

This article contains supplemental Figures, Tables, and Information. The authors declare no conflicts ofinterest.

¶ To whom correspondence should be addressed: Quantitative Proteomics, Director, Proteome Center Tuebingen, Interfaculty Institute for Cell Biology, University of Tuebingen, Auf der Morgenstelle 15, 72076 Tuebingen, Germany, Tel.: +49/(0)7071/29-70556; Fax: +49/(0)7071/29-5779; E-Mail: boris.macek@uni-tuebingen.de.

Author contributions: M.S., T.S., B.S., and B.M. designed research; M.S., T.S., and A.M. performed research; M.S., T.S., N.C.N., and B.M. analyzed data; M.S., T.S., B.S., and B.M. wrote the paper; B.S. contributed new reagents/analytic tools.

REFERENCES

1. Chang, A. E., Hynds Karnell, L., and Menck, H. R. (1998) The National Cancer Data Base Report on cutaneous and noncutaneous melanoma Cancer 83, 1664–1678
2. American Cancer Society (2017) Cancer Facts & Figures 2017. Atlanta: American Cancer Society
3. Larkin, J., Ascierto, P. A., Dreno, B., Atkinson, V., Liszkay, G., Maio, M., Mandalà, M., Demidov, L., Stroyakovskiy, D., Thomas, L., Cruz-merino, Dutriaux, L. D. C., Garbe, C., Sovak, M. A., Chang, I., Choong, N., Hack, S. P., McArthur, G., and Ribas, A. (2014) Combined vemurafenib and cobimetinib in BRAF-mutated melanoma. N Engl J. Med. 371, 1867–1876
4. Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd, Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson Ba Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W. C., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R., and Futreal, P. A. (2002) Mutations of the BRAF gene in human cancer. Nature 417, 949–954
5. Chapman, P. B., Hauschild, A., Robert, C., Haanen, J. B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., Hogg, D., Lorigan, P., Lebbe, C., Jouary, T., Schadendorf, D., Ribas, A., O'Day, S. J., Sosman, J. A., Kirkwood, J. M., Eggermont, A. M. M., Dreno, B., Nolop, K., Li, J., Nelson, B., Hou, J., Lee, R. J., Flaherty, K. T., and McArthur, G. A. (2011) Improved survival with vemurafenib in melanoma with BRAF V600E mutation. N. Eng. J. Med. 364, 2507–2516
6. Sosman, J. A., Kim, K. B., Schuchter, L., Gonzales, R., Pavlick, A. C., Weber, J. S., Joe, A. K., and Ribas, A. (2012) Survival in BRAF V600–mutant advanced melanoma treated with vemurafenib. N. Eng. J. Med. 366, 707–714
7. Flaherty, K. T., Robert, C., Hersey, P., Nathan, P., Garbe, C., Milhem, M., Demidov, L. V., Hassel, J. C., Rutkowski, P., Mohr, P., Dummer, R., Trefzer, U., Larkin, J. M. G., Utikal, J., Dreno, B., Nyakas, M., Middleton, M. R., Becker, J. C., Casey, M., Sherman, L. J., Wu, F. S., Ouellet, D., Martin, A.-M., Patel, K., and Schadendorf, D. (2012) Improved survival with MEK inhibition in BRAF-mutated melanoma. N. Engl. J. Med. 367, 107–114
8. Roesch, A. (2015) Tumor heterogeneity and plasticity as elusive drivers for resistance to MAPK pathway inhibition in melanoma. Oncogene 34, 2951–2957
9. Kemper, K., Krijgsman, O., Cornelissen-Steijger, P., Shahrabi, A., Weeber, F., Song, J.-Y., Kuilman, T., Vis, D. J., Wessels, L. F., Voest, E. E., Schumacher, T. N., Blank, C. U., Adams, D. J., Haanen, J. B., and Peeper, D. S. (2015) Intra- and inter-tumor heterogeneity in a vemurafenib-resistant melanoma patient and derived xenografts. EMBO Mol. Med. 7, e201404914
10. Nazarian, R., Shi, H., Wang, Q., Kong, X., Koya, R. C., Lee, H., Chen, Z., Lee, M.-K., Attar, N., Sazegar, H., Chodon, T., Nelson, S. F., McArthur, G., Sosman, J. A., Ribas, A., and Lo, R. S. (2010) Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. Nature 468, 973–977
11. Allen, E. M. V., Wagle, N., Sucker, A., Treacy, D., Goetz, E. M., Place, C. S., Taylor-weiner, A., Kryukov, G., Hodis, E., Rosenberg, M., McKenna, A., Cibulskis, K., Farlow, D., Zimmer, L., Hillen, U., Gutzmer, R., Goldinger, M., Ugurel, S., Gogas, H. J., Egberts, F., and Berking, C. (2014) The genetic landscape of clinical resistance to RAF inhibition in metastatic melnaoma. Cancer Discov. 4, 94–109

12. Shi, H., Hugo, W., and Kong, X. (2014) Acquired resistance and clonal evolution in melanoma during BRAF inhibior therapy. Cancer Discov. 144, 724–732
13. Wagle, N., Emery, C., Berger, M. F., Davis, M. J., Sawyer, A., Pochanard, P., Kehoe, S. M., Johannessen, C. M., MacConaill, L. E., Hahn, W. C., Meyerson, M., and Garraway, L. A. (2011) Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. J. Clin. Oncol. 29, 3085–3096
14. Lendahl, U., Zimmermann, L. B., and McKay R. D. (1990) CNS stem cells express a new class ofintermediate filament protein. Cell 60, 585–595
15. Sahlgren, C. M., Pallari, H. M., He, T., Chou, Y. H., Goldman, R. D., Eriksson, J. E. (2006) A nestin scaffold links Cdk5/p35 signaling to oxidant-induced cell death. EMBO J. 25, 4808–4819
16. Wiese, C., Rolletschek, A., Kania, G., Blyszczuk, P., Tarasov, K. V., Tarasova, Y., Wersto, R. P., Boheler, K. R., and Wobus, A. M. (2004) Nestin expression–a property of multi-lineage progenitor cells? Cell Mol. Life Sci. 61, 2510–2522
17. Kawamoto, M., Ishiwata, T., Cho, K., Uchida, E., Korc, M., Naito, Z., and Tajiri, T. (2009) Nestin expression correlates with nerve and retroperitoneal tissue invasion in pancreatic cancer. Hum. Pathol. 40, 189–198
18. Matsuda, Y., Naito, Z., Kawahara, K., Nakazawa, N., Korc, M., and Ishiwata, T. (2011) Nestin is a novel target for suppressing pancreatic cancer cell migration, invasion and metastasis. Cancer Biol Ther. 11, 512–523
19. Kleeberger, W., Bova, G. S., Nielsen, M. E., Herawi, M., Chuang, A. Y., Epstein, J. I., and Berman, D. M. (2011) Roles for the stem cell associated intermediate filament Nestin in prostate cancer migration and metastasis. Cancer Res. 67, 9199–9206
20. Li, H., Cherukuri, P., Li, N., Cowling, V., Spinella, M., Cole, M., Godwin, A. K., Wells, W., and Direnzo, J. (2007) Nestin is expressed in the basal/myoepithelial layer of the mammary gland and is a selective marker of basal epithelial breast tumors. Cancer Res. 67, 501–511
21. Chinnaiyan, P., Wang, M., Rojiani, A. M., Tofilon, P. J., Chakravarti, A., Ang, K. K., Zhang, H.-Z., Hammond, E., Curran, W., Jr, and Mehta, M. P. The prognostic value of nestin expression in newly diagnosed glioblastoma: Report from the Radiation Therapy Oncology Group. Radiat. Oncol. 8, 1–8
22. Yang, X. H., Wu, Q. L., Yu, X. B., Xu, C. X., Ma, B. F., Zhang, X. M., Li, S. N., Lahn, B. T., and Xiang, A. P. (2001) Nestin expression in different tumours and its relevance to malignant grade. J. Clin. Pathol. 6, 467–474
23. Misago, N., Mori, T., and Narisawa, Y. (2010) Nestin expression in stromal cells of trichoblastoma and basal cell carcinoma. J. Eur. Acad. Dermatol. Venereol. 24, 1354–1358
24. Akiyama, M., Matsuda, Y., Ishiwata, T., Naito, Z., and Kawana, S. (2013) Nestin is highly expressed in advanced-stage melanomas and neurotized nevi. Oncol. Rep. 29, 1595–1599
25. Akiyama, M., Matsuda, Y., Ishiwata, T., Naito, Z., and Kawana, S. (2013) Inhibition of the stem cell marker nestin reduces tumor growth and invasion of malignant melanoma. J. Investigative Dermatol. 133, 1384–1387
26. Brychtova, S., Fiuraskova, M., Hlobilkova, A., Brychta, T., and Hirnak, J. (2007) Nestin expression in cutaneous melanomas and melanocytic nevi. J. Cutan. Pathol. 34, 370–375
27. Piras, F., Perra, M. T., Murtas, D., Minerba, L., Floris, C., Maxia, C., Demurtas, P., Ugalde, J., Ribatti, D., and Sirigu, P. (2010) The stem cell marker nestin predicts poor prognosis in human melanoma. The stem cell marker nestin predicts poor prognosis in human melanoma. Oncol. Reports 23, 17–24
28. Klein, W. M., Wu, B. P., Zhao, S., Wu, H., Klein-szanto, A., JP, and Tahan, S. R. (2007) Increased expression of stem cell markers in malignant melanoma. Mod. Pathol. 133, 102–107
29. Qendro, V., Lundgren, D. H., Rezaul, K., Mahony, F., Ferrell, N., Bi, A., Latifi, A., Chowdhury, D., Gygi, S., Haas, W., Wilson, L., Murphy, M., and Han, D. K. (2014) Large-scale proteomic characterization of melanoma expressed proteins reveals nestin and vimentin as biomarkers that can potentially distinguish melanoma subtypes. J. Proteome Res. 13, 5031–5040
30. Lee, C.-W., Zhan, Q., Frank, M., Huang, J., Larson, A., Jie, S., and Xiao, D. (2015) Nestin depletion induces melanoma matrix metalloproteinases and invasion. Lab. Invest. 94, 1382–1395
31. Hyder, C. L., Lazaro, G., Pylva, J. W., Qvarnstro, S. M., and Eriksson, J. E. (2014) Nestin regulates prostate cancer cell invasion by influencing the localisation and functions of FAK and integrins. J. Cell Sci. 127, 2161–2173
32. Doxie, D. B., Greenplate, A. R., Gandelman, J. S., Diggins, K. E., Roe, C. E., Dahlman, K. B., Sosman, J. A., Kelley, M. C., and Irish, J. M. (2018) BRAF and MEK inhibitor therapy eliminates Nestin-expressing melanoma cells in human tumors. Pigment Cell Melanoma Res. 31, 708–719
33. Sinnberg, T., Makino, E., Krueger, M. A., Velic, A., Macek, B., Rothbauer, U., Groll, N., Potz, O., Czemmel, S., Niessner, H., Meier, F., Ikenberg, K., Garbe, C., and Schittek, B. (2016) A Nexus Consisting of Beta-Catenin and Stat3 Attenuates BRAF Inhibitor Efficacy and Mediates Acquired Resistance to Vemurafenib. EBioMedicine 8, 132–149
34. Herlyn, D., Iliopoulos, D., Jensen, P. J., Parmiter, A., Baird, J., Hotta, H., Adachi, K., Ross, A. H., Jambrosic, J., Koprowski, H., and Herlyn, M. (1990) In Vitro properties of human melanoma cells metastatic in nude mice. Cancer Res. 50, 2296–2302
35. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat. Protoc. 4, 484–494
36. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature Protocols 2, 1896–1906
37. Batth, T. S., and Olsen, J. V. (2016) Offline high pH reversed-phase peptide fractionation for deep phosphoproteome coverage. In: von Stechow, L., ed. Phospho-Proteomics: Methods Protocols , pp. 179–192, Springer New York, New York
38. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnol. 26, 1367–1372
39. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. 10, 1794–1805
40. Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N. T., Roth, A., Bork, P., Jensen, L. J., and von Mering, C. (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 45, D362–D368
41. Long, G. V., Fung, C., Menzies, A. M., Pupo, G. M., Carlino, M. S., Hyman, J., Shahheydari, H., Tembe, V., Thompson, J. F., Saw, R. P., Howle, J., Hayward, N. K., Johansson, P., Scolyer, R. A., Kefford, R. F., and Rizos, H. (2014) Increased MAPK reactivation in early resistance to dabrafenib/trametinib combination therapy of BRAF-mutant metastatic melanoma. Nat. Commun. 5, 5694
42. Rizos, H., Menzies, A. M., Pupo, G. M., Carlino, M. S., Fung, C., Hyman, J., Haydu, L. E., Mijatov, B., Becker, T. M., Boyd, S. C., Howle, J., Saw, R., Thompson, J. F., Kefford, R. F., Scolyer, R. A., and Long, G. V. (2014) BRAF inhibitor resistance mechanisms in metastatic melanoma: spectrum and clinical impact. Clin. Cancer. Res. 20, 1965–1977
43. RCore Team. (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
44. Dunning, M., Lynch, A., and Eldridge, M. (2015) illuminaHumanv4.db: Illumina HumanHT12v4 annotation data (chip illuminaHumanv4). R package version 1.26.0 Ed.
45. Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., and Smyth, G. K. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47
46. Ran, F. A., Hsu, P. D. P., Wright, J., Agarwala, V., Scott, D. A., Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. Nat. Protocols 8, 2281–2308
47. Bae, S., Park, J., and Kim, J.-S. (2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473–1475
48. Turečková, J., Vojtěchová, M., Krausová, M., S ˇloncová, E, Korínek, V. (2009) Focal adhesion kinase functions as an Akt downstream target in migration of colorectal cancer cells. Translational Oncol. 2, 281–290
49. Zhao, X., and Guan, J. L. (2011) Focal adhesion kinase and its signaling pathways in cell migration and angiogenesis. Adv. Drug Deliv. Rev. 63, 610–615
50. Kim, M. H., Kim, J., Hong, H., Lee, S. H., Lee, J. K., Jung, E., and Kim, J. (2016) Actin remodeling confers BRAF inhibitor resistance to melanoma cells through YAP/TAZ activation. EMBO J. 35, 462–478
51. Narita, K., Matsuda, Y., Seike, M., Naito, Z., Gemma, A., and Ishiwata, T. (2014) Nestin regulates proliferation, migration, invasion and stemness of lung adenocarcinoma. Int. J. Oncol. 44, 1118–1130

52. Chernoivanenko, I. S., Minin, A. A., and Minin, A. A. (2013) [Role of vimentin in cell migration]. Ontogenez 44, 186–202
53. Bakos, R. M., Maier, T., Besch, R., Mestel, D. S., Ruzicka, T., Sturm, R. A., and Berking, C. (2010) Nestin and SOX9 and SOX10 transcription factors are coexpressed in melanoma. Exp. Dermatol. 19, e89–e94
54. Monzani, E., Facchetti, F., Galmozzi, E., Corsini, E., Benetti, A., Cavazzin, C., Gritti, A., Piccinini, A., Porro, D., Santinami, M., Invernici, G., Parati, E., Alessandri, G., and La Porta, C. A. M. (2007) Melanoma contains CD133 and ABCG2 positive cells with enhanced tumourigenic potential. Eur. J. Cancer 43, 935–946
55. Luo, Y., Dallaglio, K., Chen, Y., Robinson, W. A., Robinson, S. E., McCarter, M. D., Wang, J., Gonzalez, R., Thompson, D. C., Norris, D. A., Roop, D. R., Vasiliou, V., and Fujita, M. (2012) ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. Stem Cells 30, 2100–2113
56. Griffin, M., Scotto, D., Josephs, D. H., Mele, S., Crescioli, S., Bax, H. J., Pellizzari, G., Wynne, M. D., Nakamura, M., Hoffmann, R. M., Ilieva, K. M., Cheung, A., Spicer, J. F., Papa, S., Lacy, K. E., Karagiannis, S. N. (2017) BRAF inhibitors: resistance and the promise of combination treatments for melanoma. Oncotarget 8, 78174–75192
57. Ludwig, T. (2005) Local proteolytic activity in tumor cell invasion and metastasis. Bioessays 27, 1181–1191
58. Moreno-Layseca, P., and Streuli, C. H. (2014) Signalling pathways linking integrins with cell cycle progression. Matrix Biol. 34, 144–153
59. Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.-W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 44, D447–D456

Supplementary Information

Schmitt, M., Sinnberg, T., Nalpas, N. C., 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 18, 1096-1109

doi.org/10.1074/mcp.RA119.001302

Table I1: List of all used patient derived melanoma cell lines in this study. ATCC: American Type Culture Collection

Name	Tissue	Morphology	Source	BRAF mutation
A375	skin	epithelial	ATCC (CRL-1619)	BRAF V600E
Mel1617	skin	epithelial	M. Herlyn [1]	BRAF V600E
451lu	skin	epithelial	M. Herlyn [1]	BRAF V600E
SKMel28	skin	polygonal	ATCC (HTB-72)	BRAF V600E
SKMel19	skin	epithelial	C. Garbe [2]	BRAF V600E

1 Herlyn, D. *et al.* Properties of Human Melanoma Cells Metastatic in Nude Mice. *Cancer Research* **50**, 2296-2302 (1990).

2 Sinnberg, T. *et al.* A Nexus Consisting of Beta-Catenin and Stat3 Attenuates BRAF Inhibitor Efficacy and Mediates Acquired Resistance to Vemurafenib. *EBioMedicine* **8**, 132-149, doi:10.1016/j.ebiom.2016.04.037 (2016).

Table I2: Top five down-and up-regulated proteins of vermurafenib-resistant and -sensitive A375 proteome.

Uniprot ID	Gene name	Protein name	SILAC ratio (log_2)	Intensity (log_{10})
P02751	FN1	Fibronectin	-2.81	10.22
P48681	NES	Nestin	-2.71	10.20
P35625	TIMP3	Metalloproteinase inhibitor 3	-1.92	9.44
Q13938	CAPS	Calcyphosin	-1.79	9.58
P04233	CD74	HLA class II histocompatibility antigen gamma chain	-1.38	9.96
A0A0G2JJ56	SPANXB1	Sperm protein associated with nucleus on X chromosome B/F	2.10	8.29
Q9GZP8	IMUP	Immortalization up-regulated protein	2.01	7.69
P80723	BASP1	Brain acid soluble protein 1	2.00	8.21
P00533	EGFR	Epidermal growth factor receptor	1.65	7.83
Q52LW3	ARHGAP29	Rho GTPase-activating protein 29	1.34	7.61

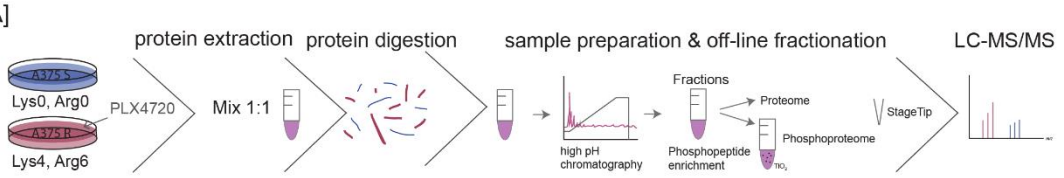
Table I3: List of FFPE specimens used in this study.

Patient	Surgery date	Localization	Treatment	Remarks
#1	09.11.2011	Abdomen		Immunohistochemistry
	05.06.2014	Buttocks (gluteal, left)	Vemurafenib (12 months)/ Ipilimumab (3 months)	Immunohistochemistry
#2	23.11. 2011	Tight (left)		Immunohistochemistry and proteomics
	25.02.2013	Lower leg (right)	Vemurafenib (3 months)	Immunohistochemistry and proteomics

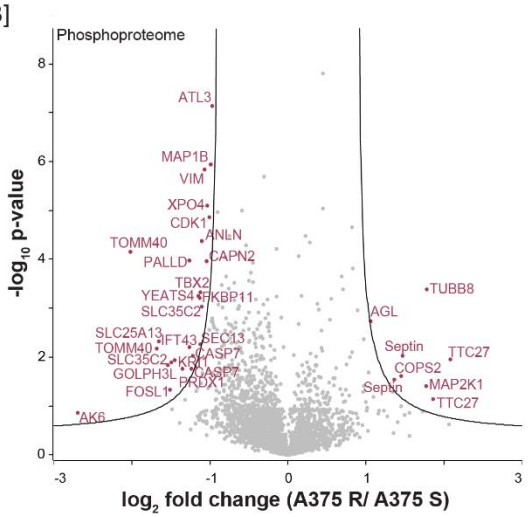
Table I4: Predicted off-target effects of crRNA guide CCTCGACGGCGCGCCGGTTG using Cas-offinder. Chr: chromosome.

crRNA	DNA	Chr	Position	Direction	mismatch
CCTCGACGGCGCGCC GGTTGNGG	CCTCGACGGCGCGC CGGTTGCGG	chr1	1570000	+	0
CCTCGACGGCGCGCC GGTTGNGG	CCcCGACGGCGCGgC GGTTcCGG	chr17	7478944	+	3
CCTCGACGGCGCGCC GGTTGNGG	CCcCaACGGCGCGCC GGcTGTGG	chr20	62861781	-	3

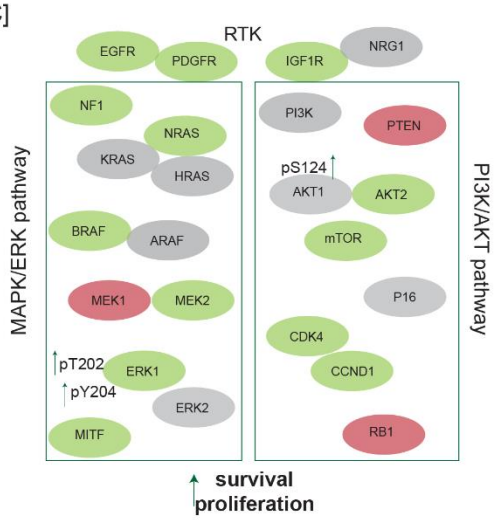
[S1A]



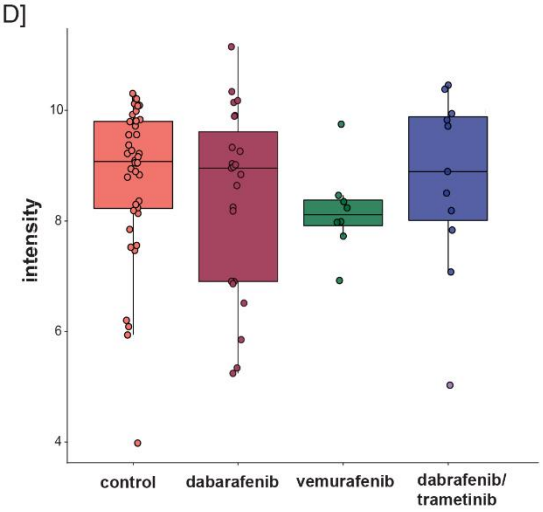
[S1B]



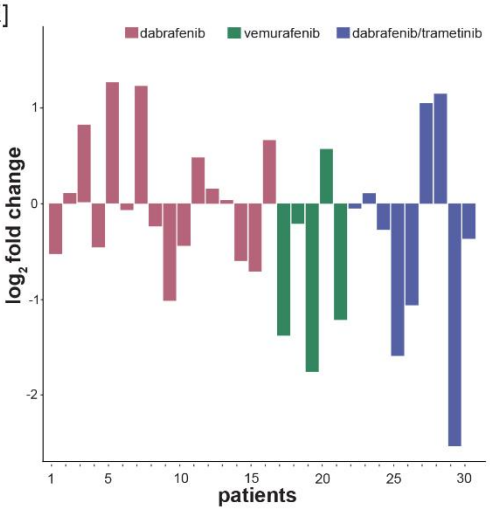
[S1C]



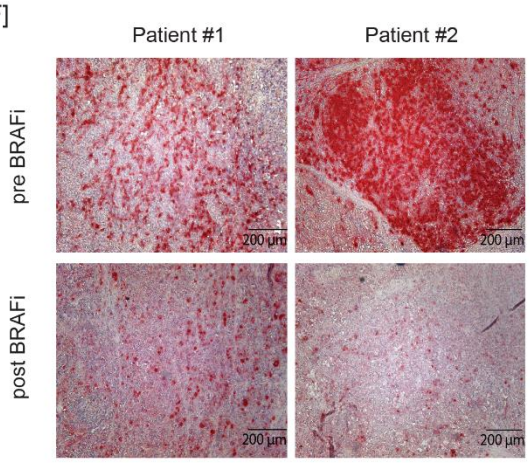
[S1D]



[S1E]



[S1F]



[S1G]

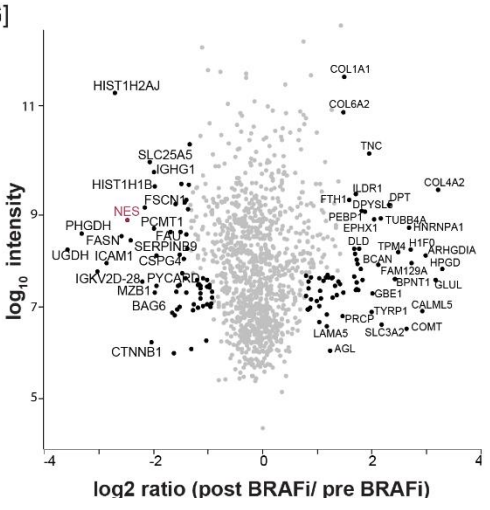
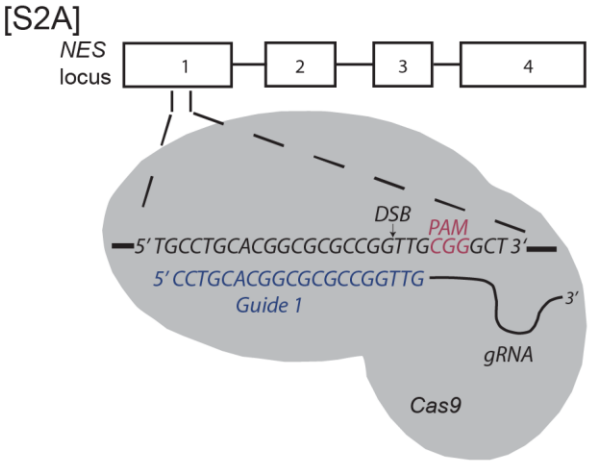
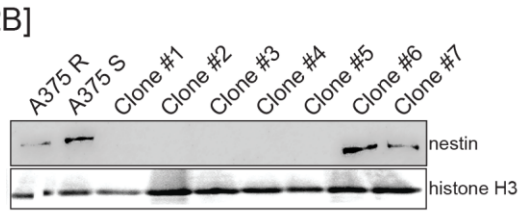
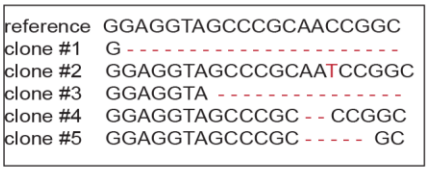
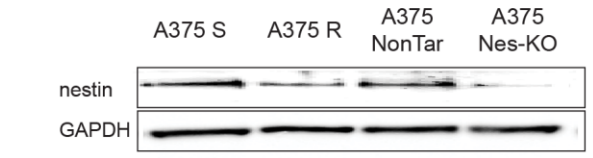


Figure S1: Quantitative proteome and phosphoproteome analysis of human melanoma cells identifies down-regulation of nestin. [A] Schematic of sample preparation workflow. Vemurafenib resistant A375 (A375 R) and A375 sensitive cells (A375 S) were 'light' (Lys0/Arg0) and 'medium' (Lys4/Arg6) SILAC labelled. After cell lysis, crude protein extracts were mixed 1:1, reduced, alkylated and trypsin digested. The resulting peptides mixture was fractionated using an off-line HPLC operated with high pH buffers. Fractions were pooled and measured directly (proteome) or applied to phosphopeptide enrichment using titanium dioxide (TiO_2) prior to liquid chromatography mass spectrometry analysis (LC-MS/MS). [B] Volcano plot of vemurafenib-resistant and -sensitive A375 proteomes for phosphoproteome. t-test difference of SILAC ratios between A375 R and A375 S (log_2) are plotted against p-value ($-log_{10}$) (n=3). Black lines indicate the significance threshold (FDR < 0.01; s_0 = 1). Significantly up- and down-regulated proteins are highlighted in magenta. [C] Identified key molecules and phosphorylation sites of the MAPK/ERK and PI3K/AKT signaling pathway. Green: up-regulated in A375 R vs. A375 S cells; red: down-regulated in A375 R vs. A375 S cells; grey: identified, but not quantified; arrows: up-regulated phosphorylation sites in A375 R vs. A375 S. [D] NES expression profile in human patients with melanoma metastases in vemurafenib, dabrafenib and dabrafenib plus trametinib treated tumors and pre-treatment control tumors (FDR ≤ 0.1). [E] mRNA expression levels of nestin protein in thirty patients with melanoma metastases after BRAF inhibitor therapy compared to control tumors. [F] Immunohistochemical staining for nestin of melanoma metastases obtained before treatment with a BRAF inhibitor vermurafenib and after resistance acquisition for two patients. Nestin levels are shown in red (Fast red substrate). [G] Proteomics of FFPE specimens pre-and post-BRAF inhibitor therapy using quantitative proteomics based dimethyl-labelling. Ratios (log_2) of post-BRAF vs. pre-BRAF inhibitor therapy are plotted against intensity (log_{10}) (p-value <0.05). Nestin is highlighted in magenta.









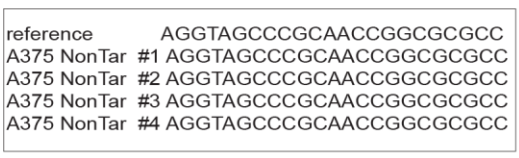


Figure S2: Nestin expression correlates with invasive properties in melanoma cell lines. [A] Schematic overview of the establishment of NES knockout cells using CRISPR/Cas9 genome editing system. Blue: guide sequence targeting Exon1 in the genomic sequence of NES; red: protospacer adjacent sequence (PAM) sequence; DSB: double strand break. [B] Western blot analysis of A375 S, A375 R and CRISPR/Cas9 genome edited cell clones Nes-KO #1-7. [C] Sanger sequencing result of reference DNA (A375 S and R) and CRISPR/Cas9 genome-edited cell clones Nes-KO #1 - 5. [D] Amino acid sequence of human nestin from Uniprot database. Grey: peptide sequences identified by LC-MS/MS. [E] Amino acid sequence of CRISPR/Cas9 genome edited cell clones Nes-KO #1-5. Grey: peptide sequences identified by LC-MS/MS. Red: truncated amino acid sequence compared to A375 S. [F] Western blot analysis of nestin and GAPDH protein in A375 S, A375 R, A375 NonTar and A375 Nes-KO cells. [G] Sanger sequencing results of reference DNA (A375 S and R) and CRISPR/Cas9 genome-edited control cell clones A375 NonTar #1 – 4.

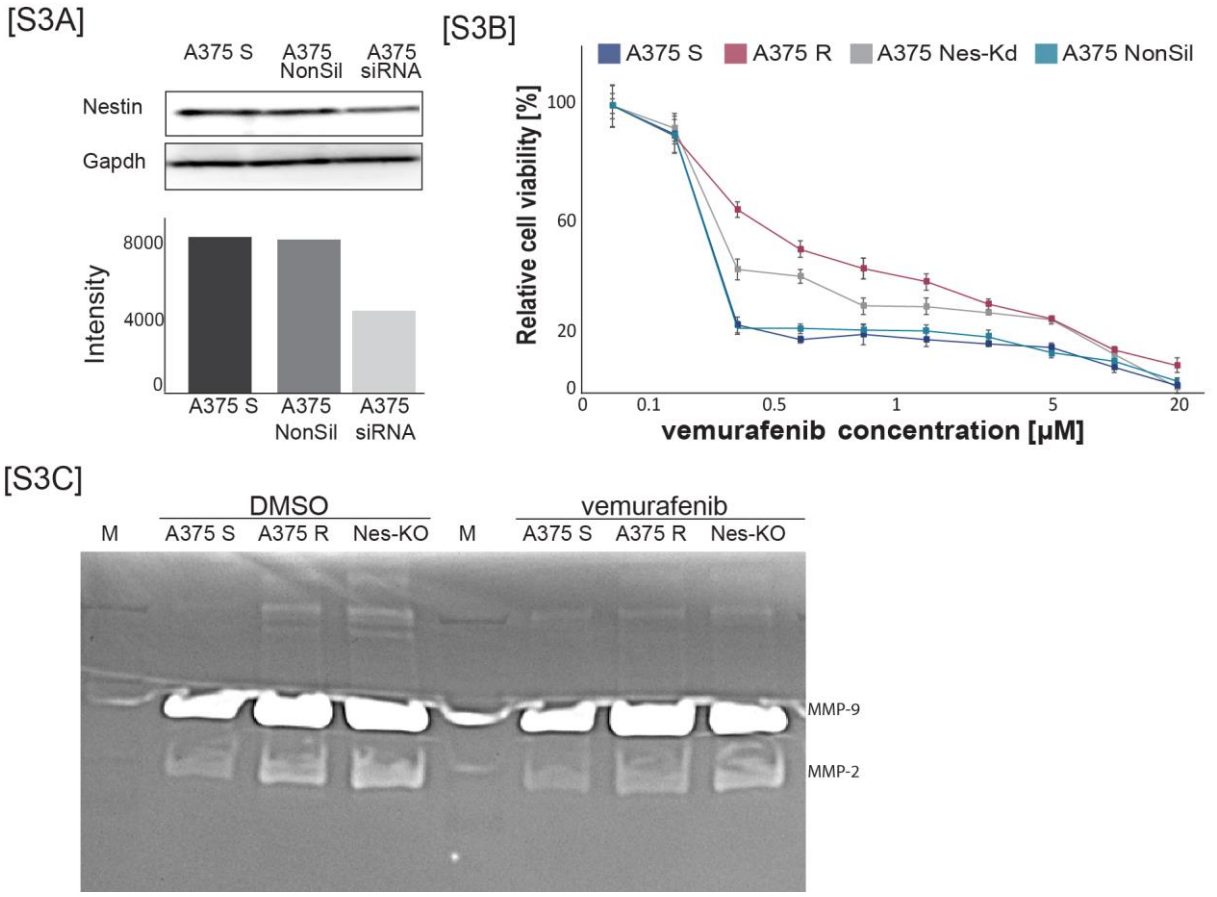
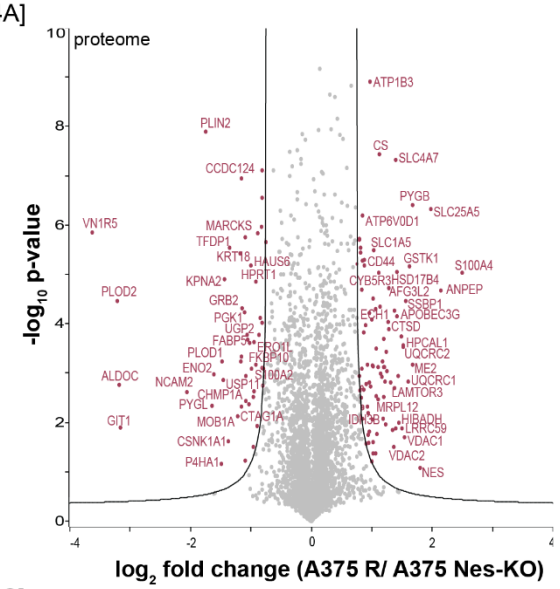
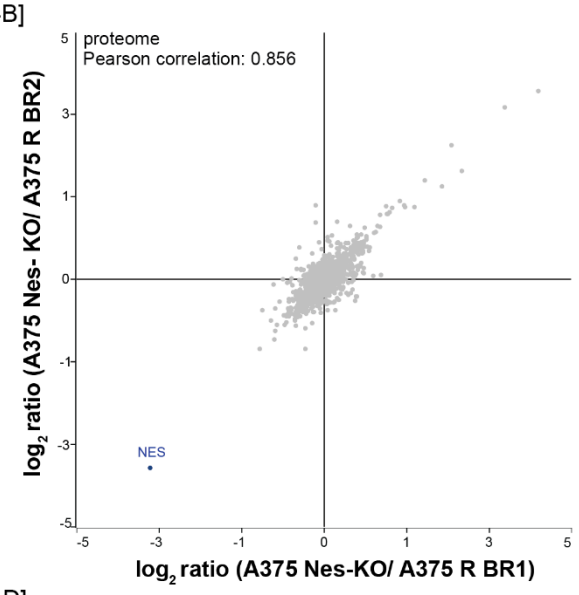


Figure S3: Depletion of nestin affects cell proliferation and colony formation upon treatment with signaling pathway inhibitors. [A] Western blot analysis of A375 S, A375 NonSil and A375 siRNA against nestin and quantification of bands intensities using ImageJ software. Nestin was down-regulated in A375 S cells by transfection of a pool of four siRNA oligos (siRNA) against human nestin. Untreated A375 S and NonSilencing siRNA (NonSil) treated A375 S cells were included as control. Cells were harvested 48 h post-transfection. [B] A375 S, A375 R, A375 NonSil and A375 Nes-Kd were cultured for 24 h, and then treated with PLX4720 at the indicated concentrations (0, 0.1, 0.25, 0.5, 1, 2.5, 5, 10 and 20 µM) or DMSO as control, respectively. Cell viability was determined by MTS assay 96 h later. Results expressed as % control represent the mean of three biological experiments (n=24). Error bar represents standard deviations of three biological replicates [C] Gelatine zymography of supernatants of A375 S, A375 R and A375 Nes-KO cell lines treated with DMSO or PLX4720 for 24 h. Image is a representative of three independent experiments.

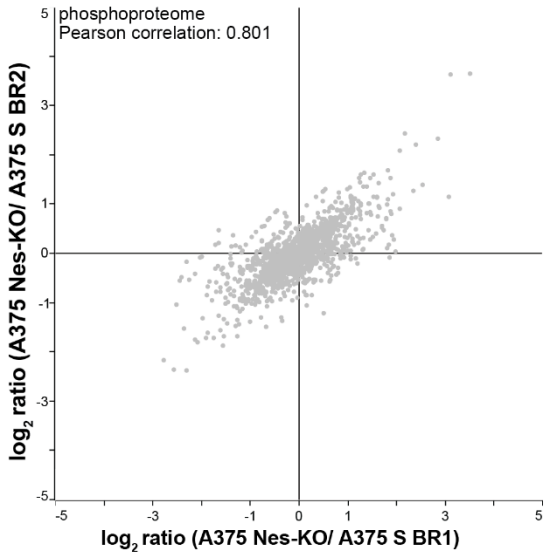
[S4A]



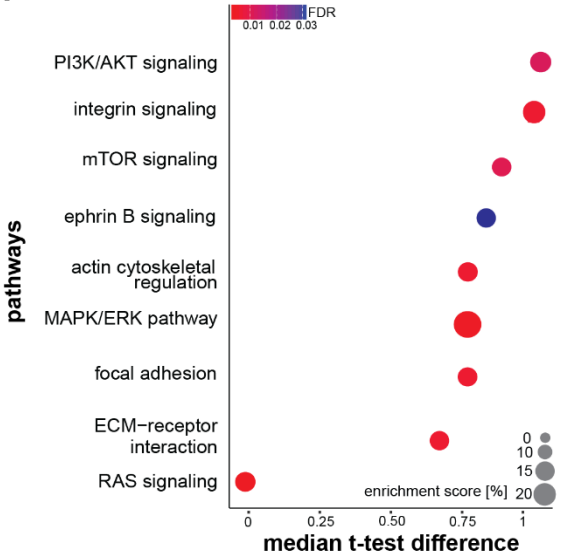
[S4B]



[S4C]



[S4D]



[S4E]

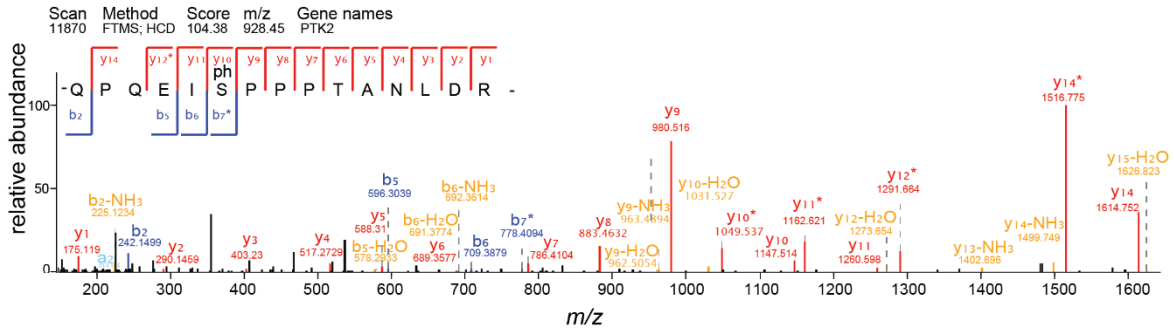


Figure S4: Quantitative proteomics comparison between nestin knockout and BRAF inhibitor sensitive and resistant cell lines. [A] Volcano plot of A375 Nes-KO and A375 R proteomes. t-test difference of SILAC Ratios between A375 Nes-KO and A375 R (log_2) are plotted against p-value ($-log_{10}$) (n=3). Black lines indicate the significance threshold (FDR < 0.01; s_0 = 1). Significantly up- and downregulated proteins are highlighted in magenta. [B] Proteome correlation of A375 Nes-KO relative to A375 R in biological replicate 1 and 2. [C] Phosphoproteome correlation of A375 Nes-KO relative to A375 R in biological replicate 1 and 2. [D] Over-representation of selected signaling KEGG pathways of A375 Nes-KO compared to A375 S cells using String database analysis. The t-test difference of SILAC ratios between A375 Nes-KO and A375 S (log_2) were plotted for each pathway (t-test, FDR < 0.1; s_0 =1). Enrichment score [%] identified significantly changing proteins mapped to the pathway by the total protein count involved in that pathway. Colour of the dots represents the FDR. [E] Annotated spectra of phosphorylated peptide LQPQEIpSPPPTANLDR containing a phosphosite at S910 on focal adhesion kinase FAK.

3.2 Manuscript II

Schmitt, M., Sinnberg, T., Bratl, K., Garbe, C., Macek, B., Nalpas N. C.

Integration of individualized proteogenomics datasets to analyse single amino acid variants in malignant melanoma

Under revision in Frontiers in Oncology, 2020

Author contributions:

The study was designed by Tobias Sinnberg, Nicolas C. Nalpas, Boris Macek and me and all experiments were performed at the Proteome Center Tuebingen, University of Tuebingen. Cell culture, sample preparation, proteomics and phosphoproteomics analysis and interaction assays were performed by me under the supervision of Boris Macek. Katrin Bratl performed two replicates of pulldown experiments under my supervision. The data was analyzed and interpreted by Nicolas C. Nalpas and me under the supervision of Boris Macek. Claus Garbe contributed in fruitful discussion of the manuscript.

A list of all used materials and extended method description can be found as Supplementary Information on the provided CD.

References of the manuscript are included in the main references.

Integration of Individualised Proteogenomics Datasets to Analyse Single 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, 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 single nucleotide variants, genomic insertions, deletions and structural variants is a cornerstone of personalised medicine. Although only about 2% of the genomic sequence is protein-coding, mutations occurring in these regions have potential to influence protein structure and can therefore have severe impact on the aetiology of many diseases. Of special interest are mutations that affect modifiable amino acid residues, as protein modifications involved in signal transduction networks cannot be analysed by genomics. Proteogenomics addresses this impact by analysing proteomes in context of patient- or tissue-specific non-synonymous single nucleotide variants (nsSNVs), insertions and deletions. Here we present a bioinformatics application termed Proteogenomics Characterisation Tools (PCTi) that enables straightforward integration of nucleotide variants into protein databases, assessment of their potential impact and subsequent visualisation of proteogenomics data. We apply PCTi to analyse the non-synonymous mutational landscape of two frequently used malignant melanoma cell lines (A375 and SkMel28) in context of resistance to commonly used BRAF inhibitor vemurafenib. We detect a disproportional impact of nucleotide variants on modifiable residues between sensitive and resistant cell lines. Approximately 35% of protein variants in both cell lines interfere with the modification status and potentially influence signal transduction networks. MS measurements confirmed mutation-driven modification changes in over 50 proteins; among these was the transcription factor RUNX1 mutated on S276L. We confirm the loss of the Ser276 phosphorylation site by MS and demonstrate the impact of this mutation on the interactome of RUNX1.

Introduction

The past decade has seen a revolution in high-throughput sequencing technologies, which provide information on DNA/RNA sequence, gene structure and expression [375]. 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 [247,248,376] and post-translational modifications (PTMs) such as phosphorylation [310,377-379]. 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 mutations. 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 [340,362,380]. However, the integration of such omics datasets is usually not straightforward and requires advanced computational skills. In this context, a number of bioinformatic tools are available to the community for proteogenomics approaches. For instance, customised protein sequence databases can be generated from NGS data using costumProDB, Galaxy-P, PPline or PGA software [351-354]. In addition, several tools compare PTM sites with mutations, such as PhosSNP, PTMvar and ReKINect [340,381,382]. Furthermore, the concurrent visualisation of omics datasets is offered by VisANT, NetGestalt and Galaxy softwares [383-385]. Currently, there is a need for workflows that combine various aspects of a proteogenomics approach in a user-friendly fashion. Proteogenomics has a potential to precisely characterise mutation-driven alterations of signal transduction pathways during tumourigenesis [386]. 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 [135]. 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 [100]. The RAS/BRAF/MEK/ERK kinase pathway is mutated to an oncogenic form in 30% of all cancers, with non-synonymous somatic missense mutations in BRAF up to 50% of cutaneous melanomas [32]. The predominant BRAF mutation is within the kinase domain with a single nucleotide substitution of valine to glutamic acid at amino acid 600 [31]. 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 [32,33]. Targeted inhibition of the mutated BRAF kinase with selective inhibitors like vemurafenib, dabrafenib or encorafenib (BRAFi) results in a reduction of MAPK pathway signalling [33]. However, almost all patients rapidly develop resistance to BRAFi monotherapy after a period of approximately five months [100,102,103]. 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, mutations 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 mutations at the proteome and PTM level are still largely unknown.

Here we present a new software termed Proteogenomics Characterisation Tools (PCTi), which streamlines multiple aspects of a proteogenomics analysis, from generation of custom protein sequence databases to visualisation of integrated proteogenomics data. We applied PCTi to analyse two immortalised human cell lines commonly used in melanoma research, A375 and SkMel28, in their parental as well as in their BRAFi resistant state. PCTi was able to reconstruct signal transduction networks specific to individual cell lines and phenotypes using their matching genomics and (phospho)proteomics datasets.

Results

To study the impact of single amino acids variants on signal transduction networks, we selected two widely exploited melanoma cells lines harbouring the $BRAF^{V600E}$ mutation, A375 and SkMel28. The cell lines were established with two different phenotypes, drug-sensitive ("S" phenotype) and drug-resistant ("R" phenotype) against the BRAF inhibitor vemurafenib, as described previously [387]. Both cell lines were subjected to exome sequencing as well as proteomics and phosphoproteomics analysis using high-resolution mass spectrometry (**Figure S1A**). To integrate and analyse genomic and proteomic datasets, we developed the software PCTi. This application is coded entirely in the R programming language and provides a user-friendly graphical interface via the Shiny package [388]. The software comprises four independent modules, which are used to (1) incorporate non-synonymous nucleotide variants into a protein sequence, (2) stratify variants according to their biological impact based on user specification, (3) integrate the nucleotide and amino acid variants identification (WES and MS), and (4) visualise the analysed proteogenomics datasets (**Figure 1**). For more details about PCTi please refer to Material and Methods section.

Application of PCTi to BRAFi-sensitive and -resistant melanoma cell lines

As a proof of principle, we used the PCTi software to characterise two melanoma cell lines, A375 and SkMel28, both presenting drug-sensitive (S) and drug-resistant (R) phenotypes to BRAF inhibitor vemurafenib. In this context, we performed Whole Exome Sequencing (WES) and deep MS-based (phospho)proteomics to get information on the mutational landscape, as well as abundance of proteins and phosphorylation sites. Number and type of nucleotide

variants detected in the WES analysis was similar between the cell lines and is summarised in **Figure S1B and C**. Each non-synonymous mutation was incorporated using the PCTincorporate module into the corresponding protein isoforms, generating several thousand additional protein sequences in respective protein databases. Despite this large increase in the number of protein isoform sequences, the database search space increased by only 2-3% (**Figure S1D and E**).

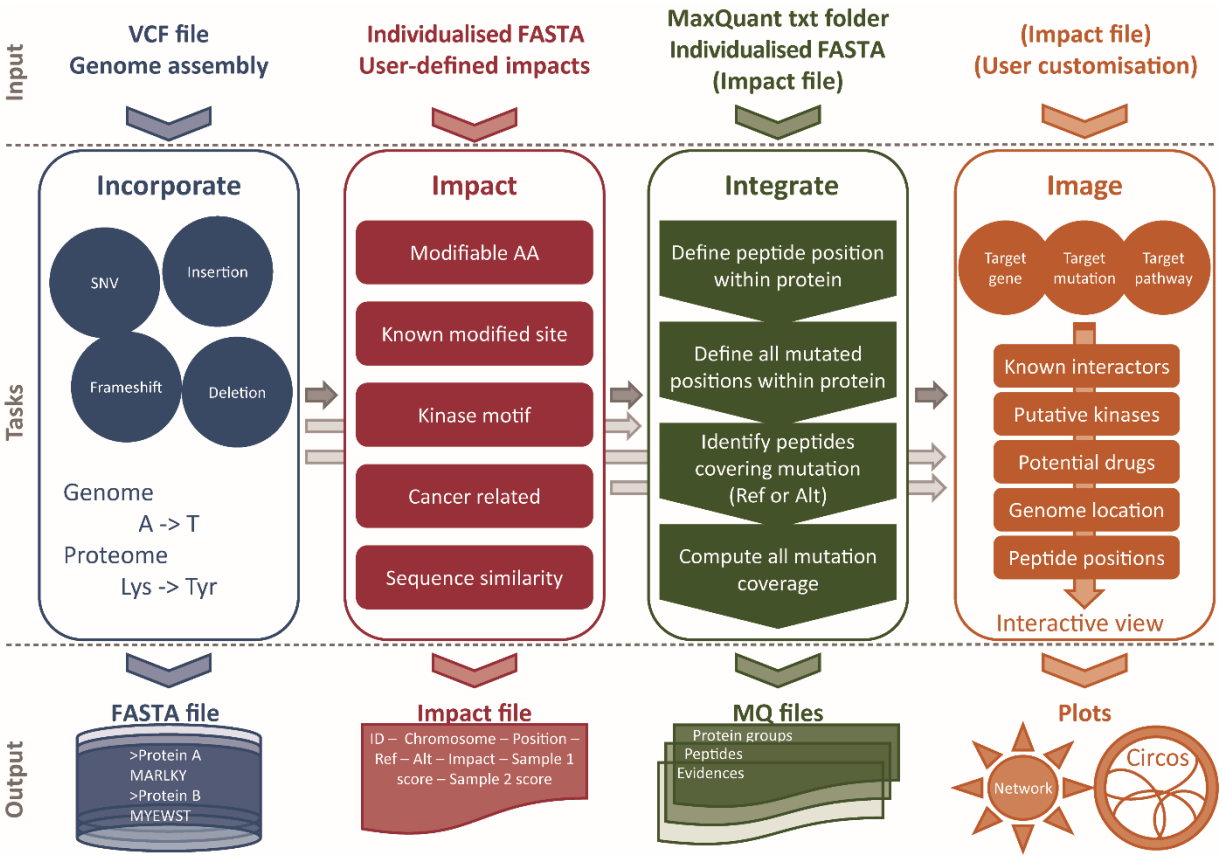


Figure 1: PCTi application provides a reactive environment for the integration of genomics with proteomics. Schematic representation of the PCTi software. The first module of PCTi allows the incorporation of a list of mutations stored in a VCF file directly into the corresponding genomic assembly. The resulting FASTA file contains protein sequences with non-synonymous mutations and is ready to use for LC-MS/MS data processing. The second module calculates the impact of the non-synonymous mutations based on user-specified parameters centered on disease type and modification status. The resulting impact scores help the prioritisation of relevant mutated proteins. The third module integrates WES and MS by identifying reference and alternate peptides resulting from a LC-MS/MS processing. It also determines the reference or alternate peptides, which contain post-translational modifications. The final module generates images of the proteogenomics integration, such as protein-based mutation/peptide coverage or proteogenomics-based interaction network.

We then used the PCTimpact module to evaluate the effect of mutated protein residues on protein phosphorylation status and on melanoma aetiology. In this context, approximately 35% of mutated protein isoforms had a loss or gain of S/T/Y modifiable amino acids and/or kinase-

substrate motifs in both cell lines. Among these, a loss of known phosphorylation sites was observed in more than one thousand protein isoforms. In addition, more than 20% of the mutated protein residues were previously reported in melanoma samples and/or present on protein encoded by known oncogene or tumour suppressor genes. In both cell lines, the impact score distribution showed that only a minority of mutated protein isoforms have a medium to high impact score (with 0 = no impact; 1 = high impact), for example the point mutations G691S on Ret proto-oncogene protein, S276L on RUNX family transcription factor 1 and a frameshift mutation V514WfsX139 on FLII actin remodelling protein (**Figure S1F and G)**.The individualised databases, discussed above, were used for the processing of deep proteomics and phosphoproteomics data from A375 and SkMel28. Each cell line was analysed separately and peptides were separated into fractions using high pH reverse phase chromatography. High resolution MS identified more than 9,300 protein groups and over 130,000 sequence-unique peptides in each cell line (**Figure S1H and I**). Interestingly, the identified coverage on the proteome level was similar, whereas the coverage of identified non-redundant phosphorylation sites strongly differed between cell lines (**Figure S1H and I**). This difference is unlikely caused by technical reasons and reflects previously reported differences [389]. In both cell lines, most of the phosphopeptides were exclusively identified in resistant cells, suggesting high relevance of PTMs in resistance to BRAF inhibition.

BRAFi-sensitive and -resistant SkMel28 cells have distinct mutational landscapes

We next compared the BRAFi-sensitive versus -resistant phenotypes for A375 and SkMel28 cell lines on the basis of the WES-identified nucleotide variants. The comparison between A375 S and R revealed almost identical number of non-synonymous nucleotide variants, as well as a very high overlap (94%) of mutations (**Figure 2A**). Although a similar number of non-synonymous nucleotide variants was identified in SkMel28, the overlap between S and R phenotypes was only 39.7% (**Figure 2B**). Comparison of the sequencing depth revealed similar results across phenotypes and cell lines, with median depth ranging from 90 to 118 reads (**Figure S2A and B**). As the same tools were used for variant calling, the difference in the overlap between S and R phenotypes in SkMel28 cells is unlikely to be a technical artefact. Variants were then 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 mutations represented approximately 65% of the total nucleotide changes and were consistent across phenotypes and cell lines (**Figure S2C and D**).

Individualised proteogenomics highlights key differences between BRAFi-sensitive and -resistant melanoma cell lines

To address hotspot mutations accumulating in specific pathways, we performed pathway enrichment of identified proteins harbouring amino acid variants. For both cell lines, we detected multiple alternate peptides and phosphopeptides including peptides harbouring single amino acid variants, InDels and frameshift mutations. Most of the identified mutations were detected in drug-resistant cells compared to drug-sensitive cells on the proteome and phosphoproteome level (**Figure 2C and D**). Next, we performed pathway enrichment of proteins containing alternate peptides and showed significant differences between cell lines and their phenotypes (**Figure 2E**). For both cell lines, MAPK signalling pathway and ErbB signalling pathway were over-represented for mutated proteins in drug-resistant cells, whereas RHO GTPase cycle was enriched for both drug-sensitive cells. In contrast, transcriptional regulation by RUNX1 was only over-represented in A375 R cells which is linked to transcriptional misregulation in cancer. Metabolism, HIPPO signalling pathway and lysosome were enriched for variant proteins in A375 S cells. The annotation enrichment results for SkMel28 S and R showed that most of the pathways were similarly over-represented for both phenotypes. In addition, we compared the proteomic and phosphoproteomic differences between both drug-sensitive and resistant cells for both cell lines and identified several significantly regulated proteins (**Figure S2E and F**). MS analysis revealed the identification of numerous significantly regulated proteins and phosphorylated proteins between cell line sand their phenotypes, including key signalling proteins like BRAF, ERK1, AKT. To characterise phenotypes between drug-sensitive and drug-resistant cells, we conducted an enrichment analysis including Gene Ontology (GO), KEGG pathway, and Reactome analysis in order to assess over-represented pathways and biological processes between significantly regulated proteins of shared identifications and unique proteins for each phenotype (**Figure S2G and H**). Based on the enrichment results for Reactome, signalling pathways like MAPK were enriched in both resistant cell lines and PI3K-AKT-mTOR pathway in A375 R and Wnt signalling in SkMel28 R cells, known to be highly activated in resistant melanoma cells. YAP/TAZ-stimulated gene expression, rRNA processing and nucleosome assembly were identified in both sensitive cells. For shared proteins, extracellular matrix organisation was over-represented in both sensitive cells and Rho GTPase activation and AURKA activation by TPX2 in A375 R cells, whereas signalling by activin, NCAM1 signalling were significantly enriched in SkMel28 R cells (**Figure S2G and H**). These findings highlighted the shared properties and differences

between melanoma cell lines and showed that alterations can be identified at the peptide and phosphopeptide level.

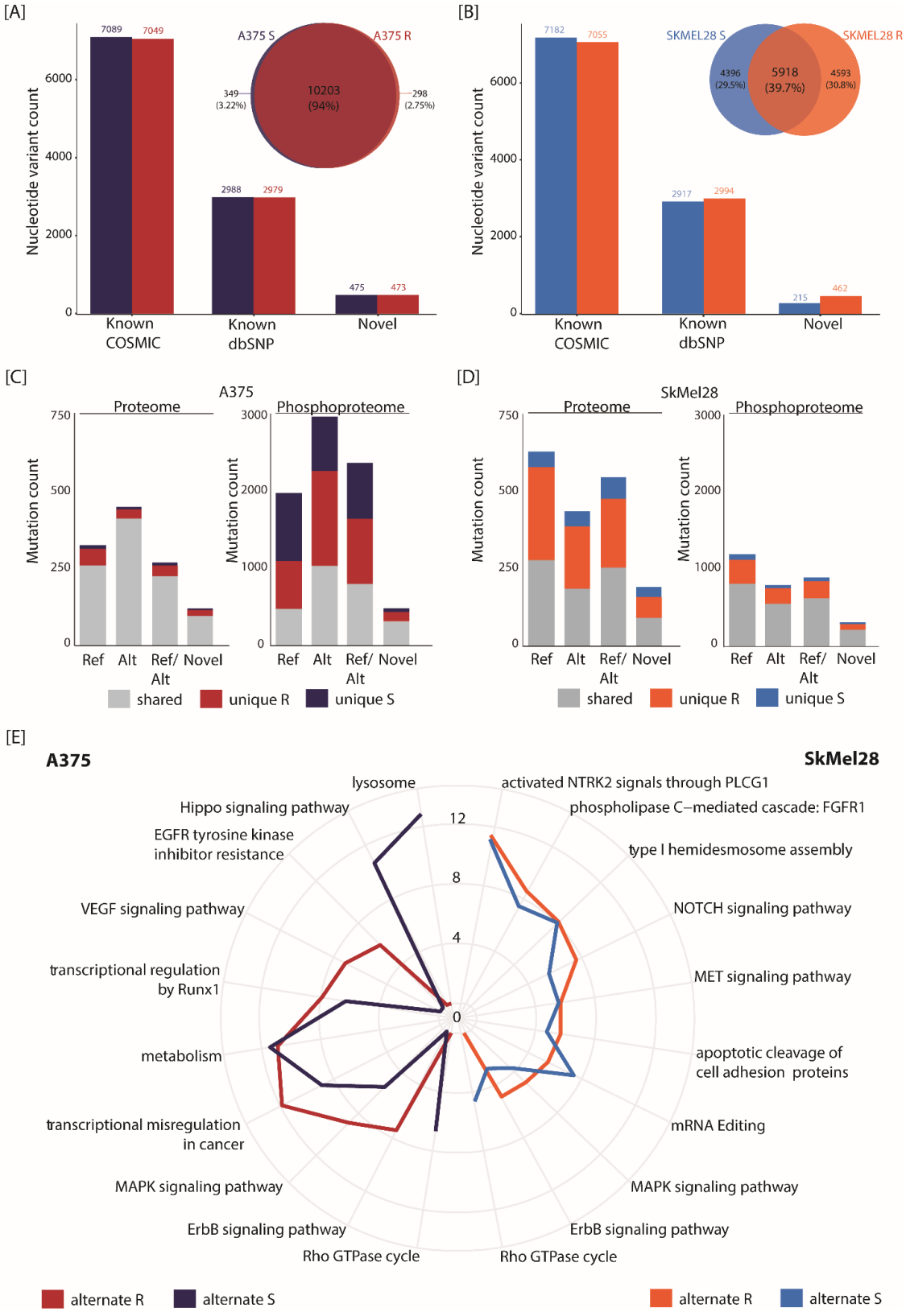


Figure 2: Individualised proteogenomics highlights key differences between dug-sensitive and drug-resistant melanoma cell lines. [A] and [B] Exome sequencing results of A375 [A] and SkMel28 [B], mapped to commonly used databases COSMIC and dbSNP and novel mutations identified in this study. Upper panel shows venn diagram of the comparison of drug-sensitive (S) and drug-resistant (R) cells. [C] and [D] Identification of reference, alternate and novel variants on the proteome and phosphoproteome level of A375 [C] and SkMel28 [D]. [E] Radial plot of over-represented pathway of proteins containing identified alternate variant peptides for A375 (left) and SkMel28 (right). The enrichment score was calculated by Fisher exact test (FDR < 0.1).

Mass spectrometry detects several alternate peptides phosphorylated on the mutation site

In the MS-based proteomics data, we focused on those mutations affecting the phosphorylation status of a protein. We identified 51 and 41 mutated isoforms with a phosphorylation event on the mutation site in A375 and SkMel28, respectively (**Figure 3A and B**). These comprised both phosphorylated reference peptides, implying a loss of the phosphorylation site due to the mutation, and phosphorylated alternate peptides, involving a gain of phosphorylation site. Notably, among the phosphorylated alternate peptides, approximately half were found phosphorylated only in the resistant phenotypes. The rest were phosphorylated either only in the sensitive phenotype or shared across phenotypes. Several of these mutations were among the top impact ranked mutations for each cell line, such as FLII, RUNX1, SCRIB, PPP1CA, KLF16, OGFR, RBMX and ANLN (**Figure S3A and B**). We checked whether these phosphorylated reference and alternate peptides contained kinase substrate motifs. While several peptides did contain a motif, significantly over-represented were only the ERK1,2 and Casein kinase II substrate motifs in A375 R (**Table S1**).

We then investigated the potential influence of phosphorylation site gain or loss on the protein-protein interactome of each cell line. Interestingly, the generated interaction network was highly enriched in proteins showing an up-regulation trend in A375 R. Conversely, for the interaction network of SkMel28, both phenotypes were equally represented in up-regulated proteins (**Figure 3C and D**). Among the mutated proteins used to generate these interaction networks, several had notably more connections, such as PCM1, RUNX1 and DTL for the A375 cell line and PPP1CA for SkMel28 (**Figure S3C and D**). Thus, the gain of a phosphorylation site on these proteins could be of importance not only for the affected protein but also for their numerous interactors. To characterise the interactome of each phenotype, we performed an over-representation analysis against GO functions, KEGG pathways and Reactome pathways (**Figure S3E and F**). The analysis revealed notable differences between phenotypes, with A375 R characterised by several pathways connected to transcription, mitosis and SUMOylation, whereas SkMel28 R was represented by telomerase activation.

Taken together these results suggest different phosphorylation landscape and possible rewiring of signal transduction networks in the resistant and sensitive phenotypes, an observation that was consistent across the two cell lines used.

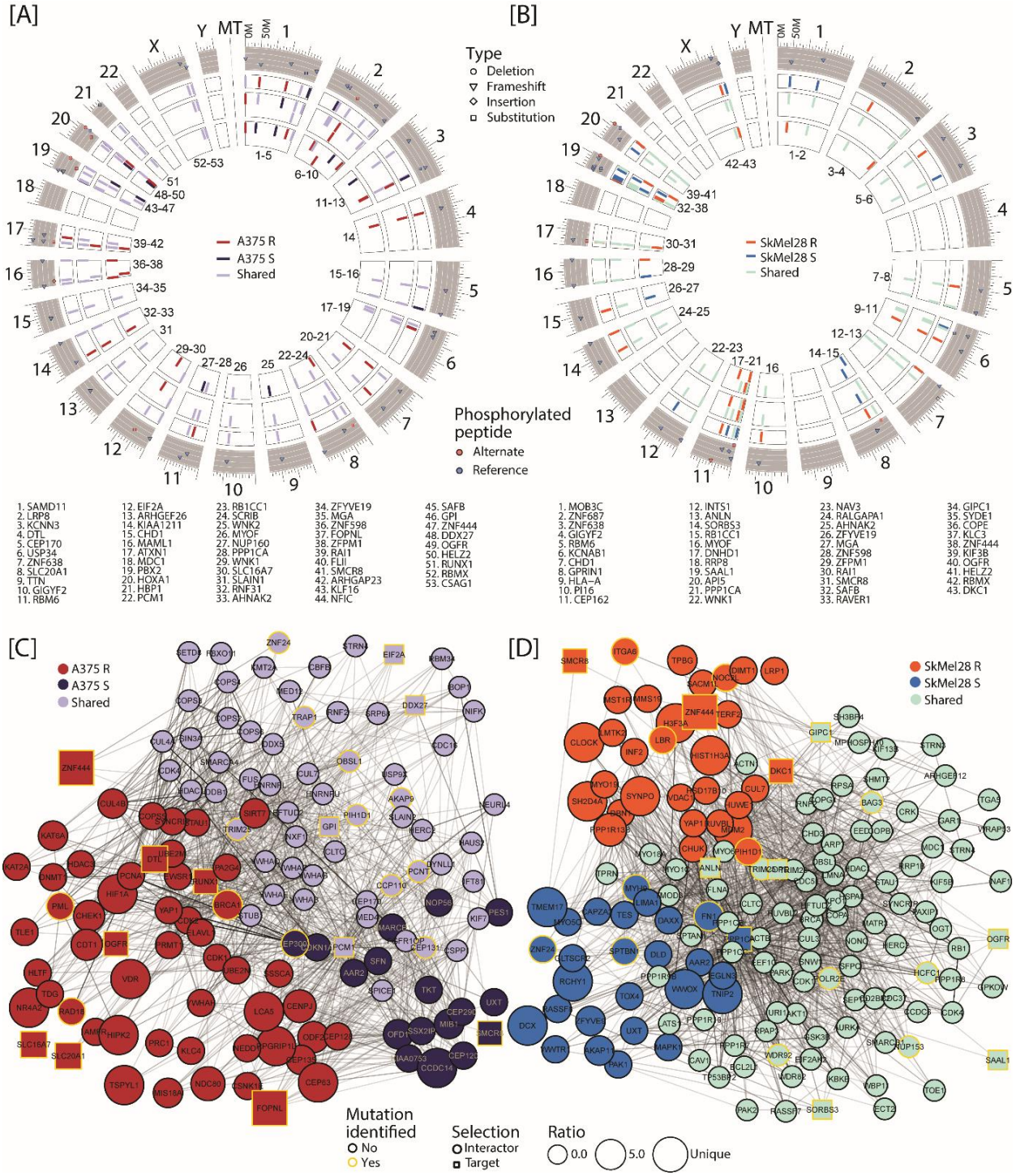


Figure 3: Mass spectrometry detects several alternate peptides phosphorylated on the mutation site. [A] and [B] The mutated proteins, identified with a phosphopeptide covering the mutation site either as reference or alternate, are displayed for A375 [A] and SkMel28 [B]. From the outer to the inner track, the first track represents the mutation genomic location in context of its impact score; with the phosphorylated mutation sites located on a reference peptide (blue) or on an alternate peptide (red). The second track depicts mutation genomic location for the WES dataset, with the colour highlighting whether this mutation was identified only in the sensitive cells, resistant or in both. The third and fourth tracks display the same type of information based on identified reference and alternate peptides, respectively. Similarly, the identified phosphorylated peptides are represented on the fifth and sixth tracks. [C] and [D] The protein-protein interactomes of the confirmed protein with loss/gain of phosphorylation sites (as well as their interactors) for A375 [C] and SkMel28 [D]. Only identified proteins are displayed and coloured based on up-regulation trend in one or another or both phenotypes.

Key molecules of MAPK pathway harbour different alterations in melanoma cells

Next, we selected the most commonly altered genes in melanoma and assessed their alterations on the exome and protein level as well as their expression and phosphorylation status for both cell lines and phenotypes (**Figure S3G**). Melanomas most frequently harbour alterations in BRAF, CDKN2A, NRAS, TP53 and NF1. Several mutations were identified on the exome and the proteome level including BRAF and EGFR in the MAPK pathway, and ERBB2 and ROS1 in the PI3K-AKT-mTOR pathway. Both phenotypes of the cell lines showed differently altered genes and proteins identified by our proteogenomics workflow. Mutations in PTEN were only identified in A375 R cells; CDK4 was only identified to be altered in SKMel28 S cells and the alternate peptide was identified by LC-MS/MS. Protein expression analysis based on label-free quantification revealed significant differences between cell lines and phenotypes. For example, KMTD2 showed a higher intensity in drug-sensitive A375 S cells compared to A375 R and TP53 is up-regulated in SkMel28 R cells compared to SkMel28 S cells. Several of the most commonly altered proteins in melanoma were identified to be phosphorylated on known but also on new modification sites, demonstrating the power of proteogenomic data integration to detect sample-specific changes in protein modifications.

Loss of a known phosphorylation site on RUNX1 leads to changed interactome and altered transcriptional activity

Among high impact mutations identified with our proteogenomic approach, one was a previously reported mutation on transcription factor RUNX1, a key transcription factor involved in cell proliferation, differentiation and apoptosis [156]. We identified several mutations on RUNX1 with a potential to change the modification status of the protein (**Figure 4A**). One of the identified alterations on *RUNX1* gene results in a loss of known phosphorylation site at S276 to L276 on RUNX1 protein. The reference and alternate peptides were identified with high resolution mass spectrometry 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 [390]. We therefore hypothesised that this mutation is likely to influence the interactome of RUNX1. To study the impact of the loss of a modifiable amino acid, we performed immunoprecipitation of Flag-tagged RUNX1_wt and RUNX1_S276L in CRISPR/Cas9-mediated RUNX1 knockout (KO) SILAC labelled cells (**Figure 4B**). 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 S4D and E**). Interestingly, histone deacetytranferase HDAC1 was enriched in RUNX1_wt interactome and depleted in the RUNX1_S276L interactome (**Figure 4B**). HDAC1 is a known interaction partner of RUNX1 and acts as a transcriptional repressor by removing acetylation modification from histone [391]. Another protein that significantly differed in the interactome of RUNX1 was the protein RAB21, a Ras-related protein. To confirm these findings, we performed pulldown assays with synthetic peptides harbouring the amino acid sequence for reference, alternate or phosphorylated reference peptides of RUNX1 in A375 cells (**Figure 4C**). As in the interactome study, HDAC1 was significantly depleted in the pulldown of alternate peptide versus reference peptide indicating that the interaction between HDAC1 and RUNX1 is disturbed due to the alteration. In addition, we also identified SIN3A to be significantly depleted in alternate peptide pulldown compared to reference pulldown. Similar to HDAC1, SIN3A acts a transcriptional repressor and both are forming a corepressor complex with RUNX1 which regulates the transcription of hematopoietic genes [176]. Besides the transcriptional regulators, we also identified several other proteins and known interactions partners of RUNX1. RIN1, a RAS and Rab interactor and PTPN23, a tyrosine-protein phosphatase, showed the same trend as HDAC1 and SIN3A, both proteins are known to act as regulator of RAS-mediated mitogenic activity [392,393]. In addition, we compared the alternate peptide pulldown to the phosphorylated peptide pulldown and identified several proteins involved in the nuclear core complex to be significantly enriched in the phosphopeptide pulldown compared to the alternate peptide pulldown (**Figure S4F**). Enrichment analysis of KEGG pathways indicated that proteins identified in alternate peptide pulldown compared to reference peptide were enriched in TGFβ signalling pathway, melanogenesis and insulin pathway signalling (**Figure S4G**). These include key molecules like PML, CTBP2 and YAP1. GOBP enrichment revealed that the proteins identified in the reference pulldown proteins are enriched for catenin import into nucleus and histone deacetylation. Proteins identified in the alternate pulldown are also involved in regulation of gene expression and cytokine biosynthetic processes. Next, we mapped a list of transcriptionally regulated proteins by RUNX1 from

BioGrid to our dataset and showed that the moiety of identified proteins is enriched in alternate pulldown compared to reference pulldown (**Figure S4H**). Taken together, we could show that the loss of known phosphorylation site has an impact on the interactome of RUNX1 (**Figure S4I**) and postulate that it leads to altered transcriptional activity (**Figure 4D**).

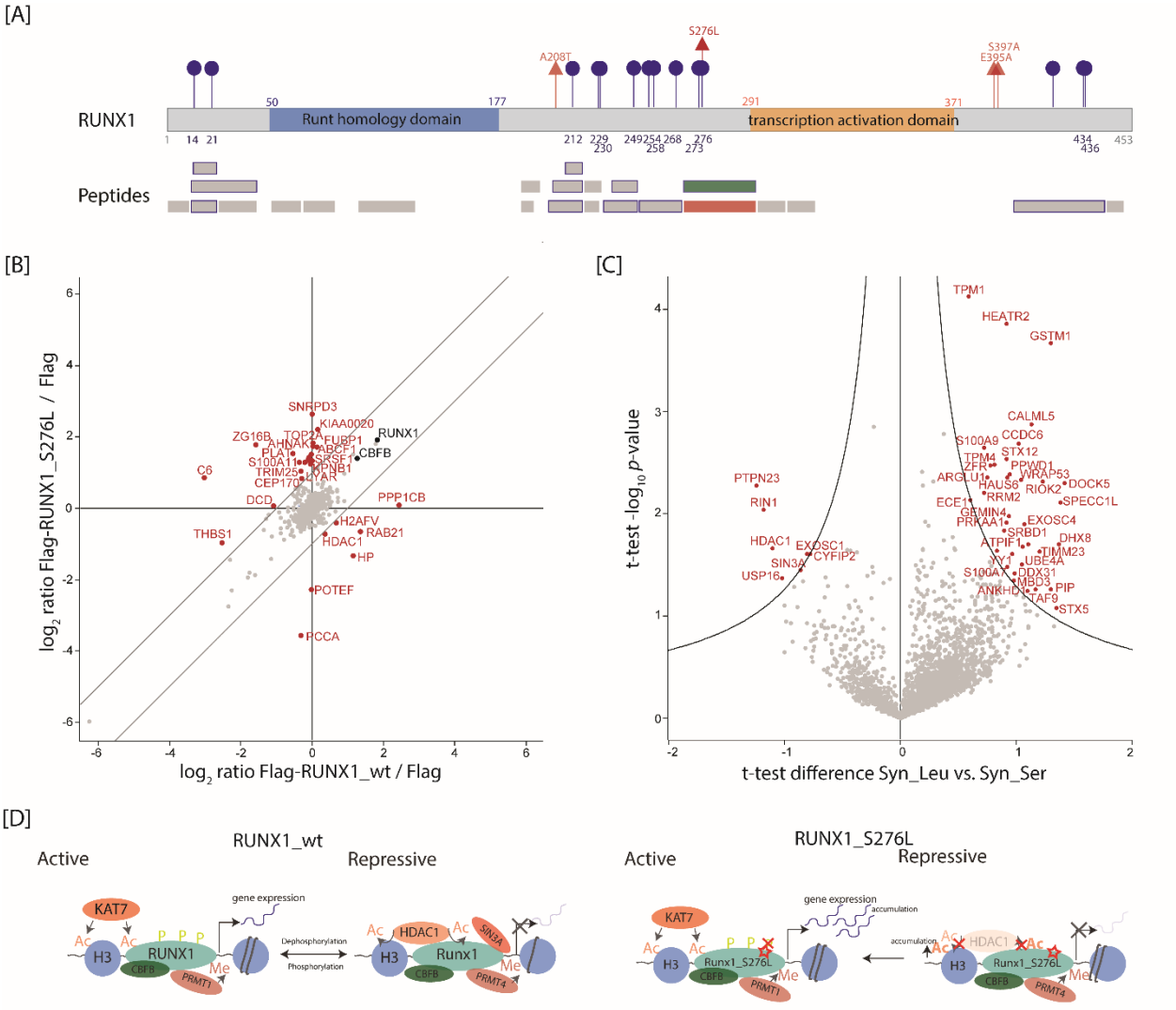


Figure 4: Loss of a known phosphorylation site leads to a changed interactome and altered transcriptional activity of RUNX1. [A] Schematic overview of the transcription factors RUNX1 protein. Numbers indicate the positions of amino acids residues within the protein. Identified phosphorylation sites are highlighted in blue and mutations by our proteogenomic workflow 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 peptide 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 (log2) of RUNX1_wt or FLAG-tagged RUNX1_S276L relative to the corresponding control cell line (FLAG tag only). RUNX1 and its core binding factor CBFC are marked in black. Significantly up and down regulated 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. Fold change of ratios between Syn_Leu and Syn_Ser (log2) are plotted against p-value (-log10) (n=3). Black lines indicate the significance threshold based on student t-test (FDR < 0.01; S0=1.2). Significantly up and down regulated proteins are highlighted in red. [D] Schematic overview of proposed interaction of RUNX1_wt and RUNX1_S276L with main transcriptional regulators.

Discussion

In the current study, we presented a new web application, PCTi, which allows a simple and user-friendly proteogenomics analysis in an individualised fashion. Onco-proteogenomics can help to identify clinical biomarkers or actionable drug targets [357,359,360]. Therefore, as a proof of principle we applied our tool to two cell line models of melanoma, a cancer that is well known for its high mutation load [135] and the potential for rewiring cellular networks [332,340]. 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 [360-363]. However, proteogenomics studies are still relatively rare and, due to their complexity, out of reach of most proteomics (or genomics) laboratories.

PCTi application provides a reactive environment for proteogenomics analysis

We hope that PCTi will facilitate proteogenomics projects by providing a simple reactive environment to construct tools (e.g. protein databases) and integrate different omics datasets. Among the key features of PCTi is a user-friendly interface that can be used as a part of existing WES and MS data processing workflows. We have modularised our application so that users are free to use all or only part of it (e.g. possible to generate a custom database without pursuing with the rest of PCTi). Nevertheless, all modules are designed to work together in order to create a homogeneous workflow, leading to interactive visualisations. Importantly, our tool is suitable for a broad range of biomedical questions and is not limited to oncology studies as it depends on user input (i.e. WES data, MS data, selected impacts). PCTi currently has some restrictions, i.e., the variants must be called against *H. sapiens* genome assembly GRCh38, and MS data processed using MaxQuant software suite. We expect these limitations to disappear as the application matures. While the application can in principle be used for non-human organisms, it was never tested in such context.

Regarding PCTincorporation module, we decided to incorporate non-synonymous nucleotide variants into a reference genome for the creation of custom protein sequences database. While this approach is not novel, it has the advantage to preserve the reference assembly used during variant calling and to allow creation of a reference protein sequences database (as opposed to retrieving it from online sources) [394]. We also decided to incorporate variants mapping to the same gene locus separately, resulting in separate protein sequence entries (e.g. PCM1 gene mutated at position 159, 597 and 691). Among our reasons is the fact that few variants are located close enough from each other to be detected by MS (typical tryptic peptide length is comprised between 7 and 30 amino acids). Furthermore, variants phasing information are not

necessarily available following variant calling, without which it is impossible to predict which variants belong to the same allele. In the context of our melanoma cell lines, the nucleotide variants incorporation revealed very similar numbers (~10,000 variants) across cell lines/phenotypes, the large majority being SNVs (**Figure S1A and B**), which is consistent with previous study [395]. We also observed characteristic nucleotide substitutions, whereby two thirds of substitutions are comprised of transitions, with the remaining one third being transversions (**Figure S2A and B**). The C to T transition was highly represented and has been reported as a result from sun-light exposure, which is highly relevant for skin cancer [396].

Due to the large number of nucleotide variants identified by WES, representing nearly twice as many variant isoforms, we implemented an impact scoring in PCTimpact module to enable stratification of variants, e.g. in clinical context. For this purpose, our impact method relies almost entirely on user input, and enables pre-definition and prioritisation of an unlimited number of impacts (via parameters customisation). With regards to application of PCTimpact to melanoma cell lines, we prioritised mutations that have been reported in melanoma patients (based on CGDS resource) [397]; and those that fall on oncogenes and tumour suppressor genes. We also focused on mutations that affect phosphorylation status. Around 14.8% of all amino acids in the human proteome are serine, threonine or tyrosine [344], which are predominantly modified by phosphorylation. Several studies have reported that these three amino acids are disproportionally affected by missense mutations [366,367].

Here, we predicted the impact of several thousand mutated isoforms, including the impact of mutated protein residues in context of loss or gain of S/T/Y in A375 and SkMel28. 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 mutations [340,362,371]. Among the mutated isoforms with impact score medium to high (**Figure S3A and B**), two thirds were known oncogenes or tumour suppressors genes; and more than 80% had a loss of known phosphorylation sites. Interestingly, about 10 nucleotide variants in A375 and SkMel28 were known with frequencies ranging from 0.8% to nearly 30% of melanoma samples in CGDS. The affected genes were *THBS1*, *ZFHX3*, *OGFR*, *RET*, *SYNE2*, *BRCA1*, *MUC16*, *SELENOP* and *RUNX1*; most of which have been shown to be involved in tumour progression and maturation or development of drug-resistance in melanoma or other cancer types [398-406]. Thus, we believe that careful selection of impacts by users (i.e. based on online resources, experiment) can lead to meaningful prioritisation of nucleotide variants, as demonstrated here for melanoma.

Integration of genomics and proteomics reveals differing mechanisms linked to BRAFi resistance in melanoma

While we detected near identical number of nucleotide variants across cell lines/phenotypes (**Figure 2A and B**), we observed very different overlap between phenotypes. Indeed, A375 S and R had very few unique nucleotide variants, as opposed to SkMel28 S and R where nearly two thirds of variants were unique to one or the other phenotype. Checking the sequencing depth of these samples revealed similar distribution between phenotypes, suggesting that these differences may originate from the underlaying biology of these cell lines [389]. 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 [352,354,407,408]. Over-representation of variant proteins revealed that cancer mutations are accumulating in MAPK and PI3K/AKT/mTOR pathways in drug-resistant cell lines, whereas YAP/TAZ stimulated gene expression was enriched in sensitive cells. The identified enriched pathways are known to be highly activated in melanoma cells with acquired resistance [30,119,409]. In addition, we report stronger differences in enriched pathways between A375 R and S compared to Skmel28 S and R. Thus, we suggest the usage of A375 cell line for the study of resistance in melanoma.

Following integration and visualisation of the different omics data for the variants affecting phosphorylation status, we observed a similar trend as described above whereby most mutations are shared between A375 R and S, whereas a large number of mutations were unique to SkMel28 R or S (**Figure 3**). Notably, the agreement between omics was good, there was no case of variant found unique to a phenotype at the WES level that was found unique to the other phenotype at the proteome or phosphoproteome levels (e.g. nucleotide variant unique to A375 S while alternate peptide found unique to A375 R). Among interesting variant, RUNX1 was one of the top impact scoring entries for A375, while a frameshift in PPP1CA was found unique to SkMel28 R (**Figure S3A and B**). Several over-represented pathways highlighted striking differences between the sensitive and resistant phenotypes as a result of confirmed loss or gain of phosphorylated residues. For example, pathways connected to transcription, mitosis and SUMOylation were over-represented in A375 R, these functions when dysregulated have been shown to induce drug resistance [410,411]. Whereas, SkMel28 R was over-represented in telomerase activation pathway, which is also known in context of melanoma resistance as well as a potential drug target to overcome such resistance [412,413].

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

We experimentally validated one striking example of a loss of a known phosphorylation sites on RUNX1 and showed that this mutation 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 [165]. In this study, we also identified several mutations that may have an influence on the modification status of the protein and one was also confirmed by LC-MS/MS (**Figure 4 and S4**). The mutation site S276L does not fall on a domain on the protein RUNX1 but is located in a highly modified region of the protein and thus may influence the 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 [178]. Here, we could not identify p300 in the interactome studies of RUNX1 by immunoprecipitation of overexpressed RUNX1 or synthetic peptide pulldowns. Interestingly, we identified the transcriptional activator WWTR1 (TAZ) and KAT7 and the corresponding transcriptional repressors HDAC1 and Sin3A to be significantly changing between reference and alternate pulldown of RUNX1 **(Figure 4B and C)**. The loss of the interaction to HDAC1 by mutating RUNX1 at S48, S303 and S424 to aspartic acid *in vitro* has been also described previously [414]. However, we could link it to a different mutation site and show that the interaction is associated with the modification status of the protein. The crosstalk between acetylation/deacetylation mediated by KAT7 and HDAC1 and phosphorylation/dephosphorylation may alter the transcriptional activity by RUNX. We postulate that the regulation of transcriptional activity of RUNX1 is p300 independent and may suggest a NAMPT dependent regulation through sirtuins, mainly SIRT2. Here, we show the enrichment of NAMPT in reference pulldown analysis. Taken together, we postulate that the mutation, which influence the modification status of the protein, change 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, proteome and PTM level. Here, we developed a new software tool, termed PCTi, and applied it to study mutational landscape of two melanoma cell lines sensitive and resistant to BRAF inhibition. Our approach revealed key differences between BRAFi-sensitive and -

resistant melanoma cell lines, such as over-representation of MAPK signalling and ErbB signalling pathways in resistant phenotypes. We also confirmed the loss or gain of several phosphorylation events on mutation sites, that may lead to the rewiring of signal transduction network in context of melanoma resistance to BRAFi. We further investigated the mutation S276L on the transcription factor RUNX1 and confirmed the loss of the Ser276 phosphorylation site by MS. Our results suggest that this mutation has an impact on the interactome of RUNX1 and may be responsible for change in its transcriptional activity. This dataset can be used as a resource to identify network-attacking mutations to improve patient's survival.

Materials and Methods

Extended method description can be found in **Supplementary Information.**

Cell culture

The human metastatic $BRAF^{V600E}$-mutated melanoma cell lines A375 and SKMel28 were used in this study. The generation of the cell lines with acquired resistance to vemurafenib analogue PLX4720 (Selleckchem) (for simplicity referred to as "vemurafenib" in the Results section) was conducted as described previously [387]. A375 S/R and SkMel28 S/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.

SILAC –labelling of cells was performed as described previously [415].

DNA extraction and Whole Exome Sequencing

Cells were harvested by centrifugation and DNA was extracted using QIAamp DNA Mini (QIAGEN) according to the manufacturers' instructions. WES libraries were prepared using SureSelect Human All Exon (Agilent) according to manufacturers' instructions. Paired-end sequencing was performed on an HiSeq 2500 instrument (Illumina). The WES measurements were performed at c.ATG Core Facility in Tuebingen. Raw sequence data were then processed using an in-house pipeline developed at the Proteome Centre Tuebingen according to GATK best guidelines [416].

Incorporation of non-synonymous variants into protein databases

PCTi application is coded entirely in the R programming language [417] and uses Shiny package [388] for its graphical user interface. The first module, PCTincorporate takes as input a Variant

Call Format (VCF) file containing nucleotide variants from one or more samples. 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 were injected into each overlapping reference transcript nucleotide sequences and then *in silico* translated. The computed information was reported directly within the alternate FASTA header to facilitate interaction with the rest of PCTi (e.g. mutation positions, reference ID). The output consists in two FASTA files containing reference protein sequences and sample-specific alternate protein sequences, which are used as protein databases for processing of LC-MS/MS data.

Annotation of the biological impact of detected variants

The FASTA file discussed above is used as the main input to the PCTimpact module, in conjunction with several user-defined specifications that define how the mutations impact are computed. In the current study, the focus was on the impact of amino acid variants on protein phosphorylation-based signal transduction networks in melanoma. PCTimpact module was used to annotate each reference/alternate protein sequence 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 [418] and these motifs were searched along the reference/alternate protein sequences. Located kinase motifs were overlapped with mutation position to determine loss/gain of the motifs. Known human phosphorylation sites were retrieved from PhosphoSitePlus and Phospho.ELM databases [382,419]. The mutations affecting these sites were annotated as loss/gain of phosphorylation. In a similar fashion, known mutation sites in melanoma were obtained from CGDS [397] and overlapped with our mutations. A list of oncogenes and tumour suppressor genes was compiled from Cosmic, ONGene, Bushman lab and Uniprot [344,420,421]. Mutations on these genes were annotated as cancer-relevant. A Levenshtein similarity score was calculated between reference and alternate protein sequences and mutated sequences with less than 90% similarity to their reference were flagged. Each impact was scored with the application default (i.e. no factor was applied to one or another impact). A summed score was calculated for each alternate sequences' amino acid, and the maximum summed score was reported for that mutated isoform. Because the score depends on the number of impacts tested by the user, we also computed a scaled maximum score (between 0 and 1), to allow comparison between processings. Following the computation of all impacts, each mutated protein isoform is scored and ranked to allow prioritisation for follow up studies (tab-separated file as output).

Overexpression and immunoprecipitation of RUNX1

Immunoprecipitation (IP) of overexpressed Flag-tagged RUNX1 in A375 S RUNX1_KO SILAC cells was performed with Flag M2 antibody in three biological replicates. *RUNX1* gene knockout was carried out by CRISPR/Cas9-mediated genome editing according to the published protocol [422].

Pulldown assays with synthetic peptides and on-bead digestion

Synthetic peptides comprising 17 amino acids and a biotinylated linker in the N-terminus were coupled to Pierce streptavidin magnetic beads and synthetic peptide pulldowns in A375 S cells were performed in three independent replicates (**Supplementary Information**).

Extraction and digestion of proteins

Cells were harvested with lysis buffer and digested with trypsin essentially as described before [415].

High pH reverse phase chromatography

Prior fractionation, peptides were purified on a Sep-Pak C18 Cartridge (Waters). High pH reverse phase chromatography was conducted using off-line Ultimate 3000 high-pressure liquid chromatography (HPLC) system (Dionex, Thermo Fischer Scientific) equipped with xBridge BEH130 C_{18} 130A, 3.5 µm, 4.6 x 250 mm column (Waters), as described previously [415].

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, 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% TFA) at 40°C 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, 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) [285]. The Andromeda search engine [423] searched MS/MS data against PCTi *H. sapiens* reference (99,354 entries) and cell line-specific alternate databases (A375 = 29,104 entries; SkMel28 = 40,041 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 of melanoma cell lines, 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

The PCTintegrate module was used to integrate WES and MS datasets, specifically to check which mutations were identified across datasets. Initially, the reference and alternate 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 the determination of 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) peptide classification. On the basis of this peptide classification, we summarised the peptides identification per mutated 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 peptides found phosphorylated were flagged as such, as well as those were the phosphorylation occurred directly on the mutated sites (either on reference or alternate sequences). This coverage information is exported within MaxQuant style processing results (tab-separated file as output).

PCTimage module was used to visualise mutated position along *H. sapiens* genome in a sample specific context (using ggplot and ggbio packages). The module allowed protein isoform specific representation, which includes mutation sites, impacts location and MS identification coverage. We also generated network of protein-protein (using BioGRID database), drug-target (using Uniprot database) and prediction kinase-substrate (PCTimpact results) interactions; e.g. RUNX1 network [344,424]. The generated networks were exported (using igraph and RCy3 packages) into Cytoscape for further customisation [425].

Pathway analysis

Statistical analyses were performed with Perseus software suite (version 1.6.5.0). The drug-sensitive and drug-resistant cell lines were compared using label-free quantification, after filtering of all reverse and potential contaminant hits. Significance B (p value 0.05) test was used for statistical analysis. A list of all protein identifications and phosphorylation site identifications are provided in **Table S1 and 2**.

For proteomic interaction studies, protein groups were kept for further statistical analysis only if quantified in 3 out of 3 replicates. The SILAC ratios of the three independent replicates were averaged and an arbitrary cut-off of two-fold change was used to determine significant SILAC ratios. The log2 transformed ratios were plotted against intensities (log10). For synthetic peptide pulldowns, label-free quantification between three independent replicates was

performed and ratios were subjected to t-test analysis, with a permutation-based FDR threshold of 0.01 and s0 value of 1.2. A list of known interaction partners and transcriptionally regulated targets by RUNX1 were retrieved from BioGrid and mapped to the dataset. A list of all protein identifications are provided in **Table S3**.

The resources used for annotation of proteins were Gene Ontology (GO), Biological Processes (GOBP), GO Cellular Compartment (GOCC), GO Molecular Functions (GOMF) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Reactome Pathway database (Reactome). The fisher exact test (FDR ≤ 0.5) was used to checked for over-represented categories among significantly changing proteins (between drug-sensitive and drug-resistant cell lines or between reference and alternate pull-down). A list of all over-representation results is provided in **Table S1, 2 and 3**.

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 the synthetic peptides, Prof. Dr. Stefano Stifani for the CMV_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 PCTi web application is freely available and hosted at https://shiny.pct.uni-tuebingen.de/pcti/. The high throughput nucleotide sequencing data have been deposited in the NCBI Sequence Read Archive [426] with the bioproject accession number PRJNA616103. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [427] partner repository with the dataset identifier PXD018305. The WES bioinformatics pipeline is available online [428]. Excel files containing the analysed data are provided in **Supplementary Information**.

Supplementary Information

Schmitt, M., Sinnberg, T., Bratl, K., Garbe, C., Macek, B., Nalpas N. C.

Integration of individualized proteogenomics datasets to analyse single amino acid variants in malignant melanoma

Under revision in Frontiers in Oncology, 2020

Table 1: List of used patient derived melanoma cell lines in this study. ATCC: American Type Culture Collection

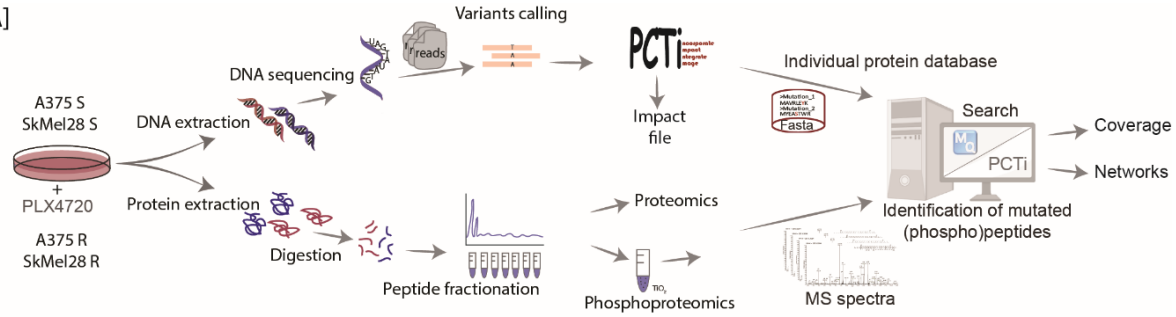
Name	Tissue	Morphology	Source	BRAF mutation
A375	skin	epithelial	ATCC (CRL-1619)	BRAF V600E
SKMel28	skin	polygonal	ATCC (HTB-72)	BRAF V600E

Table 2: List of used synthetic peptides in this study.

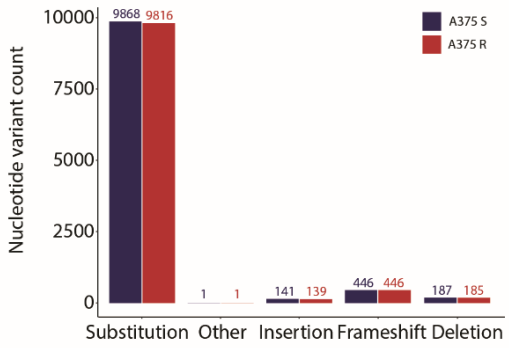
Peptide	Sequence
Runx1_S276	SGSGSPSVHPATPISPGRASGM
Runx1_L276	SGSGSPSVHPATPILPGRASGM
Runx1_p276	SGSGSPSVHPATPISPGRASGM
Runx1_pT273pS276	SGSGSPSVHPATPISPGRASGM

Kind gift of Prof. Dr. Stefan Stefanovic, University of Tuebingen

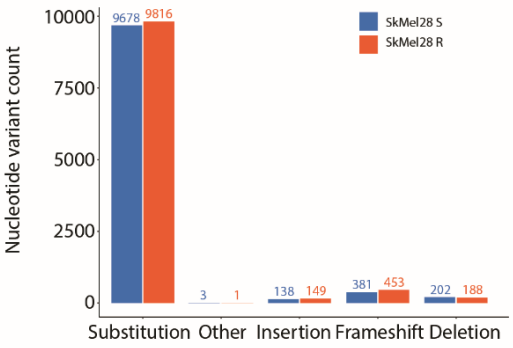
[S1A]



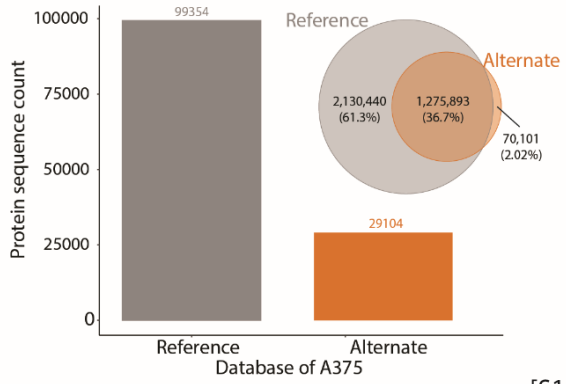
[S1B]



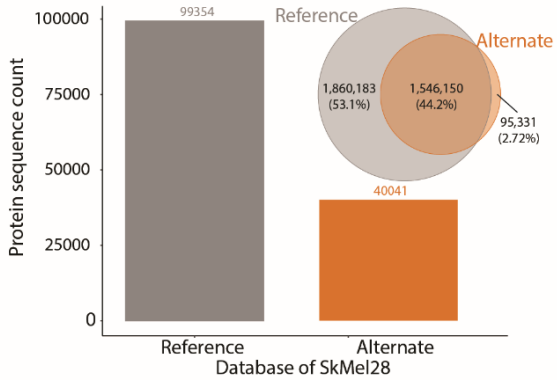
[S1C]



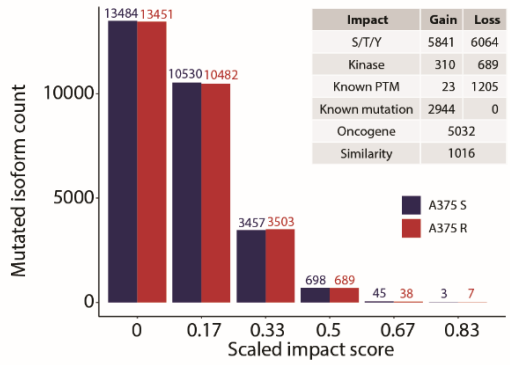
[S1D]



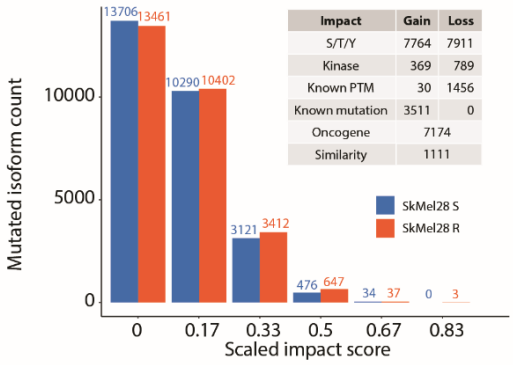
[S1E]



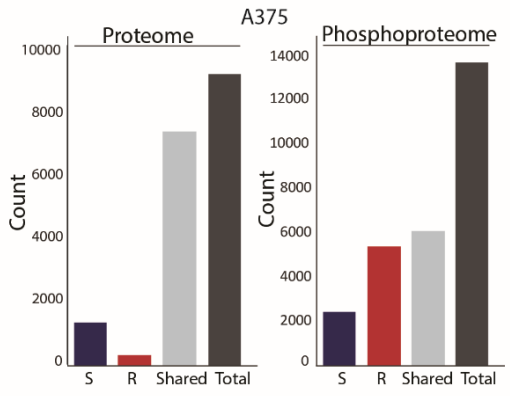
[S1F]



[S1G]



[S1H]



[S1I]

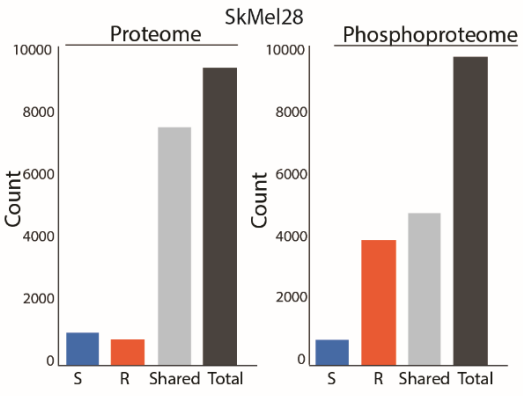


Figure S1: PCTi application provides a reactive environment for the integration of genomics with proteomics. [S1A] Schematic overview of the proteogenomic workflow. Vemurafenib sensitive (S) and resistant (R) melanoma cell lines A375 and SkMel28 were used in this study. For exome sequencing, DNA was extracted and sequenced on Illumina HiSeq 2000. Variants were called using GATK software. For the proteomic and phosphoproteomic workflow, cells were lysed and proteins were digested using trypsin. The resulting peptide mixture was fractionated using an off-line RP HPLC operated at high pH. Fractions were pooled and measured directly (proteome) or applied to phosphopeptide enrichment using titanium dioxide (TiO2) prior to LC-MS/MS. MS raw data was processed with MaxQuant software and analysed by PCTi. [S1B] and [S1C] Number of non-synonymous nucleotide variants per mutation types identified by WES for A375 [S1B] and SkMel28 [S1C]. [S1D] and [S1E] The number of protein sequences per reference or alternate databases for A375 [S1D] and SkMel28 [S1E], as well as the overlap in search space between databases (up to two missed cleavages). [S1F] and [S1G] Mutated isoform count per impact score for A375 [S1F] and SkMel28 [S1G], whereby 0 means no impact and 1 means high impact. The table displays mutated isoform count for each type of lost/gained impact. [S1H] and [S1I] Protein and phosphosite counts identified by MS for A375 [S1H] and SkMel28 [S1I].

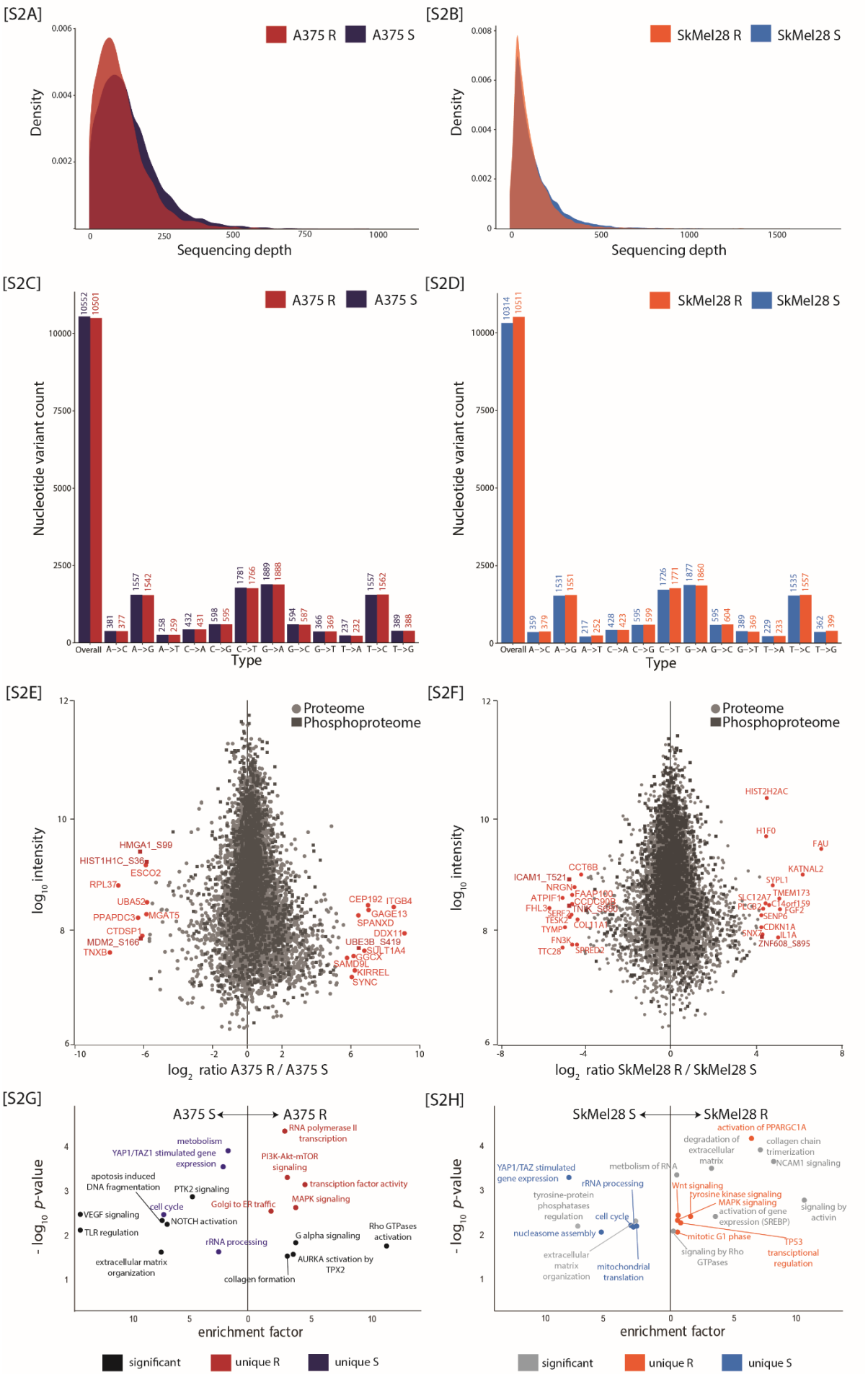

[S2A]
A375 R
A375 S
Density
Sequencing depth
[S2B]
SkMel28 R
SkMel28 S
Density
Sequencing depth
[S2C]
A375 R
A375 S
Nucleotide variant count
Overall A->C A->G A->T C->A C->G C->T G->A G->C G->T T->A T->C T->G
Type
[S2D]
SkMel28 R
SkMel28 S
Nucleotide variant count
Overall A->C A->G A->T C->A C->G C->T G->A G->C G->T T->A T->C T->G
Type
[S2E]
Proteome
Phosphoproteome
log10 intensity
log2 ratio A375 R / A375 S
[S2F]
Proteome
Phosphoproteome
log10 intensity
log2 ratio SkMel28 R / SkMel28 S
[S2G]
A375 S
A375 R
- log10 p-value
enrichment factor
significant
unique R
unique S
[S2H]
SkMel28 S
SkMel28 R
- log10 p-value
enrichment factor
significant
unique R
unique S

Figure S2: Individualised proteogenomics highlights key differences between dug-sensitive and drug-resistant melanoma cell lines. [S2A] and [S2B] Sequencing depth density of the WES data obtained for A375 [S2A] and SkMel28 [S2B]. [S2C] and [S2D] Count of nsSNV transitions and transversions for A375 [S2C] and SkMel28 [S2D]. [S2E] and [S2F] Scatter plot of A375 R and A375 S [S2E] and SkMel28 R and SkMel28 S [S2F] proteomes (filled circles) and phosphoproteome (filled squares). The log2 transformed ratios are plotted against intensities (log10). The top significantly regulated proteins and phosphoproteins are highlighted in red. [S2G] and [S2H] Over-representation of selected Reactome pathways for A375 [S2G] and SkMel28 [S2H]. The enrichment score calculated by Fisher exact test were plotted against the p-value (-log10) (FDR < 0.1).

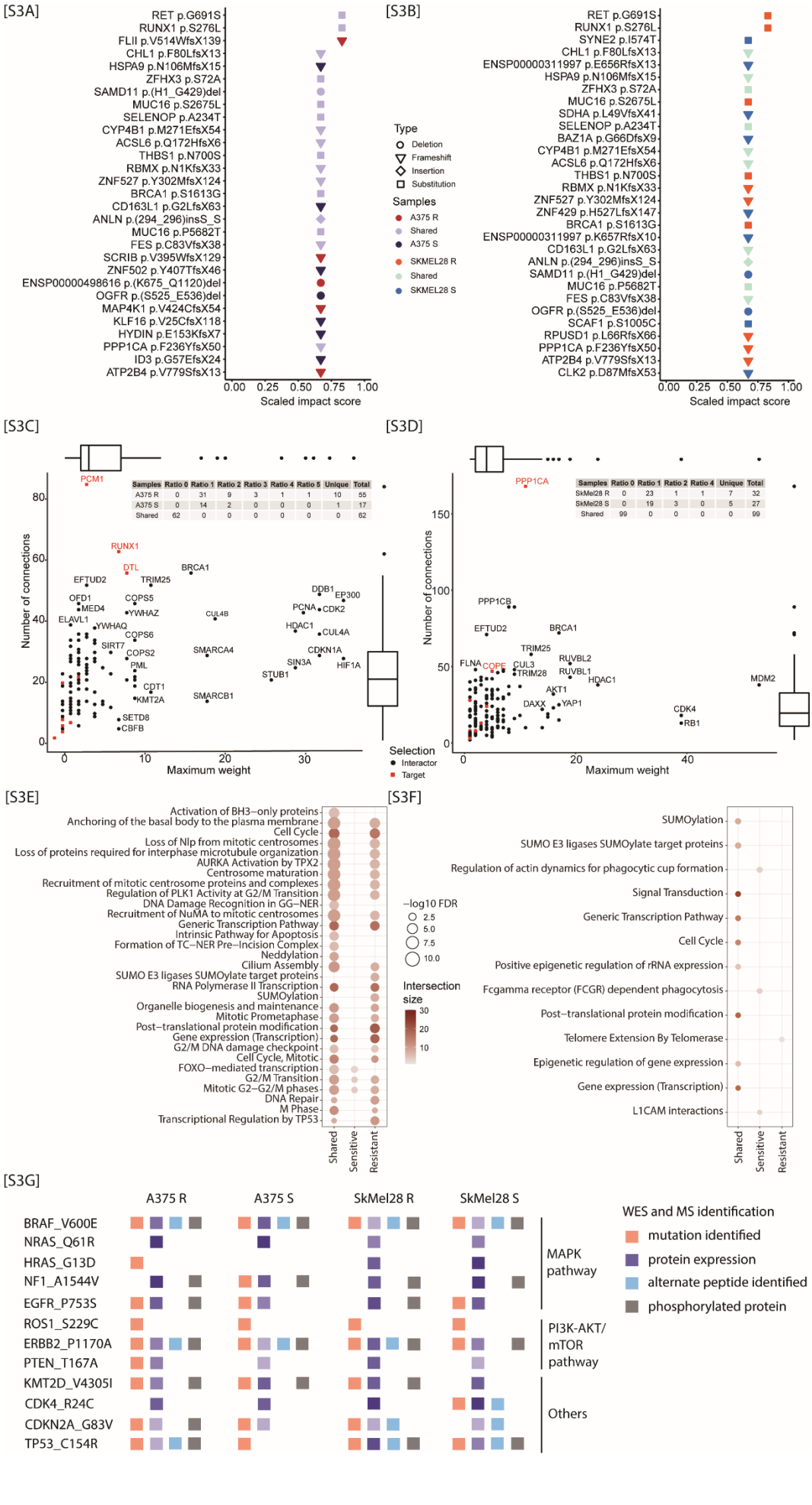
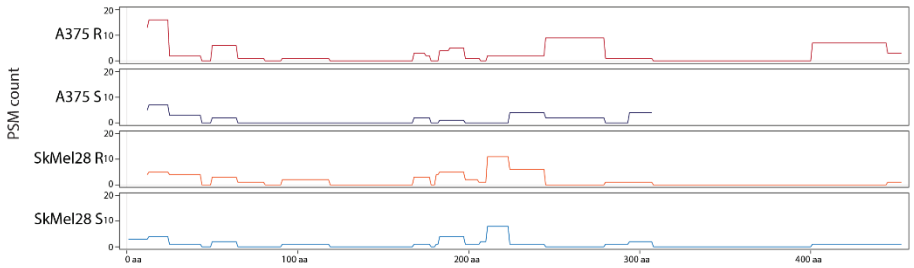
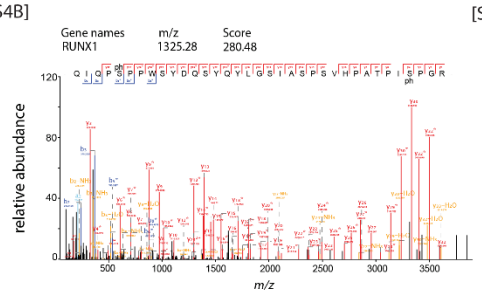

[S3A]
RET p.G691S
RUNX1 p.S276L
FLII p.V514WfsX139
CHL1 p.F80LfsX13
HSPA9 p.N106MfsX15
ZFHX3 p.S72A
SAMD11 p.(H1_G429)del
MUC16 p.S2675L
SELENOP p.A234T
CYP4B1 p.M271EfsX54
ACSL6 p.Q172HfsX6
THBS1 p.N700S
RBMX p.N1KfsX33
ZNF527 p.Y302MfsX124
BRCA1 p.S1613G
CD163L1 p.G2LfsX63
ANLN p.(294_296)insS_S
MUC16 p.P5682T
FES p.C83VfsX38
SCRIB p.V395WfsX129
ZNF502 p.Y407TfsX46
ENSP00000498616 p.(K675_Q1120)del
OGFR p.(S525_E536)del
MAP4K1 p.V424CfsX54
KLF16 p.V25CfsX118
HYDIN p.E153KfsX7
PPP1CA p.F236YfsX50
ID3 p.G57EfsX24
ATP2B4 p.V779SfsX13
0.00 0.25 0.50 0.75 1.00
Scaled impact score
Type
Deletion
Frameshift
Insertion
Substitution
Samples
A375 R
Shared
A375 S
SKMEL28 R
Shared
SKMEL28 S
[S3B]
RET p.G691S
RUNX1 p.S276L
SYNE2 p.I574T
CHL1 p.F80LfsX13
ENSP00000311997 p.E656RfsX13
HSPA9 p.N106MfsX15
ZFHX3 p.S72A
MUC16 p.S2675L
SDHA p.L49VfsX41
SELENOP p.A234T
BAZ1A p.G66DfsX9
CYP4B1 p.M271EfsX54
ACSL6 p.Q172HfsX6
THBS1 p.N700S
RBMX p.N1KfsX33
ZNF527 p.Y302MfsX124
ZNF429 p.H527LfsX147
BRCA1 p.S1613G
ENSP00000311997 p.K657RfsX10
CD163L1 p.G2LfsX63
ANLN p.(294_296)insS_S
SAMD11 p.(H1_G429)del
MUC16 p.P5682T
FES p.C83VfsX38
OGFR p.(S525_E536)del
SCAF1 p.S1005C
RPUSD1 p.L66RfsX66
PPP1CA p.F236YfsX50
ATP2B4 p.V779SfsX13
CLK2 p.D87MfsX53
0.00 0.25 0.50 0.75 1.00
Scaled impact score
[S3C]
PCM1
Samples Ratio 0 Ratio 1 Ratio 2 Ratio 3 Ratio 4 Ratio 5 Unique Total
A375 R 0 31 9 3 1 1 10 55
A375 S 0 14 2 0 0 0 1 17
Shared 62 0 0 0 0 0 0 62
RUNX1
DTL
BRCA1
EFTUD2
TRIM25
OFD1
COPS5
MED4
YWHAZ
CUL4B
ELAVL1
YWHAQ
COPS6
SIRT7
COPS2
SMARCA4
PML
CDT1
KMT2A
SMARCB1
SETD8
CBFB
DDB1
EP300
PCNA
CDK2
HDAC1
CUL4A
CDKN1A
SIN3A
HIF1A
STUB1
Number of connections
Maximum weight
Selection
Interactor
Target
[S3D]
PPP1CA
Samples Ratio 0 Ratio 1 Ratio 2 Ratio 4 Unique Total
SkMel28 R 0 23 1 1 7 32
SkMel28 S 0 19 3 0 5 27
Shared 99 0 0 0 0 99
PPP1CB
EFTUD2
BRCA1
TRIM25
RUVBL2
FLNA
COPE
CUL3
TRIM28
RUVBL1
HDAC1
AKT1
DAXX
YAP1
MDM2
CDK4
RB1
Number of connections
Maximum weight
[S3E]
Activation of BH3−only proteins
Anchoring of the basal body to the plasma membrane
Cell Cycle
Loss of Nlp from mitotic centrosomes
Loss of proteins required for interphase microtubule organization
AURKA Activation by TPX2
Centrosome maturation
Recruitment of mitotic centrosome proteins and complexes
Regulation of PLK1 Activity at G2/M Transition
DNA Damage Recognition in GG−NER
Recruitment of NuMA to mitotic centrosomes
Generic Transcription Pathway
Intrinsic Pathway for Apoptosis
Formation of TC−NER Pre−Incision Complex
Neddylation
Cilium Assembly
SUMO E3 ligases SUMOylate target proteins
RNA Polymerase II Transcription
SUMOylation
Organelle biogenesis and maintenance
Mitotic Prometaphase
Post−translational protein modification
Gene expression (Transcription)
G2/M DNA damage checkpoint
Cell Cycle, Mitotic
FOXO−mediated transcription
G2/M Transition
Mitotic G2−G2/M phases
DNA Repair
M Phase
Transcriptional Regulation by TP53
Shared
Sensitive
Resistant
−log10 FDR
2.5
5.0
7.5
10.0
Intersection size
30
20
10
[S3F]
SUMOylation
SUMO E3 ligases SUMOylate target proteins
Regulation of actin dynamics for phagocytic cup formation
Signal Transduction
Generic Transcription Pathway
Cell Cycle
Positive epigenetic regulation of rRNA expression
Fcgamma receptor (FCGR) dependent phagocytosis
Post−translational protein modification
Telomere Extension By Telomerase
Epigenetic regulation of gene expression
Gene expression (Transcription)
L1CAM interactions
Shared
Sensitive
Resistant
[S3G]
A375 R
A375 S
SkMel28 R
SkMel28 S
BRAF_V600E
NRAS_Q61R
HRAS_G13D
NF1_A1544V
EGFR_P753S
ROS1_S229C
ERBB2_P1170A
PTEN_T167A
KMT2D_V4305I
CDK4_R24C
CDKN2A_G83V
TP53_C154R
MAPK pathway
PI3K-AKT/ mTOR pathway
Others
WES and MS identification
mutation identified
protein expression
alternate peptide identified
phosphorylated protein

Figure S3: Mass spectrometry detects several alternate peptides phosphorylated on the mutation site. [S3A] and [S3B] The highest ranked impact scoring mutations for A375 [S3A] and SkMel28 [S3B] cell lines, including the type of mutation. [S3C] and [S3D] The protein-protein interactomes of the confirmed protein with loss/gain of phosphorylation sites (as well as their interactors) for A375 [S3C] and SkMel28 [S3D]. Each node is plotted on the basis of the maximum interaction weight (number of evidences for an interaction) and the number of connections (number of interactors). Nodes are colour coded in red if they represent a target protein used to generate the network or in black if they are an interactor. [S3E] and [S3F] Over-representation analysis of the interactome of each phenotype based on GO functions, KEGG pathways and Reactome pathways for A375 [S3E] and SkMel28 [S3F]. Each node represents a pathway with the colour gradient encoding the number of identified proteins in that pathway, while the node size equals the Benjamini-Hochberg adjusted p-value (-log10). [S3G] The top twelve mutations for malignant melanoma were selected and mutation status displayed based on exome sequencing, proteome and phosphoproteome identification.

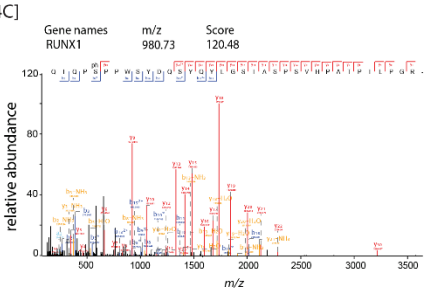
[S4A]



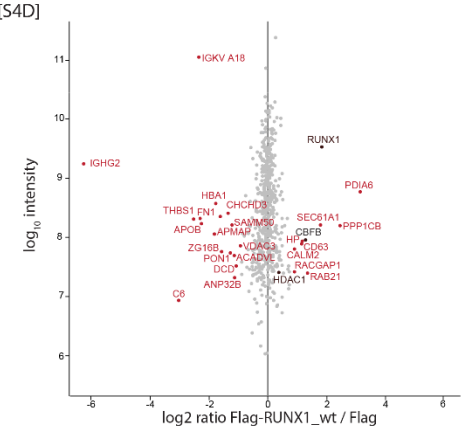
[S4B]



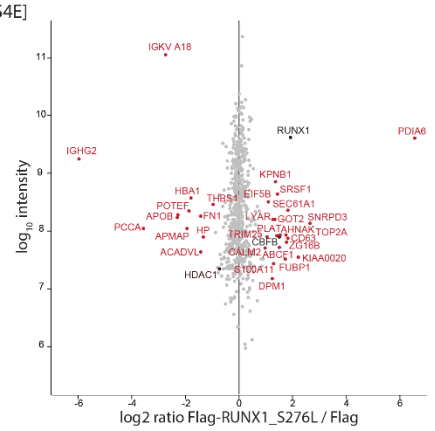
[S4C]



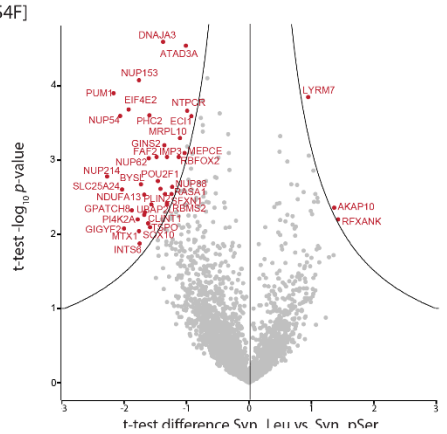
[S4D]



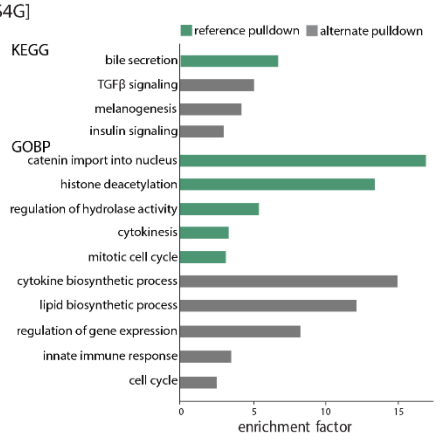
[S4E]



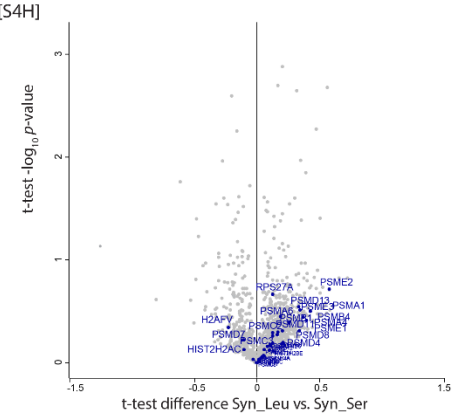
[S4F]



[S4G]



[S4H]



[S4I]

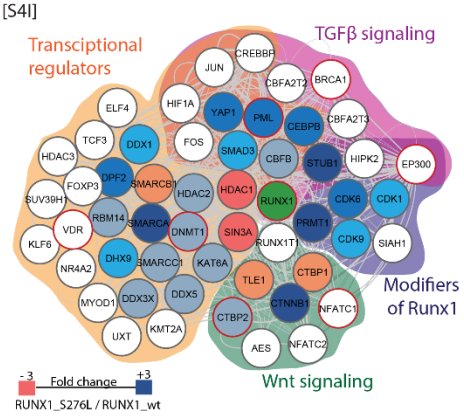


Figure S4: Loss of a known phosphorylation leads to a changed interactome and altered transcriptional activity of RUNX1. [S4A] PSM count for A375 and SkMel28 of identified peptides of RUNX1 to visualise protein coverage. [S4B] and [S4C] Deconvoluted MS/MS spectrum of phosphorylated reference [S4B] and alternate [S4C] peptide of RUNX1 identified by high resolution mass spectrometry. [S4D] and [S4E] Interaction proteomics screen in A375 RUNX1_KO cells stably overexpressing FLAG-tagged RUNX1_wt or FLAG-tagged RUNX1_S276L. SILAC protein expression (log2) of RUNX1_wt [S4D] or FLAG-tagged RUNX1_S276L [S4E] relative to the corresponding control cell line (FLAG tag only). RUNX1 and its core binding factor CBFC are marked in black. Significantly up and down regulated proteins are highlighted in red. Results represent three replicates per experiment group. [S4F] Volcano plot of synthetic alternate peptide (Syn_Leu) versus synthetic reference phosphopeptide (Syn_pSer) pulldowns of A375 cells. Fold change of ratios between Syn_Leu and Syn_pSer (log2) are plotted against p-value (-log10) (n=3). Black lines indicate the significance threshold based on student t-test (FDR < 0.01; S0=1.2; n=3). Significantly up and down regulated proteins are highlighted in red. [S4G] One-dimension annotation enrichment of KEGG pathways and GOBP for reference (Syn_Ser; green) and alternate (Syn_Leu; grey) peptide pulldown. The enrichment score calculated by Fisher exact test were plotted against the p-value (-log10) (FDR < 0.1). [S4H] Volcano plot of synthetic alternate peptide (Syn_Leu) versus synthetic reference peptide (Syn_Ser) pulldowns of A375 cells. Fold change of ratios between Syn_Leu and Syn_Ser (log2) are plotted against p-value (-log10) (n=3) based on student t-test (FDR < 0.01). Transcriptionally regulated proteins by RUNX1 based on Reactome annotation are highlighted in blue. [S4I] Protein-protein interaction network for RUNX1 based on BioGRID. Mutated genes identified by exome sequencing are circled in red. The node colour correlates with the ratio between RUNX1_S276 and RUNX1_wt. Nodes with white colour are not identified in this study. Enriched pathways are coloured as indicated.

3.3 Manuscript III

Schmitt, M., Nalpas N. C., Sinnberg, T., Niessner, H., Forchner, A., Garbe, C., Macek, B.

Individualized proteogenomics reveals mutational landscape of melanoma patients in response to immunotherapy

In preparation, 2020

Author contributions:

The study was designed by Tobias Sinnberg, Heike Niessner, Nicolas C. Nalpas, Boris Macek and me and all experiments were performed at the Proteome Center Tuebingen, University of Tuebingen. Tobias Sinnberg and Heike Niessner collected the tumor samples and generated the patient-derived xenografts. Cell culture, sample preparation, proteomics and phosphoproteomics analysis, interaction assays were performed by me under the supervision of Boris Macek. The data was analyzed and interpreted by Nicolas C. Nalpas and me under the supervision of Boris Macek. Claus Garbe and Andrea Forchner contributed in fruitful discussion of the manuscript.

A list of all used materials and extended method description can be found as Supplementary Information on the provided CD.

References of the manuscript are included in the main references.

Individualized proteogenomics reveals mutational landscape of melanoma patients in response to immunotherapy

Marisa Schmitt[1], Nicolas C. Nalpas[1], Tobias Sinnberg[2], Heike Niessner[2], Andrea Forschner[2], Claus Garbe[2], Boris Macek[1]

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

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

Keywords: Proteogenomics, Melanoma, Mass Spectrometry, Immunotherapy, Whole Exome 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

Abstract

Immune checkpoint inhibitors are used to restore or augment antitumor immune response and show great promise in treatment of melanoma and other types of cancers. However, only a relatively small percentage of patients are fully responsive to immune checkpoint inhibition, mostly due to tumor heterogeneity and primary resistance to therapy. Both of these features are largely driven by accumulation of patient-specific mutations, pointing to the need for personalized approaches in diagnostics and immunotherapy. Proteogenomics integrates patient-specific genomic and proteomic data to study cancer development and resistance mechanisms, as well as tumor heterogeneity in individual patients. Here, we use a proteogenomic approach to characterize the mutational landscape of samples derived from four clinical melanoma patients at the genomic, proteomic and phosphoproteomic level. Integration of datasets enabled identification and quantification of an extensive number of sample-specific amino acid variants, among them many that affect modifiable amino acid residues and were not previously reported in melanoma. We detected a disproportional number of alternate peptides between treated and untreated (naïve) samples with a high potential to influence signal transduction. Statistical analysis revealed accumulation of mutations in specific pathways within immune checkpoint inhibitor-treated and naïve samples, including PI3K-AKT and focal adhesion signaling. Several variants detected by MS affected the protein phosphorylation status; among them was the guanine nucleotide exchange factor DOCK1 that was selected for further validation. This is the first proteogenomic study designed to study the mutational landscape of patient-derived melanoma tissue samples in response to immunotherapy.

Introduction

Over the last decades, the role of the immune system controlling tumor progression has been established and new immunotherapeutic targets showed remarkable clinical activity. The reagents nivolumab and ipilimumab are immune checkpoint antibodies targeting PD-1 (the programmed cell death-1) and CTLA-4 (cytotoxic T lymphocyte-antigen-4) receptors [429,430]. PD-1 and CTLA-4 are co-inhibitory T cell receptors and act as negative regulatory receptors that block T cell activation and induce immune tolerance [431,432]. Subsequently, blockade of these receptors with antibodies demonstrated tumor rejections and a significant prolongation in melanoma patient survival [73,433]. However, only a minority of patients responded to ipilimumab and many patients developed immune-related toxicities. The complexity and multiplicity of

involved mechanisms, heterogeneity in the immune response across tumors, the tumor microenvironment and the varying tumor immunogenicity play roles in response and resistance to immune checkpoint blockade [434]. The clinical response to immune checkpoint inhibitors and resistance is often associated with a high mutational load and the number of expressed tumor neoantigens leading to antitumor immunity [80,81]. Several studies showed that deficiencies in antigen presentation and down-regulation of MHC class I (MHC-I) play a role in immune checkpoint resistance [86-88]. Besides mutation in β2-microtubulin, loss of JAK-STAT pathway results in acquired resistance due to down-regulation of MHC-1 [91,92]. Additionally, classic oncologic pathways like MAPK, PI3K-AKT or Wnt/β-catenin pathways can regulate immune response by influencing the tumor microenvironment. Alterations in the MAPK pathway may lead to increased expression of VEGF, a vascular endothelial growth factor, and other inhibitory cytokines, thus mediating evasion of tumor cells [93,94]. Constitutive activation of PI3K-AKT pathway due to loss of PTEN was associated with resistance to PD-1 therapy and decreased overall survival of patients with leiomyosarcoma [95,96]. The majority of these studies are performed at the genomic and transcriptomic level. Transcriptomic signatures of cytosolic markers and immune-related genes were associated with clinical response and outcome of patients with different therapies [435]. Melanoma sub-populations showed a heterogeneity in transcriptional processes and CDK4 and CDK6 regulated pathways were linked to resistance mechanisms in non-responder cells studied by single cell RNA sequencing. In contrast to transcriptomic studies, Harel *et al.* compared in a quantitative proteomic screen clinical melanoma samples treated with either tumor infiltrating lymphocyte (TIL)-based or anti PD-1 immunotherapy and could show an association between higher lipid metabolism and response to immunotherapy [436]. 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 individualized, i.e. do not contain sequence information specific for the individual patient. Commonly used protein databases therefore inherently prevent identification of individual non-synonymous mutations. By combining nucleotide sequencing and MS technologies, it is possible to simultaneously study and integrate DNA sequence, RNA expression and splicing, protein isoform abundance and PTMs in a personalized fashion. Genomic alterations due to non-synonymous single nucleotide variants (nsSNVs), insertions or deletions (InDels) of nucleotides, frameshifts and alternate splicing variants can alter cellular function at the protein level by modulating its abundance, localization and protein-protein interaction [437,438]. Clinical data have shown that oncogenic targets are aberrantly post-translationally modified during tumorigenesis and might be relevant as therapeutic targets [31].

The most prominent modification is phosphorylation, which is abnormally activated during tumorigenesis and may propagate dysregulated signals and cellular functions [439,440]. However, such alterations affecting the modification level in signaling molecules can be benign and insignificant. Here, we propose to use melanoma tissue from human tumors and patient-derived xenografts of the same cases in response to immunotherapy to study the mutational landscape and reconstruct signaling transduction network specific to individual samples using their corresponding genomics, proteomics and PTMs datasets.

Results

Individualized proteogenomics of melanoma in response to immunotherapy

In order to identify signatures and cellular mechanisms of immunotherapy response, we analyzed matching clinical samples including blood, tumor tissue and patient-derived xenografts (PDX) from four patients (**Figure 1A and S1A**). Two of the analyzed samples were naïve (no treatment at time point of sampling) and two patients were treated with immune checkpoint inhibitors (ICi) nivolumab and ipilimumab at the time point of surgery (**Figure S1A**). The progression-free survival (PFS) and overall survival (OS) were calculated based on the start of therapy and have differed in all samples (**Figure S1A**). The patient under therapy with the ID 111 showed a shorter PFS and OS compared to others. Only one patient presented the well characterized $BRAF^{V600E}$ mutation, however all patients showed NRAS mutations (G12V, Q61R, A146T, F156L) at different sites, which is the second-most mutated gene in melanoma. For proteogenomic analysis, we performed whole exome sequencing (WES) from snap-frozen tumor tissue and matching blood samples, allowing detection of germline and somatic nucleotide variants (**Figure 1A**). Using our in-house online tool PCTi, we generated individualized protein sequence databases that contain the WES-identified non-synonymous variants (**Figure 1A**). Among all non-synonymous nucleotide variants detected by WES (ca. 23,000), more than half were unique to one of the four patients, whereas only 15.9% were identified in all four samples (**Figure 1B)**. Number and type of nucleotide variants detected by WES were similar across all four samples (**Figure 1B and C**). In this context, the vast majority were substitutions, most of which have been previously reported (ca. 85%). The rest of the nucleotide variants were classified as frameshifts, deletions and insertions; these also followed a similar reporting status as the substitutions (**Figure 1C**). Next, we compared the WES analysis of blood and tissue samples in order to distinguish between germline and somatic nucleotide variants (**Figure 1D**). Here, we observed an approximate 1:10 ratio of somatic to germline

nucleotide variants. The non-synonymous mutations were incorporated using the PCTi tool into the corresponding protein isoform, thus providing protein sequence databases that are individualized for each patient. The tumor tissue was injected into immune-deficient mouse in order to generate patient-derived xenografts (**Figure 1A**). From the PDX material, proteins were extracted and separated by high pH reverse chromatography. Peptides were either measured immediately by LC-MS/MS or subjected to phosphopeptide enrichment using TiO_2 chromatography. In addition to the PDX material, we also performed proteomics of FFPE material of the same patient. The proteomic and phosphoproteomic datasets of each sample were processed against the human reference and individualized protein databases, as well as the Uniprot mouse database to assess the mouse contamination. Identified peptides were divided into three classes based on taxonomic classification: class I contains peptides annotated exclusively as human, class II are shared between human and mouse, and class III are annotated exclusively as mouse. Nearly all of the proteins and mutations were identified in class I and II and used for further analysis (**Figure 1E)**. In total in class I and II, we identified over 9,500 proteins and 120,000 sequence-specific peptides per sample (**Figure 1E, Supplementary Table 1 – 4**). The phosphoproteomic analyses revealed the identification of over 9,000 phosphorylation sites for samples 101, 111 and 129, while over 5,000 phosphorylation sites were detected for the sample 110 (**Figure 1E, Supplementary Table 1-4**). Across the different clinical samples, between 18% and 28% contained an entry in the sample-specific databases for mutated protein isoforms, including single amino acid variants, insertions, deletions and frameshifts. Up to 27% of these mutations were detected by high resolution mass spectrometry. The samples 110 (12.16%) and 111 (11.86%) showed the highest proportion of mutated isoforms compared to the overall identified proteins. In addition, we confirmed the gain of a new phosphorylation site for four protein isoforms (PCM1, TANC1, CLDN23, CTNND1) by MS by identifying the phosphorylated alternate peptide (**Figure 1D**). From the 1,687 detected reference peptides, several were confirmed by MS to be phosphorylated on the mutation site including five protein isoforms showing the phosphorylated reference peptide as well as the alternate peptide harboring the non-modifiable residue (HELZ2, LMO7, MIA2, TTN, MDC1, PLEC, ZFYVE19, ARGEF40) (**Figure 1D**). The difference of identified variant peptides detected by MS is unlikely due to technical artifacts at the WES level, since the sequencing depth was similar across samples; thus allowing for the comparison of mutations across samples (**Figure S1B**). Interestingly, the MaxQuant-derived score and intensity of identified peptides were similar between reference and alternate peptides, highlighting the overall good quality of

the mutation identifications (**Figure S1C and Figure S1D**). Together our results highlight the importance of individualized approaches in order to investigate patient specific tumors.

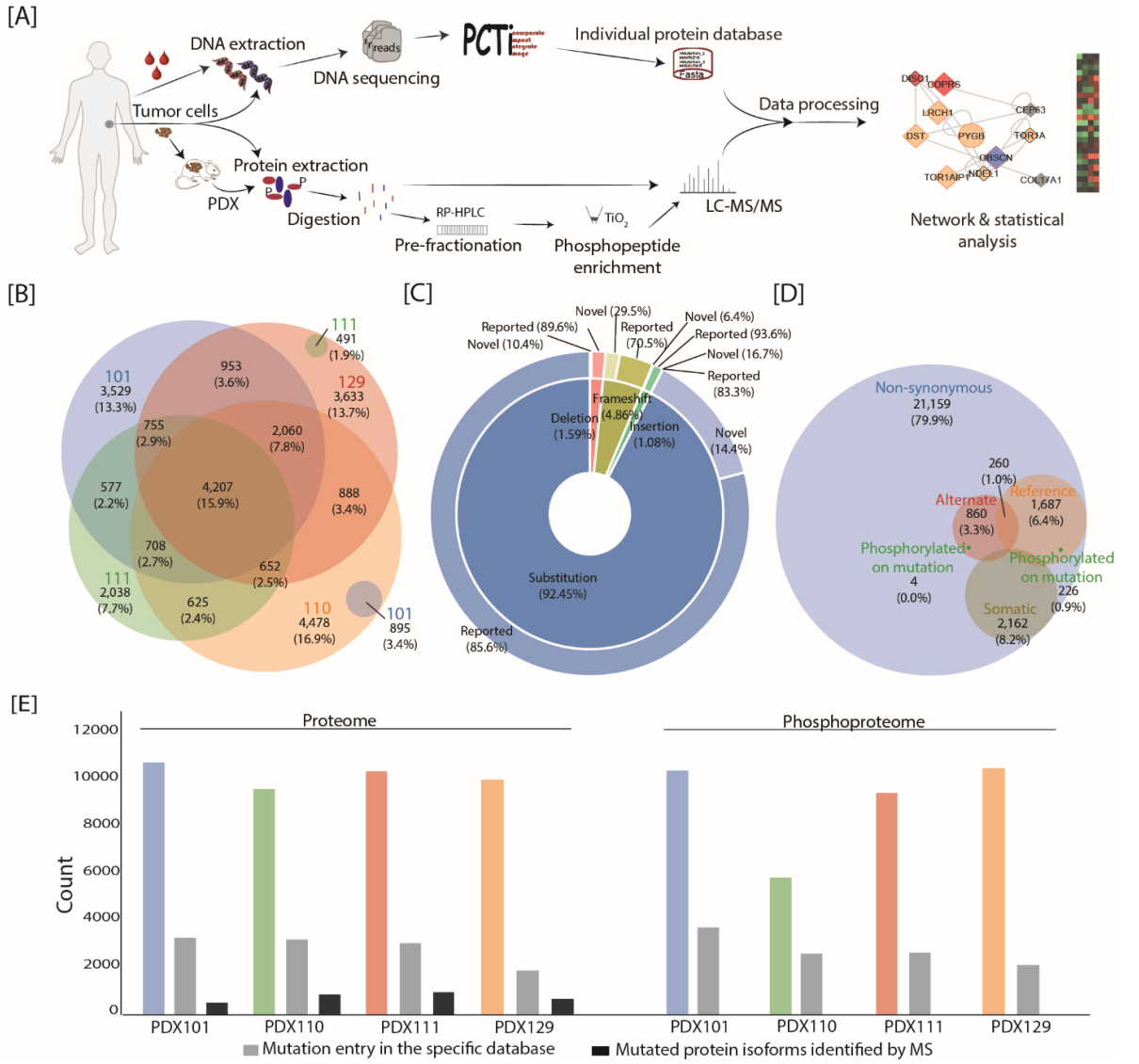


Figure 1: Proteogenomics integrates mutational landscape of melanoma cells. [A] Schematic overview of the proteogenomic workflow. Whole blood and tumor tissue of four patients were used in this study. Metastatic tumor tissue was injected into immune-deficient mouse to generate patient-derived xenografts (PDX). For whole exome sequencing, DNA was extracted from whole blood and metastatic tissue and sequenced on Illumina sequencing instrument. Individualized protein databases and impact files were generated with PCTi online tool. For the proteomic workflow, FFPE specimens from the same tissue as well as PDX tissue was used. Cells were lysed and proteins were digested using trypsin. The resulting peptide mixture from PDX material was fractionated using an off-line RP HPLC operated at high pH. Fractions were pooled and measured directly or applied to phosphopeptide enrichment using titanium dioxide (TiO_2) prior to LC-MS/MS. MS raw data was processed with MaxQuant software and analysed by PCTi. [B] Overlap of non-synonymous nucleotide variants identified by WES of four melanoma patients (tumor tissue and blood). [C] Inner donut depicts the type of all non-synonymous nucleotide variants identified by WES including substitution, insertions, deletions and frameshifts. Outer donut represents the proportion of novel nucleotide variants identified in this study. [D] Overlap of somatic (brown) proportional to all identified nucleotide variants (blue) of all samples, as well as overlap with alternate (red) and reference (orange) peptides identified by MS and peptides phosphorylated on variant site (green). [E] Identified protein groups and phosphorylation sites by MS for each sample. Identified proteins with a mutation entry in the database are depicted in grey and identified mutated protein isoforms by MS are shown in black.

Proteogenomics highlights accumulation of mutations in specific pathways linked to immunotherapy

To address mutations accumulating in specific pathways, we performed pathway enrichment analysis of identified proteins by MS harboring amino acid variants based on KEGG and Reactome resources (**Figure 2A and B**). In both naïve samples, the enriched pathways were similar and included ribosome biogenesis, apoptotic cleavage and caspase-mediated cleavage of protein and focal adhesion. While the over-representation of JAK-STAT signaling and RUNX3 regulated transcription were exclusively identified in the naïve samples 101 and 110, respectively (**Figure 2A**). The enriched pathways for immune checkpoint inhibitor (ICi) treated samples differed and only the pathways signaling by GTPase and Type I hemidesmosome assembly were identified in both samples. For the sample 111 from the patient with the short PSF and OS, VEGF and PI3K-AKT signaling as well as activation of IFN production were enriched. Whereas, G-alpha (12/13) signaling events, DAP12 signaling and signaling by MET were over-represented in mutated proteins for the sample 129 (**Figure 2B**). Next, we focused on the sample-specific mutations on alternate peptides and phosphopeptide exclusively detected in one of the samples (**Figure 2C**). We identified approximately 125 alternate peptides in the samples 101 and 129 and over 300 and 390 alternate peptides in the samples 110 and 111, respectively. Notably, several phosphorylated alternate peptides were detected by mass spectrometry with a potential to change the modification status of the protein. For the sample with ID 111, 35 identified alternate peptides were phosphorylated and six of them showed a phosphorylation event on the mutation site. These comprised three reference phosphorylated peptides, implying a loss of the phosphorylation site due to the mutation, and phosphorylated alternate peptides, involving a gain of a new phosphorylation site. For example, substitutions on the protein isoforms ZFYVE19 (S366A) and WNK1 (T1056P) showed a loss of phosphorylation sites and the phosphorylation site on the key regulator of abscission step in cytokinesis ZFYVE19 is known to get phosphorylated by the kinase NEK3. The gain of a new phosphorylation site on CTNND1 (N52S) and CLDND1 (P232S) were confirmed to be phosphorylated by mass spectrometry and the PCTi software predicted that the kinases p38MAPK and CDK5/GSK3 could act on these new kinase motifs, respectively. Moreover, two frameshifts resulting in the loss of known phosphorylation sites and gains of new phosphorylation sites were detected on the protein isoform INST1 (p.N773TfsX0) and CASP9 (p.L151SfsX0), resulting in a changed kinase motif from PLK1 to GSK3 and CAMK2G to PKC, respectively. In addition, we performed pathway enrichment of proteins containing sample-specific alternate peptides and showed accumulation of mutations in specific pathways

(**Figure 2D**). The pathway positive regulation of MAP kinase activity was enriched in both ICi-treated samples, whereas cytokine production and regulation of PI3K-AKT signaling were over-represented in the sample 111 and FGFR2b ligand binding and activation and Rho GTPase cycle in the sample 129. cAMP signaling pathway and EGFR tyrosine kinase inhibitor resistance were enriched for sample-specific variants in the sample 110 and cellular senescence and immunoregulatory interactions in the PDX101. These findings demonstrating the power of proteogenomic data integration to detect sample-specific mutations and their accumulation in specific pathways.

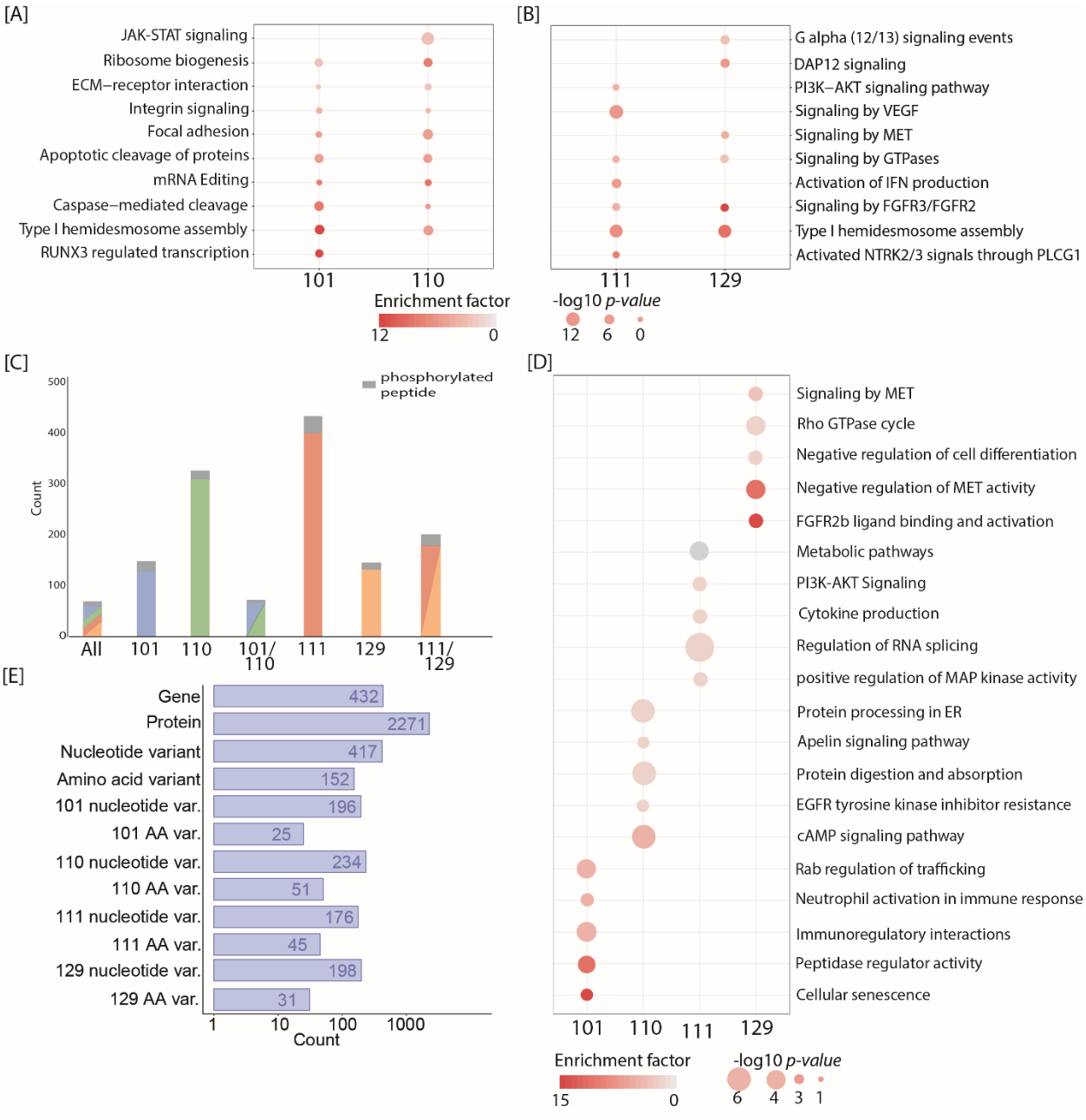


Figure 2: Proteogenomics highlights differing mechanisms of mutated proteins between naïve and treated samples. [A] and [B] Dot plot of over-represented pathway based on proteins containing identified alternate peptides for naïve samples 101 and 110 [A] and immune checkpoint inhibitor treated samples 111 and 129 [B]. Results are based on Fisher-Exact test (FDR ≤ 0.1). [C] Sample-specific alternate peptides and phosphopeptides that are exclusively identified by MS. [D] Dot plot of over-represented pathway based on proteins containing sample-specific alternate peptides and phosphopeptides. Results are based on Fisher-Exact test (FDR ≤ 0.1). [E] Histogram summarizing the content of the PI3K-AKT signaling pathway in terms of number of genes, proteins and identified mutations.

While the PI3K-AKT signaling pathway is known to be activated in melanoma cells in response to immunotherapy [96], our analysis shows that other complementary tumorigenesis –related pathways including GTPase signaling are also activated in immune checkpoint inhibitor treated samples due to specific genomic and proteomic alterations, representing alternative treatment opportunities. Visualization of our proteogenomic datasets on the PI3K-AKT signaling pathway (KEGG) showed that the vast majority of genes harbor one or more nucleotide variant in at least one of the clinical samples (**Figure S2A**). Interestingly, among these nucleotide variants identified by WES, between 13% and 26% of the corresponding amino acid variants could be confirmed by MS depending on the clinical samples (lowest in 101, highest in 111, **Figure 2E**). Similar results were observed for the regulation of actin cytoskeleton pathway, which is more specific of the naïve samples (**Figure S2A and B**). Notably, we observed a very high coverage of MS-identified amino acid variants in this pathway (between 26% and 43% confirmed by MS (**Figure S2B)**. Here, we showed that proteomics supplements genomics with protein expression profiles of mutated molecules, thus highlighting key signaling pathways involved in melanoma progression and resistance.

PDX samples are comparable to matching FFPE tumor samples

In addition to the PDX samples, we performed proteomics of FFPE material from the same patients. Notably, between 60% and 70% of FFPE proteins were also identified in the PDX samples (**Figure S2C**), which supports the use of PDX as a model to study cancer progression. Approximately 60% of the mutated protein isoforms in the FFPE specimens were also identified the PDX samples (lowest in 101, highest in 111). Next, we addressed the differences between samples and sample types of the overall categories immune response, metabolism, signal transduction and ECM interaction based on iBAQ intensity (log10) (**Figure S2D**). For immune response and signal transduction no significant differences between samples and sample types were observed, however, proteins in both categories for the PDX samples 111 and 129 were slightly higher compared to the naïve samples. Interestingly this was even based on less identified protein isoforms compared to 110. Protein isoforms involved in metabolism were significantly higher in the FFPE sample 111 compared to other FFPE samples and in the FFPE

samples of 110 and 129 the iBAQ intensities for ECM interaction were significantly higher compared to other samples. The same trend was observed for both categories in the corresponding PDX. These results show that human proteomic and phosphoproteomic profiles recapitulate in PDX models.

Proteogenomics detects several unreported mutations in melanoma with high potential to influence signal transduction

We then compared the naïve and treated samples based on the identification and abundance of shared alternate peptides (**Figure 3 and Supplementary Table 5**). These MS-identified mutations were stratified using our PCTi application based on their potential impact on cellular signal transduction. Among the alternate peptides shared across all, naïve-only or treated-only samples, most of the detected mutations were single amino acid variants and not described previously in melanoma. However, all the shown alternate peptides in this figure besides the insertion (p.(149_160)insA_A) on the protein RPL14 were described in other databases like COSMIC or dbSNP previously. We detected 41 mutated peptides shared between all samples, including known oncogenic genes like MSH6, DOCK1 or SYNE2. Most of the detected mutations had a loss or a gain of a (phosphorylatable) S/T/Y amino acid residues, of which only three had been previously reported in melanoma. Mutated protein isoforms were enriched in immune response, metabolism and GTPase Rho activation (**Figure 3A**). Among the 42 alternate peptides shared between naïve samples, two showed an insertion of 5 or 11 amino acids on the proteins OXA1L (p.(478_483)insS_S) and RPL14 (p.(149_160)insA_A), respectively (**Figure 3B**). In addition, more than 40% of the alternate peptides had a loss or gain of S/T/Y amino acid or belonged to protein encoded by known oncogene or tumor suppressor genes. Several of the mutated peptides shared between the naïve samples were involved in immune related response including the Reactome pathways innate and adaptive immune system, interleukin signaling and cytokine signaling. Interestingly, 166 alternate peptides are shared in immunotherapy treated samples and showed an accumulation of mutations in metabolism, signal transduction and cell cycle, including the oncogenic genes FN1, AKAP13 and ENO3 (**Figure 3C**). Among these mutated peptides, two insertion and one deletion mutations were observed on the proteins IRF2BP2 (p.(93_99)insQ_Q), SAAL1 (p.(6_11)insP_P) and PHLDA1 (p.(Q190_Q204)del). These results postulate the mutational landscape of melanoma cells in response to immunotherapy and the accumulation of mutations in specific pathways due to the treatment.

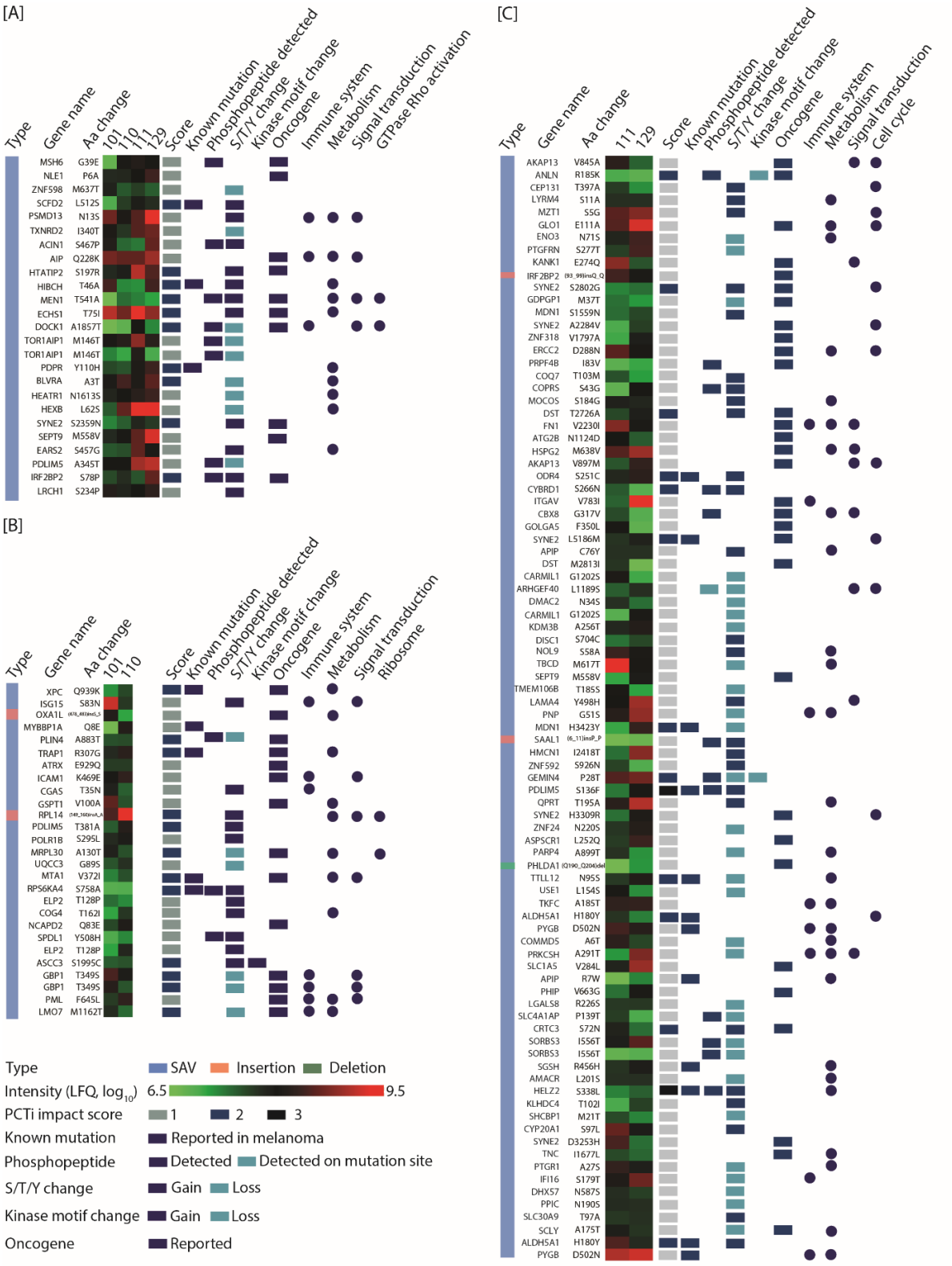


Figure 3: Impact of identified alternate peptides for all, naïve and immune checkpoint inhibitor treated samples. MS-identified alternate peptide shared between [A] all patients, [B] naïve patients only and [C] immune checkpoint inhibitor treated patients only. Across all panels, the different tracks indicate for each alternate peptide the mutation types, the peptide intensity per sample, the predicted impact score, whether the mutation is known or not in melanoma, whether the peptide is phosphorylated or not, whether the mutation results in loss/gain of S/T/Y and/or kinase motif, whether the affected proteins are encoded by a known oncogene or tumor suppressor gene and whether the affected proteins belong to immune system, metabolism, signal transduction and/or ribosome functional categories.

Integration of genomics, proteomics and drug database prioritizes actionable mutations

The protein isoforms harboring the 249 identified amino acid variants identified by MS shared in all, naïve-only and treated-only samples were used to generate a protein-protein interaction network (**Figure S3A**). The protein-protein interactions were retrieved from BioGrid and the drug-protein interactions were obtained from DrugBank. While most of the amino acid variants that originated this network were shared in treated samples only, the majority of identified proteins were actually shared among all samples. This suggest an accumulation of mutations rather than a change of proteome in treated samples compare to the naïve samples. Notably, just over ten protein isoforms harbor a mutation with a medium to high impact score, which highlights their putative role in context of protein phosphorylation-based signal transduction networks in melanoma.

We then prioritized the mutated proteins in this network through a five-dimensional scatter plot (**Figure S3B**). The dimensions of importance for prioritization were the protein abundance, the number of interactors per protein, the impact score (PCTi), whether the amino acid variant per protein was identified by MS and whether the protein is druggable. This strategy highlighted a cluster of 12 proteins with an intermediate impact score and one additional protein with a high impact (IRF2BP2). They were characterized by a median intensity (log_{10}) of 10.1 and a median of 6 interactors. Among these, three proteins were druggable, i.e. RPS6KA4, ECHS1 and ALDH5A1. Furthermore, four proteins harbor a mutation leading to gain of a phosphorylation site, i.e. PLCG1, MDN1, ALDH5A1 and DOCK1, and most proteins are encoded by known oncogene or tumor suppressor genes, with the exception of RPL14, MDN1, PDPR, RPS6KA4 and ALDH5A1. These results show how multiple levels of proteogenomic information can be integrated in order to prioritize and extract proteins that are most likely to rewire signal transduction network in context of melanoma.

Changed interactome due to gain of a new phosphorylation site on DOCK1

We decided to further investigate one striking mutation that fulfilled the following criteria: 1) has one of the highest impact scores; 2) belongs to over-represented pathways among the treated samples; 3) has the potential to influence the modification status of the protein and 4) the mutation was identified by mass spectrometry. The guanine nucleotide exchange factor (GEF) protein DOCK1 was identified to be mutated in all PDX samples and harbors a mutation from alanine to threonine at position 1857, resulting in kinase motifs for PKG or CAMII **(Figure 4A)**. The mutation is located in a highly modified region of the protein and is in close proximity to the PXXP motif, important for the binding of CRK and GRB2. A second gain of a modifiable

amino acid at position 1792 (A1792T) was detected by genomics at the binding site of NCK2. Abnormal regulation of DOCK1 is associated with various diseases, including melanoma [441]. DOCK1 is involved in RAS signaling and thereby regulates the well characterized MAPK and PI3K-AKT signaling pathways[198]. In addition, DOCK1 also regulates cytoskeletal rearrangements, Rho GTPase signaling and RTK signaling such as VEGF signaling [441-443].

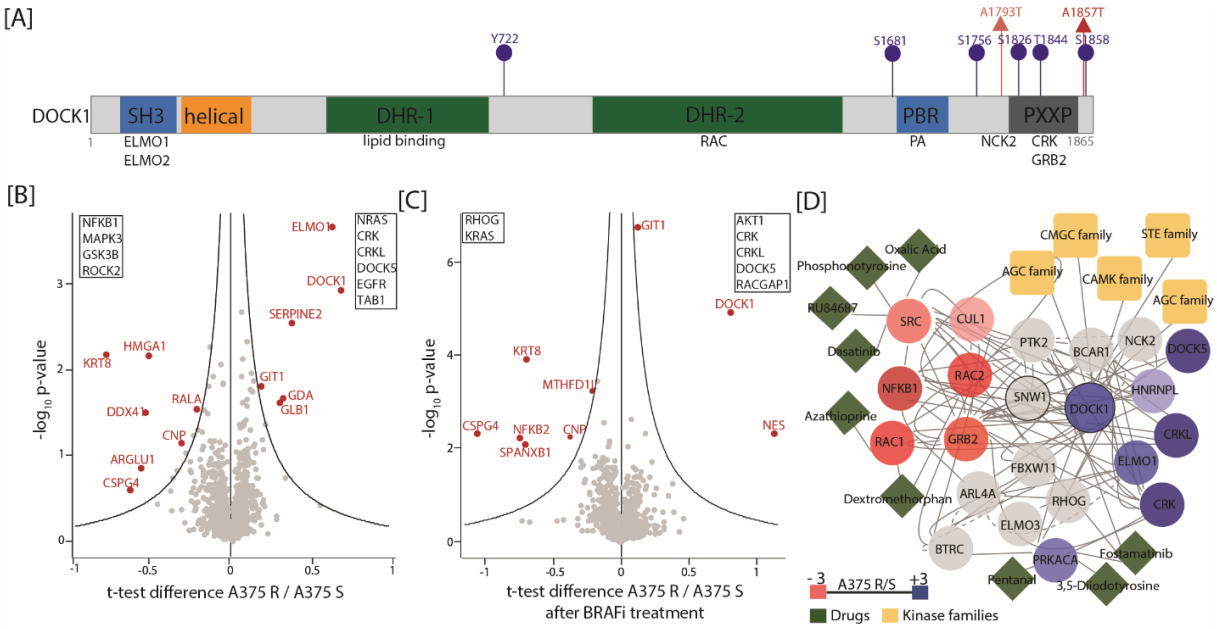


Figure 4: Gain of a new phosphosite on GEF protein DOCK1 leads to changed interactome. [A] Schematic overview of domains, mutations and identified phosphorylation sites on GEF protein DOCK1. Numbers indicate amino acid position of domains and mutation, phosphorylation sites. The known interaction partners are described under the domain. [B] and [C] Volcano plot of immunoprecipitated pulldowns of DOCK1 in A375 R versus A375 S cells in absence [B] and presence [C] of BRAF inhibitor vemurafenib treatment (2 µM PLX4720). T-test difference (log_2) are plotted against t-test p-value ($-log_{10}$) (t-test, Benjamini-Hochberg FDR ≤ 0.05, S0 = 0.9). Results are of four independent replicates. [D] Protein-protein interaction network for DOCK1 based on BioGRID. Mutated genes identified by exome sequencing are circled in black. The node color indicates the protein ratio between immunoprecipitation of DOCK1 in A375 S and R cells. Predicted kinase are depicted in yellow. Nodes with grey color are not identified in this study. Known drugs based on DrugBank are depicted in green.

To validate the impact of the mutation on the interactome of DOCK1, we performed immunoprecipitation of DOCK1 in drug-sensitive (S) and drug-resistant (R) melanoma cells using a DOCK1 antibody in absence and presence of BRAF inhibitor (BRAFi) vemurafenib (**Figure 4B and C**). While we found DOCK1 mutated in all pulldown experiments, the phosphorylation on the mutation site showed a higher occupancy in A375 R (64%) compared to A375 S cells suggesting an enriched phosphorylation status of DOCK1 in drug-resistant cells. Both the alternate peptide and the phosphorylated alternate peptide were identified by MS. The known interaction partners, ELMO1 and GIT1 were highly enriched in pulldowns of A375 R cells compared to drug-sensitive cells (**Figure 4B**). The formation of DOCK1 and ELMO

complex is crucial to achieve RAC activation, which regulates actin cytoskeletal remodeling, organization and function [444]. The ARF-GTPase-activating protein GIT1 and the RAS-related protein RALA were significantly enriched in A375 R and A375 S, respectively; both of which are involved in adhesion, proliferation and cytoskeletal organization. Interestingly, several known interaction partners of DOCK1 and signaling proteins were exclusively identified in either pulldowns, including NRAS, CRK and CRKL for A375 R pulldowns and NFKB, MAPK3 and GSKB for A375 S pulldowns. In the presence of BRAFi, GIT1 was even stronger enriched in pulldowns of A375 R compared to A375 S, while the opposite was true for NFKβ (**Figure 4C**). Next, we tested whether the significantly regulated and exclusively detected proteins in the separate pulldown experiments revealed an over-representation of KEGG and Reactome pathways (**Figure S4A and B**). Focal adhesion, MAPK and TGF-beta signaling and regulation of autophagy pathways were enriched in A375 R, whereas cytokine receptor interaction, MAPK and ErbB signaling and actin cytoskeletal regulation pathways were over-represented in A375 S cells. Following treatment of cells with BRAFi, some of the over-representation results changed, for example the cytokine receptor interaction and p53 signaling pathways were enriched in pulldowns of A375 R and S cells, respectively. Using the BioGrid protein-protein interaction database, we observed several differences in DOCK1 interactome (i.e. protein abundance) between pulldowns of A375 R and S cells (**Figure 4D**). The interactome also includes approved drugs (based on DrugBank database) to possibly counteract the DOCK1-dependent activation, notably several drugs are available for interaction partners SRC and RAC1. In addition, we added the putative kinase families that are predicted to bind to mutated DOCK1 and may phosphorylate the gained modifiable residue. Our results show a novel mutation on DOCK1, leading to gain of a modifiable amino acid, which we confirmed phosphorylated in A375 melanoma cell line. Our pulldown experiments also suggest that this mutation influences the interactome of DOCK1.

Discussion

Here, we present the individualized proteogenomic landscape of four melanoma samples of patients in response to immunotherapy. This study is, to our knowledge, the first integrative individual proteogenomic analyses of melanoma tumor tissue and matching PDX in response to immunotherapy. Malignant melanoma is predominantly studied by genomics and transcriptomics, and more recently by proteomics [445]. As the majority of drugs target proteins, proteomics allows extensive and quantitative surveys of the global proteome in order to select

targeted treatment and predict drug response in tumor therapy. However, proteomics is not individualized and publicly available databases do not contain cancer- and sample-specific mutations. The integration of genomic and proteomic data, called proteogenomics, is a powerful tool to study the mutational landscape in context of the protein abundance and post-translational modifications. Several genomics and transcriptomic studies revealed the mutational landscape and heterogeneity of melanoma cases [41,446,447]. A recent quantitative proteomic screen of melanoma patient's tumors in response to immunotherapy revealed the link between lipid metabolism and response to immunotherapy [436]. Lobas *et al.* used a proteogenomic approach to study eight melanoma cell lines; their analysis allowed discrimination between the specific cell lines based on their variant peptide profiles [448]. Our dataset is generated from a small cohort of four samples, however 14% of our identified nucleotide variants were not reported previously. We identified an extensive number of shared as well as sample-specific alternate peptides by whole exome sequencing and mass spectrometry. The identifications are in the same range or even better in comparison to other proteogenomic datasets of human cancer tissue [356,362,380]. The detected alternate variant peptides were of high quality and had MaxQuant-derived score similar to the reference variant peptides (**Figure 1 and S1**). We also did not observe a change in intensity or posterior error probability distribution between reference and alternate variant peptides, which may have been indicative of reduced quality. Several pathways were over-represented in either naïve or immune checkpoint inhibitor treated samples based on accumulation of mutations (**Figure 2**). Interestingly, the over-represented pathways in naïve samples were overall similar, whereas they differed between the IC treated samples. For example, integrin signaling and focal adhesion was observed in both naïve samples and signaling by VEGF in sample 111 and signaling by FGR2/FGR3 in sample 129. In addition, the ICi treated samples showed over-representation in the PI3K-AKT signaling and GTPase signaling pathways, highlighting potential drug-treatment options in melanoma. In this study, we identified several known mechanisms involved in response to immunotherapy including JAK-STAT and PI3K-AKT signaling, signaling by VEGF and activation of IFN production [91,449]. In addition, we showed an accumulation of sample-specific mutations on key proteins of these specific pathways. (**Figure 2 and S2**). Several individual markers were identified that can be used to predict putative drug treatments in specific patients. Based on our result, we can postulate that additional pathways, such as signaling by MET, FGFR3/FGFR2 signaling and RUNX3 regulated transcription pathways are involved in development of resistance. However, additional work is required to validate these proteogenomic data.

This study shows that the proteogenomic signatures of patient derived xenografts (PDX) confirm most findings from melanoma cancer patients and can serve as model to study cancer mutational landscapes as a source of tumor tissue closely resembling the clinical lesions. PDX are generated directly from tumor tissue and overcome several limitations over the use of monolayers of cells (cell lines), which is based on the selective proliferation of clonal cells. PDXs keep the histological features, genomic signatures and genetic heterogeneity of cells in a tumor mass. In addition, PDX tumors provide enough material to also perform phosphoproteomics. However, PDX have limitations that must be considered prior to data analysis, indeed it can take up to 6 months to generate PDX and samples are usually highly contaminated with mouse cells (due to sample preparation). In this study, the iBAQ intensities of mutated protein isoforms involved in metabolism were significantly higher in the patient with the shortest PSF and OS compared to all other. The same trend was observed in matching PDX, which highlights the fact that proteogenomic signatures of PDXs recapitulate most of the findings in human tumor samples. Harel *et al.* also identified the association between metabolism and response to immunotherapy in their proteomic study [436]. Besides signaling pathways, the microenvironment plays an important role in melanoma progression and response to immunotherapy. Intensities of mutated proteins involved in ECM interaction were also significantly higher in the samples 110 and 129 compared to other human tissue samples suggesting an involvement of ECM reorganization in cancer cells.

We experimentally validated one example of a gain of a new phosphorylation site on DOCK1 and showed that the phosphorylation status of this mutation site has an impact on the interactome of DOCK1 in vermurafenib-sensitive and -resistant cells. The GEF protein DOCK1 is mutated in 15% of melanoma patients and highly expressed in RAS-driven cancer types [198,445,447]. In this study, we identified the mutation from alanine to threonine in all samples (homozygous), which was also characterized by a high impact possibly rewiring signal transduction. The reference peptide harboring the not modifiable residue at A1856 was reported previously in PhosphoSitePlus to be phosphorylated in several studies. This highlights the use of proteogenomics to refine databases and generate sample-specific databases. Pinto *et al.* showed in a SILAC-based quantitative proteomic screen that the phosphorylation site at position T1857 was significantly up-regulated in interleukin 33 (IL-33)-stimulated compared to unstimulated RAW264.7 cells. Their study revealed that actin cytoskeletal reorganization was overrepresented in significantly changing proteins and thus was involved in immune response [450]. Here, we also identified DOCK1 to be highly phosphorylated in drug-resistant cells compared to sensitive cells and the occupancy for this specific site (T1857) suggest a

different protein phosphorylation status between phenotypes. In addition, in the pulldown experiments we identified a distinct interactome of DOCK1 and several key signaling molecules enriched in either sensitive or resistant cells. For example, ELMO1 was significantly enriched in A375 R cells and CRK was exclusively identified in pulldowns from A375 R cells. Both proteins are known to form a complex with DOCK1, thereby recruiting DOCK1 to the plasma membrane [190,444]. The binding to DOCK1 enhances its GEF activity and leads to RAC activation and subsequent downstream signaling. Our results suggest a DOCK1 dependent activation of focal adhesion and MAPK signaling pathways in vermurafenib resistant cells, whereas cytokine receptor interaction and ErbB signaling pathway are activated in sensitive cells. Further experiments need to be done to compare reference and mutated version of DOCK1 in context of BRAF and checkpoint inhibition to highlight the clinical relevance of this mutation site in melanoma. Taken together, we postulate that the phosphorylated mutation changes the interactome of DOCK1.

Conclusion

Individualized proteogenomics allows the detection of sample-specific variants at the genome, proteome and PTM level. Here, we studied the mutational landscape of four clinical samples in response to immunotherapy. Our dataset might serve the scientific and melanoma community as a resource of clinical genomic, proteomic and phosphoproteomic profiles, which is still sparse in melanoma. Our approach revealed accumulation of mutations in specific pathways in naïve and ICi-treated samples, such as over-representation of PI3K-AKT signaling or activation of Rho GTPase signaling in treated samples. We also confirmed the loss or gain of several phosphorylation events on mutation sites that possibly lead to rewiring of cell signal transduction. We further validated the mutation A1857T on the GEF protein DOCK1 and our data suggest that this gain of a new phosphorylation site has an impact on the interactome of the protein.

Materials and Methods

Skin metastasis were collected during surgery and compared to blood. In total, we analyzed four metastatic and four “control” samples. In addition, primary tissues were injected into mice to obtain patient-derived xenografts (PDX). The use of human tissue from an internal biobank

was approved by the local ethical committee (781/2018BO2) and experiments were performed in accordance with the declaration of Helsinki Principles.

Generation of patient-derived xenografts

To generate PDXs, tumor tissue was finely minced using the cross-blade technique, digested in nevi solution (HBBS (w/o Ca2+ and Mg2+) with 0.05% collagenase, 0.1% hyaluronidase and 0.15% dispase) and filtered through a 100 µm cell strainer. The melanoma cell suspension was implanted with Matrigel (Corning Life Sciences) subcutaneously in NSG (NOD.Cg-PrkdcscidIl2rgtm1Wjl/SzJ) mice, leading to patient derived xenografts (PDX). Tumor grafts were harvested when they reached a diameter of 10 to 15 mm, digested as above, resuspended in Biofreeze medium (Biochrom/Merck) and 1 ml per cryotube of the cell suspension was frozen for short-term cryoconservation in −80 °C and for long-term storage in liquid nitrogen.

Protein extraction from patient-derived xenografts

Cell lysis of snap-frozen patient-derived xenografts (PDX) was performed with lysis buffer (6 M urea, 2 M thiourea, 10 mM Tris-HCl pH 8.0) supplemented with protease inhibitor (complete Mini EDTA-free tablets, Roche) and phosphatase inhibitor buffers (5 mM glycerol-2-phosphate, 5 mM sodium fluoride, and 1 mM sodium orthovanadate). Glass beads (zirconia/glass beads 0.23 mm, Carl Roth GmbH) were added and cell lysis was performed in a BeadBug microtube homogenizer (3 cycles, 1 min at full speed, Sigma-Aldrich). Cell extracts were centrifuged at 13,000 rpm for 20 min and proteins were purified by acetone precipitation and subjected to tryptic digestion prior LC-MS/MS analysis.

Protein extraction from formalin-fixed paraffin embedded tissue preparation

FFPE tissue were first de-paraffinized by two washes in xylene (5 min, 50°C) followed by three serial washes in ethanol (100%, 95% to 70%) for 10 min each. Ethanol was removed completely and sections air-dried. Lysis was carried out in 4% SDS, 50 mM DTT, 100 mM HEPES pH 7.5 supplemented with protease inhibitor at 95°C for 60 min and by sonication for 15 min. Proteins were purified by acetone precipitation and subjected to tryptic digestion prior LC-MS/MS analysis.

Sample preparation for MS analysis

Purified protein pellets were dissolved in lysis buffer (6 M urea, 2 M thiourea, 10 mM Tris-HCl pH 8.0), reduced using 100 mM DTT, and alkylated using 50 mM iodoacetamide followed by pre-digestion using endopeptidase Lys-C (Lysyl Endopeptidase, Wako Chemicals) for 3 h.

After diluting the sample to 2 M Urea with 10 mM ammonium bicarbonate, proteins were digested into peptides using sequencing grade trypsin (1 µg per 100 mg protein, Promega Corporation) overnight. Peptides were then acidified with 1% TFA and then either purified on C18 stage tips (as described previously) or purified on Sep-Pak C18 Cartridge (Waters) and eluted in 80% ACN for high pH reverse phase chromatography.

Immunoprecipitation of DOCK1 in melanoma cell lines

Drug-sensitive (S) and drug-resistant (R) melanoma cells A375 were grown in RPMI-1640 medium containing 100 U/ml penicillin/streptavidin and 10% FBS. A375 R cells were grown in the presence of 2 µM PLX4720 (vemurafenib-analog, Selleckchem). Cells were seeded to 80% confluence and treated with PLX4720 or DMSO for 3 h. Cell lysis was performed in pulldown buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 1 mM EDTA, 0.5% Triton X100) supplemented with protease inhibitors (complete Mini EDTA-free tablets, Roche)and phosphatase inhibitor buffers (5 mM glycerol-2-phosphate, 5 mM sodium fluoride, and 1 mM sodium orthovanadate) on ice for 10 min. Cell lysates were precleared for 1 h at 4°C with washed Pierce Protein G magnetic beads (Thermo Fisher Scientific) using 5 µl per mg lysate. DOCK1 antibody (MA5-15010, Thermo Fisher Scientific) was coupled to beads by incubation at 4°C for 20 min in incubation buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 1 mM EDTA, 0.5% Triton X100). The beads were washed three times with DPBS to remove unbound peptides. Precleared cell lysates and DOCK1 antibody coupled to Pierce Protein G magnetic beads were incubated for 2 h at 4°C while shaking. Pierce Protein G magnetic beads were used as control and all pulldowns were performed four times. Beads were washed three times with incubation buffer and two times with DPBS. Proteins were eluted by incubation at 95°C for 10 min in NuPAGE LDS sample buffer (Thermo Fisher Scientific). Proteins were separated on a NuPAGE Bis-Tris 4-12% gradient gel (Thermo Fisher Scientific) and stained with Coomassie Brilliant Blue solution (Bio-Rad Laboratories). Gel lanes were cut into small pieces and washed three times with washing buffer (5 mM AmBiC, 50% ACN) to remove Coomassie stain. To reduce disulfide bonds, 10 mM DTT in 20 mM AmBiC was added and incubated at 56°C for 1 h. After alkylation of the disulfide bonds with IAA (55 mM in 20 mM AmBiC) for 30 min at RT, gel pieces were washed with washing buffer and dehydrated with 100% ACN and vacuum centrifugation (10 min). Proteins were digested with trypsin (12.5 ng/ml in 20 mM AmBiC, Promega Corporation) at 37° overnight. Digested peptides were extracted from gel pieces with 3% TFA in 30% ACN, followed by 0.5% acetic acid in 80% ACN and 100% ACN. All extracts were combined, concentrated by vacuum centrifugation and purified on C18 StageTips.

High pH reverse phase chromatography

High pH reverse phase chromatography was conducted using an Ultimate 3000 HPLC (Thermo Fischer Scientific) equipped with xBridge BEH130 C_{18} 130A, 3.5 µm, 4.6 x 250 mm column (Waters) as described previously [415]. In brief, peptides were eluted with an 80 min gradient generated from solvent A (5 mM NH_4OH) and solvent B (5 mM NH_4OH, 90% ACN) at pH 10. Fractions were collected in the first 60 min of the gradient and concatenated into 30 pools followed by vacuum centrifugation. Peptide pools were resuspended in 500 µl 80% ACN, 10 µg of the pool were concentrated and desalted on StageTips prior LC-MS/MS measurements for proteome analysis.

Phosphopeptide enrichment

Phosphopeptides were enriched using TiO_2 beads (Titansphere, 10 µm, GL Sciences) as described previously. 1 mg of beads (in 80% ACN, 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.

Liquid chromatography- mass spectrometry

LC-MS/MS runs were performed on EASY-nLC 1200 UHPLC (Thermo Scientific) coupled to Q Exactive HF and HFX Orbitrap mass spectrometers (Thermo Scientific). The peptides were separated on 20-cm analytical HPLC-columns (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)) using a water-acetonitrile gradient of 60 min and 90 min for proteomic/immunoprecipitated samples and phosphoproteomic sample fractions, respectively. The FFPE samples were measured twice with 60 min and 130 min gradient. Gradients were generated by solvent A (0.1% formic acid) and solvent B (80% ACN in 0.1% acetic acid) with a flow rate of 200 nl/min at 40°C. Peptides were ionized by nanoelectrospray ionization at 2.3 kV and a capillary temperature of 275°C. For high pH proteomic fractions, FFPE samples or immunoprecipitated samples, each full spectrum, acquired with 60,000 resolution (automated control target of 3e6; fill time 25 ms for Q Exactive HF and 20 ms for Q Exactive HFX), was followed by 12 tandem MS (MS/MS) spectra, where the 12 most abundant multiply charged ions were selected for MS/MS sequencing with a resolution of 30,000, an automated control

target of 1e5, an injection time of 45 ms, and collision energy of 27% for Q Exactive HF and 28% for Q Exactive HFX. Phosphopeptide enriched samples, full MS scans were acquired with a resolution of 60,000 (AGC target 3e6, fill time 25 ms). The seven most abundant multiply charged ions were selected for MS/MS sequencing with a resolution of 45,000 on Q Exactive HFX and 60,000 on Q Exactive HF, an AGC target of 1e5 and a fill time of 220 ms.

DNA extraction and sequencing from blood and snap-frozen primary tissue

For sample ID 110 and 129, genomic DNA was extracted from blood and snap-frozen primary tissue using GeneElute mammalian genomic DNA isolation kit (Sigma-Aldrich) according to manufactures' instructions with slight modifications. Human snap-frozen tissue was incubated in lysis solution C solution at 55°C overnight, whereas blood samples were incubated for 10 min. DNA was purified on GeneElute MiniPrep columns and eluted with nuclease-free water. For sample ID 101, genomic DNA was isolated by c.ATG Core Facility in Tuebingen using the QIAamp DNA Mini (QIAGEN) kit as recommended by the manufacturer.

At the c.ATG Core Facility in Tuebingen, the genomic DNA from each sample were assessed for quantity and quality on Nanodrop spectrophotometer (ThermoFisher Scientific), Qubit Fluorometer (ThermoFisher Scientific) and Bioanalyzer (Agilent) instruments. The exome captures and libraries were prepared using Sureselect XT Human All Exon V7 Low Input kit (Agilent) with dual indexing according to manufacturer's instructions. The resulting libraries were sequenced on a NovaSeq 6000 instrument (Illumina) using S2 FlowCell (200 cycles). Exome sequencing data for sample ID 111 was retrieved from DKTK master trial [451].

Exome sequencing data analysis

Raw sequence data were processed using an in-house pipeline developed at the Proteome Center Tuebingen. The raw reads were initially quality checked using FastQC software [452]. Illumina adapters and 5'/3' low quality bases were trimmed from reads using Trimmomatic [453]. Paired-end reads from individual libraries were then aligned to *H. sapiens* reference genome (GRCh38) using the BWA aligner [454]. Reads resulting from PCR duplication were marked using Picard package. Germline variants were called using the GATK HaplotypeCaller workflow, while the somatic variants were identified using the GATK Mutect2 workflow [416]. Variants were recalibrated for score and filtered (soft-filter) using GATK. SnpEff software was used to perform the annotation and functional effect prediction of detected variants [455].

Generation of personalized protein databases for MS analyses

We then used our in-house application, named PCTi (currently under review), in order to generate personalized protein sequence databases. Briefly, 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). Each sample called nucleotide variants were injected into the corresponding reference transcript nucleotide sequences and then *in silico* translated. The computed information was reported directly within the alternate FASTA header to facilitate interaction with the rest of PCTi (e.g. mutation positions, reference ID). The output consists in two FASTA files containing reference protein sequences and sample-specific alternate protein sequences, which are used as protein databases for processing of LC-MS/MS data.

Prediction of the detected variants biological impact

We used PCTi application to predict the mutations impact. Here, the focus was on the impact of amino acid variants on protein phosphorylation-based signal transduction networks in melanoma. Each reference/alternate protein sequence was annotated based on (1) whether phosphorylation sites (S/T/Y) were lost and/or gained (IRanges package); (2) loss/gain of kinase motifs [51]; (3) loss/gain of known human phosphorylation sites [382,419]; (4) loss/gain of known mutation sites in melanoma [397]; (5) being encoded by oncogenes or tumour suppressor genes [344,420,421]; (6) levenshtein similarity between reference and alternate protein sequences. Each impact was scored with the application default. A summed score was calculated for each alternate sequences' amino acid, and the maximum summed score was reported per mutated isoform. The mutated protein isoforms were then ranked to allow prioritization for follow up studies.

Mass spectrometry data analysis

The LC-MS/MS data were searched against PCTi *H. sapiens* reference (99,354 entries) and individualized alternate databases (101 = 29,104 entries; 110 = 40,041 entries; 111 = 40,041 entries; 129 = 40,041 entries), as well as UniProt *H. sapiens* (release 2019/02/13; 95,943 entries) databases and commonly observed contaminants using the Andromeda search engine integrated into MaxQuant software (version 1.5.2.8) [285]. The PDX samples were also searched against UniProt *M. musculus* (release 2019/02/13; 95,943 entries) database. 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%. The PDX and FFPE samples were quantified using intensity based absolute quantification (iBAQ). For label-free quantification of interaction studies, a minimum of two peptides was required.

Statistical analyses and data visualization

Statistical analyses were performed with Perseus software suite (version 1.6.5.0). The taxonomy for each ENSEMBL protein ID was annotated after filtering of all reverse and potential contaminants hits. A list of all protein and phosphorylation sites identifications for each sample are provided in **Table 1-4**. The impact file generated in PCTi was mapped to the peptide identification for each PDX in order to stratify the mutations impact (**Supplementary Table 5**). For mutated protein isoforms in class I and II pathway over-representation was performed. The resources used for annotation of proteins were Gene Ontology (GO), Biological Processes (GOBP), GO Cellular Compartment (GOCC), GO Molecular Functions (GOMF) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Reactome Pathway database (Reactome). The fisher exact test (FDR $\leq$ 0.5) was used to checked for over-represented categories among mutated protein isoforms and shared variant peptides. A list of all over-representation results is provided in **Supplementary Table 1-4**.

Venn diagrams to show the overlap between identified nucleotide variants from WES analysis and between protein identifications from PDX and FFPE material were performed with the online tool https://www.stefanjol.nl/venny. Box plot analysis of iBAQ intensities of mutated protein isoforms were prepared in GraphPad Prism and Excel. Statistical analysis was performed with two-tailed unpaired t-test in GraphPad Prism. P values < 0.05 were considered statistically significant, with * for $p < 0.05$, ** for $p < 0.01$. We also generated interaction network within the R programming environment [417]. These networks were generated using protein-protein (using BioGRID database), drug-target (using DrugBank database) and predicted kinase-substrate (PCTi results) interactions; e.g. DOCK1 network [344,424]. The generated networks were exported (using igraph and RCy3 packages) into Cytoscape for further customisation [425].

For proteomic interaction studies, label-free quantification normalization data were used after log2 transformation. The data were filtered to retain only proteins with valid values in at least 70% of the samples. Following that, two sample Students t-test was performed with Benjamini-Hochberg-based FDR p-value threshold of 0.05 and S0 of 0.9. The following gene annotation

were added: GOBP, GOMF, GOCC and KEGG The pathway annotation enrichment was performed on the fold change with Benjamini-Hochberg FDR p-value < 0.5. A list of all over-representation results is provided in **Supplementary Table 6**.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

N.C.N., M.S., T.S., H.N. and B.M. designed the study. M.S. performed the proteomics experiments. T.S. and H.N. collected the tumor samples and generated the patient-derived xenografts. M.S. and N.C.N. analyzed 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 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).

Supplementary Information

Schmitt, M., Nalpas N. C., Sinnberg, T., Niessner, H., Forchner, A., Garbe, C., Macek, B.

Individualized proteogenomics reveals mutational landscape of melanoma patients in response to immunotherapy

In preparation, 2020

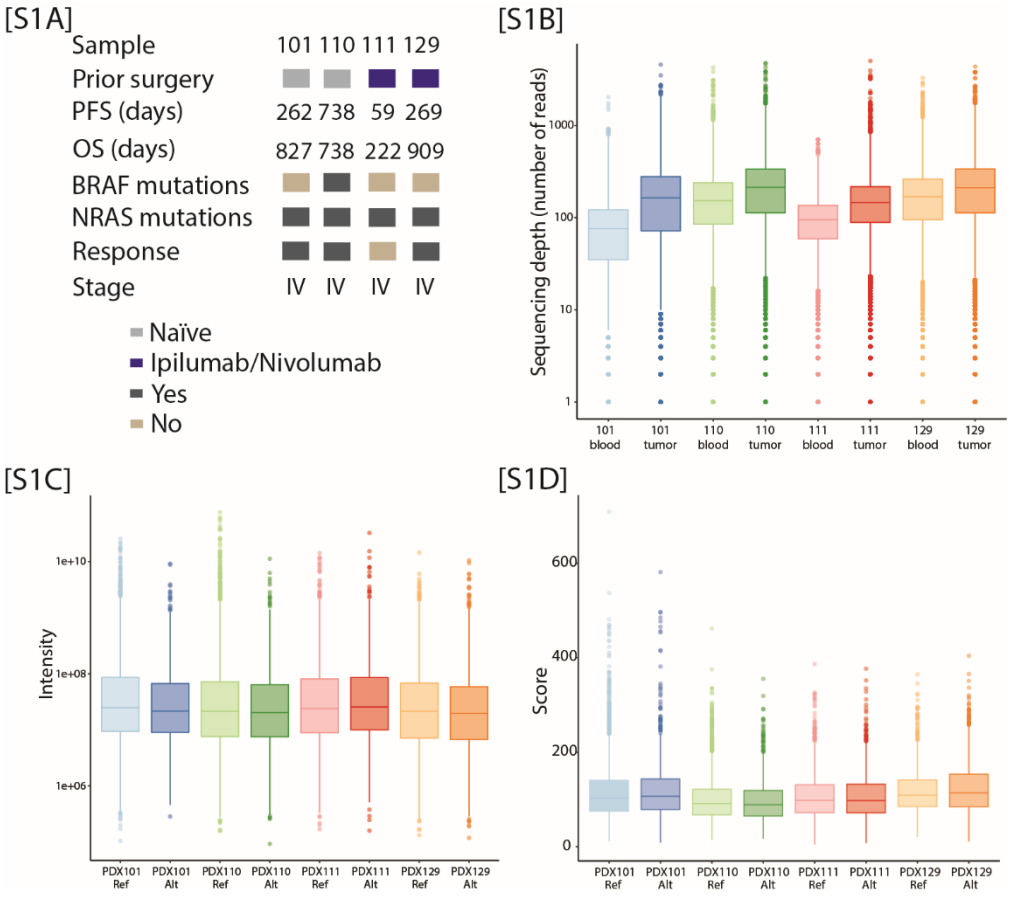
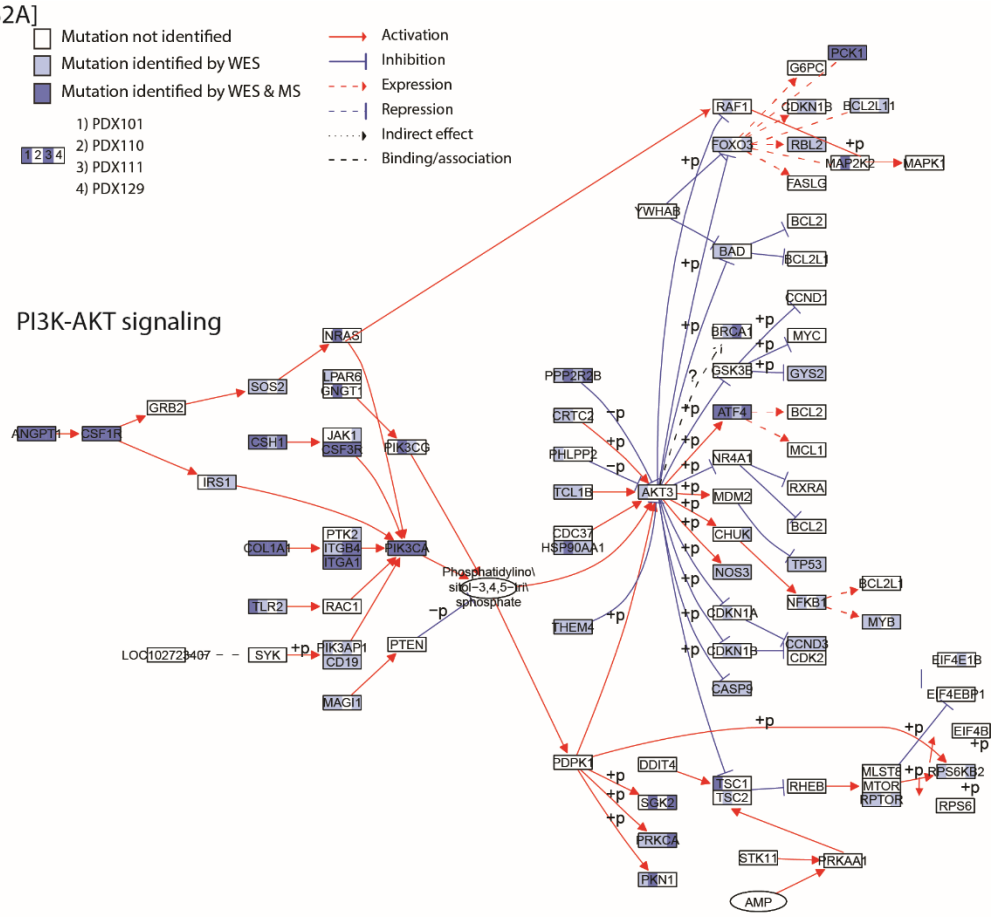
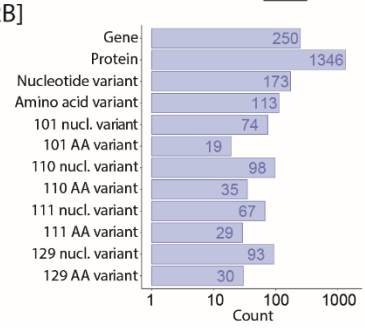


Figure S1: Proteogenomics integrates mutational landscape of melanoma cells. [S1A] Clinical information of used samples. The progression-free survival (PFS) and overall survival (OS) was calculated based on start of therapy and numbers indicate days after therapy start. [S1B] WES sequencing depth for each patient and sample type. [S1C] MS-measured intensity of identified reference (Ref) and alternate (Alt) peptides for each patient. [S1D] The MaxQuant score of reference (Ref) and alternate (Alt) peptides identified by MS for each sample.

[S2A]


Mutation not identified
Mutation identified by WES
Mutation identified by WES & MS
1) PDX101
2) PDX110
3) PDX111
4) PDX129
Activation
Inhibition
Expression
Repression
Indirect effect
Binding/association
PI3K-AKT signaling
NRAS
SOS2
GRB2
ANGPT1
CSF1R
LPAR6
GNGT1
CSH1
JAK1
CSF3R
PIK3CG
IRS1
COL1A1
PTK2
ITGB4
ITGA1
PIK3CA
TLR2
RAC1
LOC102723407
SYK
PIK3AP1
CD19
MAGI1
PTEN
Phosphatidylinositol-3,4,5-trisphosphate
PPP2R2B
CRTC2
PHLPP2
TCL1B
CDC37
HSP90AA1
THEM4
AKT3
PDPK1
DDIT4
SGK2
PRKCA
PKN1
RAF1
FOXO3
YWHAB
BAD
BRCA1
GSK3B
ATF4
NR4A1
MDM2
CHUK
NOS3
CDKN1A
CDKN1B
CASP9
PCK1
G6PC
CDKN1B
BCL2L11
RBL2
MAP2K2
MAPK1
FASLG
BCL2
BCL2L1
CCND1
MYC
GYS2
BCL2
MCL1
RXRA
BCL2
TP53
BCL2L1
NFKB1
MYB
CCND3
CDK2
EIF4E1B
EIF4EBP1
EIF4B
RPS6KB2
RPS6
TSC1
TSC2
RHEB
MLST8
MTOR
RPTOR
STK11
PRKAA1
AMP
actin cytoskeleton signaling
Acetylcholine
CHRM2
GNG12
KNG1
FGD1
CDC42
PAK1
DIAPH3
WASL
BAIAP2
ENAH
ACTB
ARPC3
PFN1
TMSB4X
VCL
PIP4K2A
SCIN
APC
ARHGEF4
SSH3
CFL1
LIMK1
MYLK
F2
F2R
CXCL12
CXCR4
GNA12
TIAM1
LPAR1
1-acyl-sn-glycerol 3-phosphate
PIK3CA
ARHGEF7
ARHGEF6
DOCK1
QGAP1
RAC1
ARHGEF7
GIT1
PXN
EGF
EGFR
INS
SOS2
NRAS
CYFIP1
NCKAP1
ABI2
BRK1
WASF1
ARPC3
WASF2
FN1
ITGA1
PTK2
BCAR1
CRK
RHOA
DIAPH2
BAIAP2
ACTB
PFN1
SRC
RHGAP35
MYH9
MYL5
BRAF
MAP2K2
MAPK1
ROCK2
PPP1R12A
SLC9A1
MSN
PIP4K2A
ACTN1
VCL

[S2B]


Gene 250
Protein 1346
Nucleotide variant 173
Amino acid variant 113
101 nucl. variant 74
101 AA variant 19
110 nucl. variant 98
110 AA variant 35
111 nucl. variant 67
111 AA variant 29
129 nucl. variant 93
129 AA variant 30
1
10
100
1000
Count

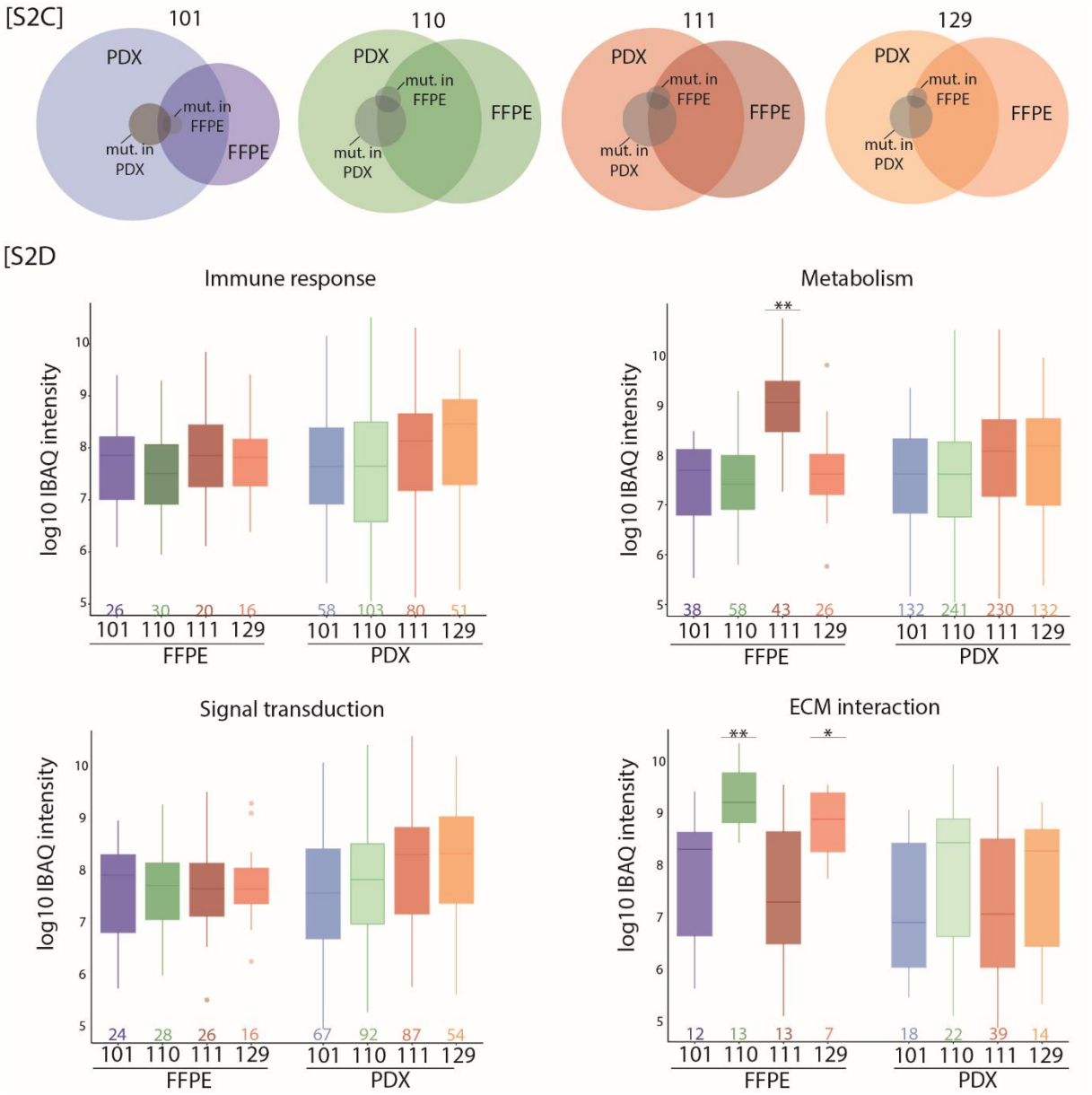
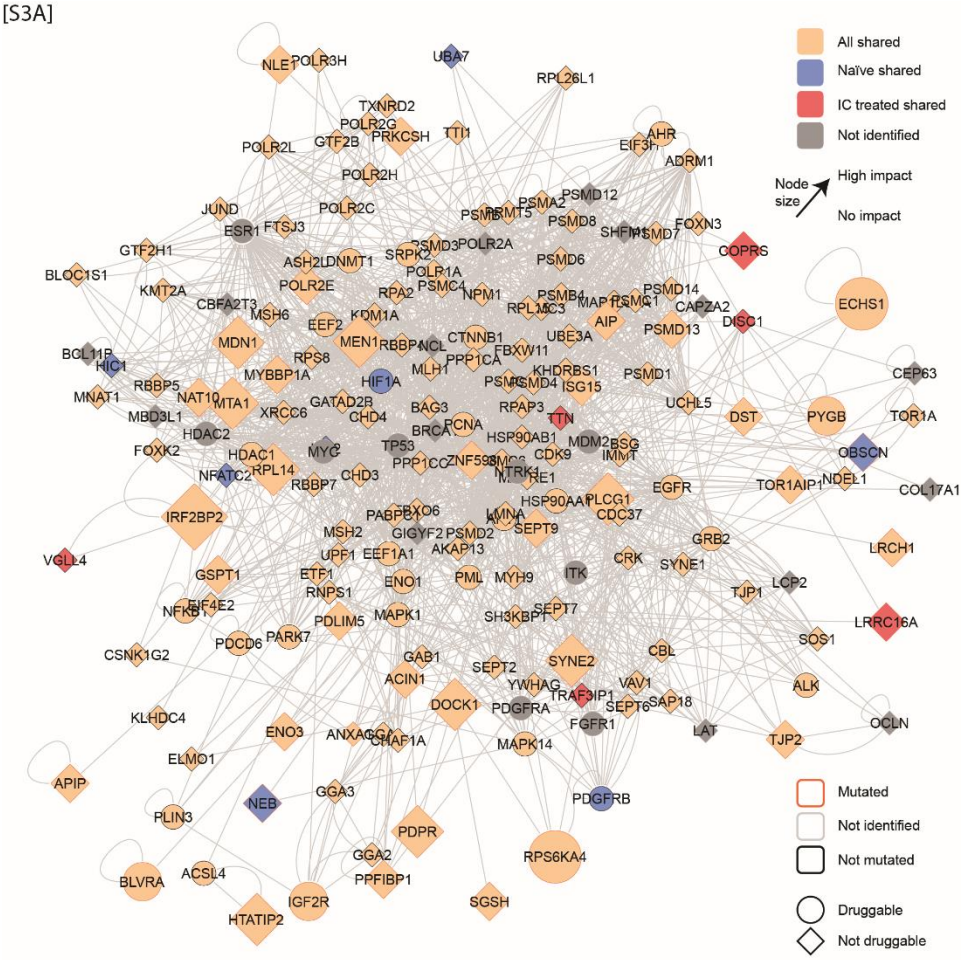


Figure S2: Proteogenomics highlights differing mechanisms of alternate protein isoforms between naïve and treated samples. [S2A] Visualization of WES- and MS-identified alternate variants (WES only = light blue, WES and MS = dark blue) for each patient on the PI3K-AKT signaling pathway (KEGG) and actin cytoskeleton pathway (KEGG). [S2B] Histogram summarizing the content of the regulation of actin cytoskeleton pathway in terms of number of genes, proteins and identified mutations [S2C] Overlap of identified protein groups and proteins containing alternate peptides (grey) for each sample. [S2D] Box plot of over-represented categories for each sample. The MaxQuant iBAQ intensity (log10) is plotted for each identified alternate protein isoform.

[S3A]


All shared
Naïve shared
IC treated shared
Not identified
Node size
High impact
No impact
Mutated
Not identified
Not mutated
Druggable
Not druggable

[S3B]

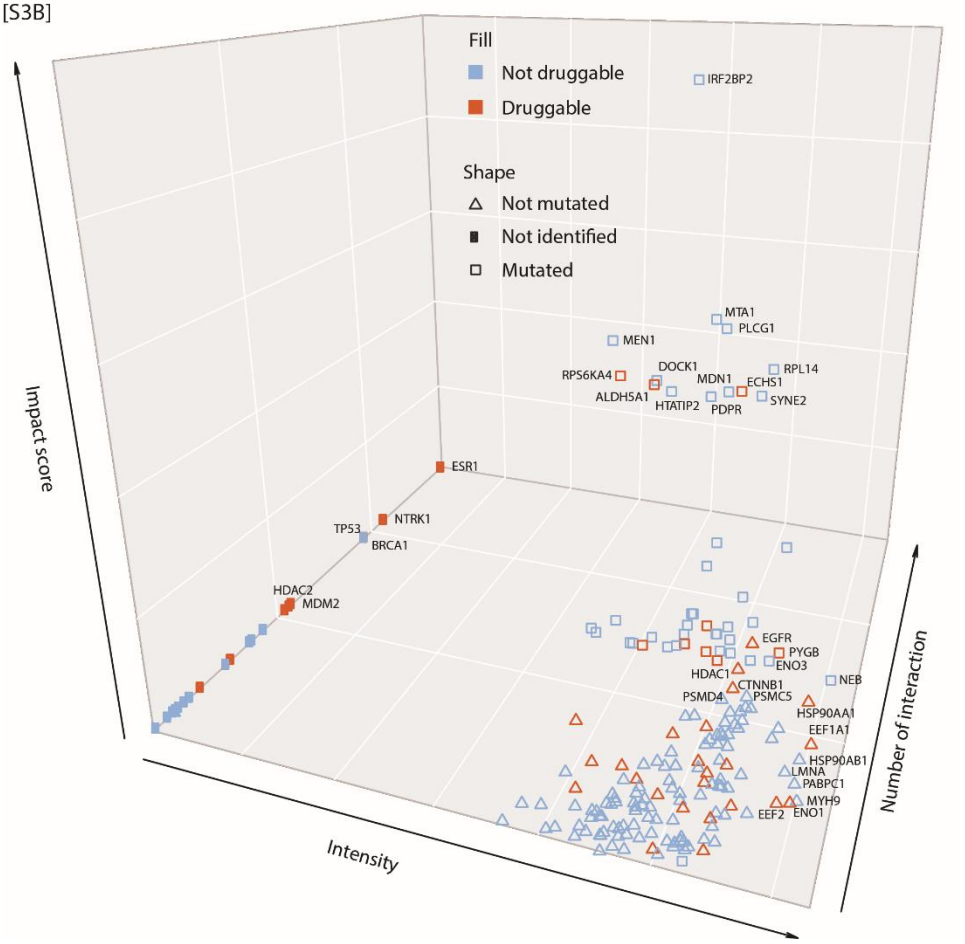

Fill
Not druggable
Druggable
Shape
Not mutated
Not identified
Mutated
Impact score
Intensity
Number of interaction

Figure S3: Interactome of proteins identified with alternate variant peptides. [S3A] Protein-protein interaction network for the MS-identified alternate isoform based on BioGRID database. Mutated genes identified by exome sequencing are circled in black. The node fill color indicates whether the isoform was identified in all, naïve only, IC treated only or none of the patients. The node edge color indicates whether the mutation was identified based on MS. The node shape shows which isoforms can be targeted by a drug treatment based on DrugBank database. The node size represents the predicted impact score (computed through PCTi application). [S3B] Each protein isoform from the interaction network is plotted on a 5-dimensional scatter plot to allow ranking. The dimensions represent the MS intensity (x-axis), the number of interactions (y-axis), the predicted impact score (z-axis), whether the isoforms are drug targets (color-code) and whether the mutation was identified based on MS (shape-code).

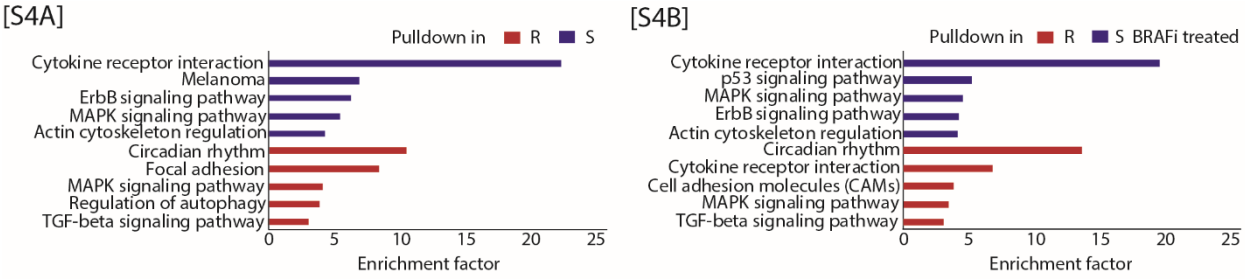


Figure S4: Gain of a new phosphosite on GEF protein DOCK1 leads to changed interactome. [S4A] and [S4B] One-dimension annotation enrichment of KEGG and Reactome pathways for pulldown of DOCK1 in A375 S and R cells in absence [S4A] and presence of BRAF-inhibitor vermurafenib [S4B]. The enrichment score calculated by Fisher exact test were plotted (Benjamini-Hochberg FDR < 0.5).

4 Discussion

Parts of this chapter were taken from Schmitt M. et al., Quantitative proteomics links the intermediate filament nestin to resistance to targeted BRAF inhibition in melanoma cells. Mol Cell Proteomics (2019), Schmitt M. et al., Integration of Individualised proteogenomics datasets to analyse single amino acid variants in malignant melanoma (Under revision) and Schmitt M. et al., Individualized proteogenomics reveals mutational landscape of melanoma patients in response to immunotherapy (In preparation).

The aim of this thesis was to identify proteomic changes and associated mutations in malignant melanoma in response to targeted and immunotherapy. First, quantitative (phospho)proteomics analyses of drug-sensitive and -resistant melanoma cells were utilized to identify additional resistance mechanisms and new molecular targets that can overcome resistance against targeted therapy. Changes in proteomic profiles revealed the down-regulation of cytoskeletal proteins in vemurafenib-resistant cells, including the intermediate filament nestin. CRISPR/Cas9 knockout of the nestin gene was associated with an invasive phenotype and resistance to targeted therapy. In the second part, an individualized proteogenomics approach was established to analyze the non-synonymous mutational landscape of two malignant melanoma cell lines in context of resistance to targeted therapy. The reported data highlighted the mutational profile differences and commonalities between both cell lines and phenotypes. Validation by interaction studies in combination with proteomics showed that the interactome of RUNX1 was changed due to a loss of a known phosphorylation site. Finally, the individualized proteogenomics approach was applied to characterize human tumor samples in response to immunotherapy. Data integration allowed the identification of an extensive number of sample-specific variants, among which several showed a potential to rewire signal transduction. The analysis revealed that variants are accumulating in specific pathways, for example PI3K/AKT signaling pathways or GTPase activation, including key molecules harboring a mutation on a modifiable amino acid. The gain of a new phosphorylation site on the GEF protein DOCK1 was confirmed to change DOCK1 interactome by proteomic interaction studies.

4.1 Proteomic and proteogenomic approaches to study malignant melanoma

Melanoma accounts for 79% of skin cancer-related deaths in Germany and due to ineffective treatment options in the late stage melanoma, patients have an overall poor prognosis [1]. In about 50% of melanoma cases the non-synonymous mutation $BRAF^{V600E}$ (or more rarely

$BRAF^{V600K}$), located within the kinase domain, leads to a hyperactive form of the BRAF kinase [31]. Consequently, the mitogen activated-protein kinase (MAPK)-pathway is constitutively active and becomes the main oncogenic pathway in melanoma [32,33,100]. In 2010, a new promising class of drugs, the mutant-specific BRAF kinase inhibitors such as vemurafenib (PLX4072) or dabrafenib, were designed and resulted in inhibition of the MAPK pathway [99]. These drugs have approximately 50% response rate and result in an average survival benefit of four months [100,101]. Another drug, cobimetinib, inhibits MEK1/2 kinases and has also been shown to improve survival of $BRAF^{V600E}$-carrier individuals [108,110]. However, tumors treated with BRAF or MEK inhibitors rapidly adapt and find ways to escape the treatment by developing acquired resistance. Such resistance is characterized by a partial or complete initial reduction of the tumor followed by recurrence of metastases within a period of five to seven months [101-104]. More recently, alternative therapeutic strategies involving immune checkpoint inhibitors have been developed. Ipilimumab, a monoclonal antibody targeting the cytotoxic T-lymphocyte antigen 4 (CTLA-4), was the first therapy to show an improved survival in metastatic melanoma patient, with 15% response rate [432]. Upon binding to CD80, which is present on surface of antigen presenting cells, CTLA-4 negatively regulates T cells priming and lead to down-regulation of anti-tumor immune response. Therefore, inhibition of CTLA-4 by ipilimumab leads to increased anti-tumor response [73,432]. Other immune checkpoint inhibitors include pembrolizumab, nivolumab or atezolizumab, which target either program death-1 receptor (PD-1) or its ligand (PD-L1), and have been shown to have response rate of up to 30%. These compounds inhibit the interaction between PD-1 and PD-L1 during the effector stage of the T cell anti-tumor response, therefore limiting the immune system suppression from PD-1/PD-L1 interaction. While current trials suggest durable responses in patient under immunotherapy, there is increasing evidence pointing towards existence of innate and acquired resistance to therapy [434]. With the increasing number of kinase inhibitor or acquired resistance mechanisms reported in the literature, it is now clear that personalized medicine will be critical to effective patient therapy.

In the last decade, a revolution in the field of genomic techniques has led to the emergence of high-throughput sequencing technologies, which provide information on DNA/RNA sequence, gene structure and gene expression. Several studies identified genomic changes after disease progression in melanoma including secondary mutations, gene fusions, mechanisms of resistance and predictive biomarkers for disease stages [456]. Berger *et al.* used a systematic approach to study cancer-associated mRNA alterations by integrating transcriptomics and structural-genomic data and identified several novel melanoma gene fusions and novel read

through transcripts [446]. The results demonstrated that point mutations are major drivers of melanoma progression. In another transcriptomics studies, data from primary samples of stage I melanoma and patients undergoing immunotherapy revealed a six-class signature to predict patient's outcome based on the expression of prognostic biomarkers, such as the transcription factor *JUN* [457]. The genomic and transcriptomics landscape was analyzed to study resistance mechanism against treatment even cross-resistance between kinase inhibitors and PD-1 therapy in metastatic melanoma [447].

Mass spectrometry based proteomics is experiencing a technological revolution similar to that of the high-throughput sequencing. The current state-of-the-art proteomics workflows are capable of routine, comprehensive analysis of proteomes and post-translational modifications such as phosphorylation. In the first part of this thesis, quantitative proteomics of melanoma cell lines (A375 S and A375 R) with differing phenotypes of acquired resistance to the $BRAF^{V600E}$ inhibitor vemurafenib was performed in order to identify additional resistance mechanisms and reveal new molecular targets to overcome resistance. This study is, to our knowledge, one of the largest global (phospho)proteomic analyses assessing the differentially expressed proteins in drug-sensitive and drug-resistant melanoma cells. Our analysis revealed that several proteins involved in cytoskeletal organization and signaling were down-regulated in drug-resistant cells compared to sensitive cells including nestin, vimentin and gelsolin. Nestin has been reported for its involvement in cancer cell migration, invasion, and metastasis [148,153,458]. Quendro and colleagues showed in a large scale proteomic study that nestin and vimentin are both up-regulated in melanoma cells and tissue materials compared to control melanocytes [141]. This was confirmed in the present study in A375 melanoma cells and further showed that nestin and vimentin are down-regulated in resistant cells compared to sensitive cells. In addition, phosphorylation analysis demonstrated that key signaling proteins are phosphorylated in resistant cells including ERK1/2 (T202/Y204) and AKT (S124).

In standard proteomics approaches, peptides and proteins are identified by matching MS/MS spectra against protein databases derived from public repositories (e.g. UniProt) that are not "individualized", i.e. do not contain sequence information specific to the individual patient. Commonly used protein databases therefore inherently prevent identification of individual non-synonymous somatic mutations. By combining nucleotide sequencing and MS technologies, it is possible to simultaneously study and integrate DNA sequence, RNA expression and splicing, protein isoform abundance and PTMs in a personalized fashion. In the present thesis, two immortalized human melanoma cell lines, as well as patient derived xenografts (PDX) and patient tumor materials were used to reconstruct signaling transduction network specific to

individual samples using their matching genomics, proteomics and PTMs datasets. The individualized proteogenomics databases were generated with PCTi, an in-house software, and highlighted a disproportional impact of nucleotide variants on modifiable amino acids. Hundreds of amino acid variants were identified by high resolution mass spectrometry including single amino acid variants, InDels and frameshifts. In addition, the loss or gain of several phosphorylation events due to point mutations could be confirmed by mass spectrometry. These in turn may lead to rewiring of signal transduction networks in context of melanoma resistance to BRAFi. Statistical analysis of mutated proteins revealed significant differences between drug-resistant and drug-sensitive phenotypes and cell lines, such as over-representation of MAPK and PI3K/AKT/mTOR signaling in drug-resistant cells and YAP/TAZ stimulated gene expression in sensitive cells. This dataset is one of the first proteogenomics datasets in melanoma including genomic and phosphoproteomic data, and can be used as a resource to identify network-attacking mutations in melanoma. The proteogenomic analyses resulted in the same range (or higher) as other studies investigating amino acid variants using custom protein sequence databases [351,354]. For example, Shenykman *et al.* used a proteogenomic approach based on RNA sequencing data and identified 544 peptides containing single amino acid variants specific to a sample [352].

In the third part of the thesis, the proteogenomic approach was further applied to human tumor material undergoing immunotherapy. This study is the first integrative individual proteogenomic analysis of melanoma tumor tissue and matching PDX in response to immunotherapy. Hundreds of sample-specific mutations were identified in human tumor tissue and patient-derived xenografts by matching exome sequencing data and (phospho)proteomic data. Most of the detected mutations were not previously reported in melanoma, among which a few had a high potential to rewire signal transduction. In addition, the proteogenomic approach identified key differences between naïve and immune checkpoint inhibitor treated samples, such as over-representation of mutated proteins in PI3K/AKT signaling pathway and metabolism in context of melanoma response to immunotherapy. Harel *et al.* analyzed large-scale proteomic screen from 116 melanoma primary tissues and showed that the mitochondrial lipid metabolism was associated with tumor response to immunotherapy [436]. In the present study, pathways involved in metabolism were also enriched in responders and could be linked to mutations in MEN1, ECHS1 and PDLIM5. Most of the findings from melanoma cancer patient could be recapitulated in PDX tumors, highlighting that PDXs can serve as a model to study mutational landscape in different cancer types. This dataset demonstrates that

proteogenomics is a powerful tool to study the mode of action of disease-associated mutations at the genome, proteome and PTM level in a sample-specific fashion.

In both proteogenomic datasets, numerous mutations were identified, which were not reported previously in melanoma or other publicly available datasets, such as COSMIC or dbSNP. Therefore, these mutations could not have been identified with a standard shotgun proteomics workflow and are of fundamental biological interest for melanoma community. The predicted impact score of these mutations may help to stratify and identify sample-specific mutations for further investigations and as actionable drug-targets. All three datasets can be used as a resource to identify key (phospho)proteins and network-attacking mutations in context of kinase inhibition and immunotherapy in order to improve patient's survival.

4.2 Resistance mechanisms and biomarkers in melanoma

In this thesis, several known resistance mechanisms including MAPK and PI3K/AKT signaling were identified in melanoma cell lines, patient derived xenografts and patient tumor materials in response to kinase inhibitors and immunotherapy. Several key proteins and phosphorylation sites within these pathways were identified with high confidence, such as EGFR, BRAF, NRAS and ERK1/2 proteins, as well as phosphorylation on AKT, ERK1/2 and MEK. In addition, in the proteogenomics dataset, a number of proteins and phosphoproteins were found to be mutated with the potential to rewire signal transduction. For example, NRAS mutations were identified in all patient-derived xenografts, while PTEN and CDK4 were found in the drug-resistant cell lines A375 and SkMel28, respectively. Interestingly, most of the identified mutations were not previously reported in melanoma and therefore may be of high interest for further investigations. Taken together, this highlights the good coverage of our datasets and their utility as a resource for the melanoma community. In addition, an accumulation of mutations in specific pathways, such as focal adhesion, PI3K/AKT signaling and GTPase activation in response to BRAFi resistance and immunotherapy were detected in this thesis suggesting their relevance in both treatment options.

The tumor microenvironment and remodeling of the cytoskeletal organization have also been reported to play an important role in the development of acquired resistance. In the quantitative proteomics analysis, the intermediate filament nestin was identified as one of the most down-regulated proteins in drug-resistant cells. Several studies have been shown that nestin is involved in cancer cell migration, invasion, and metastasis [148,153,458]. In addition, the down-

regulation of nestin in FFPE specimens using immunohistochemistry was confirmed; however, expression of nestin differed in the immunohistochemistry staining between tumor specimens. Quantitative proteomics of one pair of pre-and post-treated tumors identified nestin as one of the most down-regulated proteins. These results highlight the significance of nestin expression in human tumors. Doxie and colleagues also showed that nestin expression was completely depleted in human tumors after BRAF and MEK inhibitor therapy. To study the effect of nestin, a CRISPR/Cas9 approach was used to generate a *NES* gene knockout in drug-sensitive melanoma cells. This study provides novel data showing that nestin expression significantly correlates with cell survival and colony formation upon MAPK signaling pathway inhibitor treatment. Indeed, cell survival and colony formation ability in knockout cells was increased following BRAF inhibitor treatment, but also upon inhibition with the MEK inhibitor cobimetinib. Several studies have suggested that combined therapy with BRAF and MEK inhibitors are promising to delay MAPK-driven acquired resistance, but may lead to other resistance mechanisms, for example PI3K/AKT/mTOR signaling pathways [459]. Depletion of nestin may activate these resistance mechanisms and increase cell survival upon mono- or combined therapy. The proteomic and phosphoproteomic analysis provide evidence that nestin depletion is associated with signaling through focal adhesion, integrin and PI3K/AKT/mTOR pathways. Interestingly at the proteome level, ECM interacting proteins, such as Laminin-B or Filamin-B, the integrins β1 and β4, Proteinkinase C, FAK and other downstream signaling proteins were significantly up-regulated in the genome edited cells compared to drug sensitive cells. Furthermore, the phosphoproteome analysis revealed differentially regulated phosphorylation sites on the key players of the integrin signaling pathway and downstream proteins. In conclusion, nestin protein levels could be linked not only with an invasive phenotype, but also with acquired drug resistance in melanoma.

The proteogenomic analysis led to the identification of a number of network-attacking mutations, most of which were unique to specific cell line or PDX. For example, a non-synonymous single amino acid variant with the potential to change the protein modification status was found on the runt-related transcription factor 1 protein (RUNX1) in A375 cell line. The mutation at position S276L resulted in the knock-out of a known phosphosite on this critical protein that was previously identified to be involved in cancer development. RUNX1 plays an important role in cell proliferation, differentiation and apoptosis [156] and several of its interaction partners were also mutated in drug-sensitive and drug-resistant cells, including PML, CTBP2 and YAP1. Interactome studies of wild-type and mutated RUNX1 revealed that the interactome of RUNX1 was altered due the loss of phosphorylation sites. Indeed, several interaction

partners, such as the transcriptional repressors HDAC1 and Sin3A, were significantly enriched in RUNX1_wt interactome and depleted in RUNX1_S276L interactome. These results also suggested a shift in transcriptional activity of mutated RUNX1 compared to wild-type protein.

In addition to nestin, another protein, involved in signaling pathways that alter the microenvironment of cancer cells, was identified to be mutated in both melanoma cell lines and patient-derived xenografts. The dedicator of cytokinesis protein 1 (*DOCK1*) was mutated at A1857T and led to gain of a new modifiable residue (phosphorylation) that could be confirmed by high resolution mass spectrometry. DOCK1 is mutated in 15% of melanoma patients and is involved in integrin signaling through FAK activation, MAPK pathway and cytoskeletal rearrangement. DOCK1 also regulates the activity of RAC and MAPK2, which are key molecules in resistance mechanisms and among the main drivers of melanogenesis. The results of the interactome studies in drug-sensitive and drug-resistant melanoma cells suggest that this mutation has an impact on the interactome of DOCK1. Notably, the intermediate filament nestin was also identified to enriched in the pulldown of DOCK1 in drug-resistant melanoma cells. Both proteins are involved in cytoskeletal reorganization, which is known to play an important role in cancer progression and development of resistance.

This thesis highlights that proteogenomics provides valuable insights into cancer biology and how the proteome is regulated by genetic effects.

5 Conclusions

Malignant melanoma is characterized by mutation in key signaling pathways including MAPK and PI3K/AKT pathways. Over the last decade several treatments including kinase and immune checkpoint inhibitors have been developed. However, acquired resistance hampered the prolongation of progression-free survival and response rates. In this work, mechanisms underlying resistance were investigated in melanoma cell lines, patient-derived xenografts and tumor tissue, via exhaustive characterization of mutational landscape, proteome and phosphoproteome profiling. Results obtained in the three parts of the thesis lead to the following conclusions:

1. Comprehensive (phospho)proteomics analysis of vemurafenib-sensitive and resistant melanoma cells
   a) Distinct proteome and phosphoproteome changes were observed between BRAF inhibitor resistant- and sensitive cells
   b) Intermediate filament protein nestin was identified as one of the highest down-regulated proteins in melanoma cells and tumors.
   c) CRISPR/Cas9-mediated knockout of nestin gene led to increased cellular proliferation and colony formation upon BRAF and MEK inhibition
   d) Depletion of nestin led to increased invasiveness and metalloproteinases activity similar to vermurafenib-resistant cells
   e) Phosphoproteome analysis revealed that nestin depletion is associated with integrin and PI3K/AKT/mTOR signaling and also with acquired drug resistance in melanoma

2. Individualized proteogenomic characterization of melanoma cell lines in response to vemurafenib resistance
   a) Integration of genomics and proteomics led to the identification of numerous non-synonymous variants at the genomic and (phospho)proteomic level
   b) Disproportional impact of nucleotide variants on modifiable residues were detected between sensitive and resistant cell lines
   c) Functional investigation of over-represented pathways of alternate proteins between phenotypes and cell lines showed differing mechanism linked to BRAFi resistance
   d) Several alternate peptides, which are phosphorylated on the mutation site were identified by mass spectrometry

e) Rewiring of signal transduction due to the loss of a known phosphorylation site on RUNX1 protein was confirmed by interactome studies

3. Individualized proteogenomic characterization of human melanoma tumor cells in response to immunotherapy
   a) Distinct mutational landscapes were identified between naïve and immune checkpoint inhibitor treated tumor samples
   b) Proteogenomics highlighted accumulation of mutations in specific pathways that are linked to immunotherapy
   c) Proteogenomic profiles of PDXs resembled the main findings in human tumor tissue
   d) Actionable mutations with high potential to rewire signal transduction were detected by integrating genomic, proteomic and drug database data
   e) Interactome studies confirmed change in interactome due to gain of new phosphorylation site on the GEF protein DOCK1

6 Future perspectives

Late stage malignant melanoma is often treated with selective kinase inhibitors targeting mutated BRAF or MEK kinases and/or immunotherapy using immune checkpoint inhibitors. However, after a period of a progression-free disease, most patients develop resistance to the therapy, which is followed by rapid progression of cancer. Many underlying factors behind development of resistance to kinase inhibition have been studied by genomics and transcriptomics and more recently by proteomics. Yet, to better understand cancer biology underlying resistance, new tools are needed to investigate mutational profiles at the genome, proteome and PTM level. Likewise, individualized approaches and personalized therapeutic predictors are of utmost importance. Moreover, the lack of insights into the disease mechanisms, as well as the complexity of cancer and patient heterogeneity calls for proteomic and proteogenomic approaches applied in context of personalized medicine and should be the focus in future investigations. In addition, PTMs are still largely unexplored in context in cancer resistance and progression and should be studied further due to their transient importance in cancer. This thesis provides datasets that can be used as a resource to identify key (phospho)proteins and network-attacking mutations in context of kinase inhibition and immunotherapy in order to improve patient's survival, as well the frequency and impacts of mutations on all levels and provides insights into molecular etiology of melanoma. Further validation and functional assessment of selected regulated proteins and mutations will provide insights into the importance of specific mutations for establishment of drug resistant and signaling pathways. Predicted actionable drug targets should be further validated by drug target assays in drug-sensitive and drug-resistant cells as well as in cell lines generated from the same patient tumor tissue. Additionally, further experiments need to be done to compare reference and mutated version of DOCK1 in context of BRAF and checkpoint inhibition to highlight the clinical relevance of this mutation site in melanoma. Moreover, proteogenomic approaches are a quite new research area and only a few reports in melanoma translational research have been utilized. The results of this thesis and other showed that proteogenomics has the potential to reveal variants that alter protein abundance and modification status linked to the disease and are of critical importance to determine the feasibility of proteogenomics approaches in individualized cancer medicine. However, efficient and easy-to use tool for the bioinformatic analysis and visualization of proteogenomic data must be optimized to become a true prognostic tool. Only then, proteogenomics will be able to improve the treatment of patients in personalized manner and may be used routinely in clinical context.

7 References

1 Robert Koch, I. *Cancer in Germany Cancer in Germany*. (2014).

2 Domingues, B., Lopes, J. M., Soares, P. & Pópulo, H. Melanoma treatment in review. *ImmunoTargets and therapy* **7**, 35 (2018).

3 Gilchrest, B. A., Eller, M. S., Geller, A. C. & Yaar, M. The pathogenesis of melanoma induced by ultraviolet radiation. *N Engl J Med* **340**, 1341-1348, doi:10.1056/NEJM199904293401707 (1999).

4 Pennello, G., Devesa, S. & Gail, M. Association of surface ultraviolet B radiation levels with melanoma and nonmelanoma skin cancer in United States blacks. *Cancer Epidemiol Biomarkers Prev* **9**, 291-297 (2000).

5 Elwood, J. M. & Jopson, J. Melanoma and sun exposure: an overview of published studies. *Int J Cancer* **73**, 198-203, doi:10.1002/(sici)1097-0215(19971009)73:2<198::aid-ijc6>3.0.co;2-r (1997).

6 Lazovich, D. *et al.* Indoor tanning and risk of melanoma: a case-control study in a highly exposed population. *Cancer Epidemiol Biomarkers Prev* **19**, 1557-1568, doi:10.1158/1055-9965.EPI-09-1249 (2010).

7 Archier, E. *et al.* Carcinogenic risks of psoralen UV-A therapy and narrowband UV-B therapy in chronic plaque psoriasis: a systematic literature review. *J Eur Acad Dermatol Venereol* **26 Suppl 3**, 22-31, doi:10.1111/j.1468-3083.2012.04520.x (2012).

8 Bauer, J. & Garbe, C. Acquired melanocytic nevi as risk factor for melanoma development. A comprehensive review of epidemiological data. *Pigment Cell Res* **16**, 297-306, doi:10.1034/j.1600-0749.2003.00047.x (2003).

9 Bevona, C., Goggins, W., Quinn, T., Fullerton, J. & Tsao, H. Cutaneous melanomas associated with nevi. *Arch Dermatol* **139**, 1620-1624; discussion 1624, doi:10.1001/archderm.139.12.1620 (2003).

10 Goldstein, A. M. & Tucker, M. A. Genetic epidemiology of cutaneous melanoma: a global perspective. *Arch Dermatol* **137**, 1493-1496, doi:10.1001/archderm.137.11.1493 (2001).

11 Chang, A. E., Karnell, L. H. & Menck, H. R. The National Cancer Data Base report on cutaneous and noncutaneous melanoma: a summary of 84,836 cases from the past decade. The American College of Surgeons Commission on Cancer and the American Cancer Society. *Cancer* **83**, 1664-1678, doi:10.1002/(sici)1097-0142(19981015)83:8<1664::aid-cncr23>3.0.co;2-g (1998).

12 Gray-Schopfer, V., Wellbrock, C. & Marais, R. Melanoma biology and new targeted therapy. *Nature* **445**, 851-857, doi:10.1038/nature05661 (2007).

13 Tolleson, W. H. Human melanocyte biology, toxicology, and pathology. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev* **23**, 105-161, doi:10.1080/10590500500234970 (2005).

14 Markova, E., Petrova, N., Razin, S. & Kantidze, O. Transcription factor RUNX1. *Molecular Biology* **46**, 755-767 (2012).

15 Dong, L. *et al.* Melanocyte-stimulating hormone directly enhances UV-Induced DNA repair in keratinocytes by a xeroderma pigmentosum group A-dependent mechanism. *Cancer Res* **70**, 3547-3556, doi:10.1158/0008-5472.CAN-09-4596 (2010).

16 Brenner, M. & Hearing, V. J. The protective role of melanin against UV damage in human skin. *Photochem Photobiol* **84**, 539-549, doi:10.1111/j.1751-1097.2007.00226.x (2008).

17 Carsberg, C. J. W., H.M.; Friedmann, P.S. Ultraviolet Radiation-Induced Melanogenesis in Human Melanocytes Effects of Modulating Protein Kinase C. *J Cell Sci* **107**, 2591-2597 (1994).

18 Bonaventure, J., Domingues, M. J. & Larue, L. Cellular and molecular mechanisms controlling the migration of melanocytes and melanoma cells. *Pigment Cell Melanoma Res* **26**, 316-325, doi:10.1111/pcmr.12080 (2013).
19 Flaherty, K. T., Hodi, F. S. & Fisher, D. E. From genes to drugs: targeted strategies for melanoma. *Nat Rev Cancer* **12**, 349-361, doi:10.1038/nrc3218 (2012).
20 Eves, P. C. & Haycock, J. W. Melanocortin signalling mechanisms. *Adv Exp Med Biol* **681**, 19-28, doi:10.1007/978-1-4419-6354-3_2 (2010).
21 Balch, C. M. *et al.* An evidence-based staging system for cutaneous melanoma. *CA Cancer J Clin* **54**, 131-149; quiz 182-134, doi:10.3322/canjclin.54.3.131 (2004).
22 Balch, C. M. *et al.* Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* **27**, 6199-6206, doi:10.1200/JCO.2009.23.4799 (2009).
23 American Cancer, S. Cancer Facts & Figures 2017. (2017).
24 E., S. J. a. W. http://www.pathophys.org/melanoma/.
25 Cargnello, M. & Roux, P. P. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol Mol Biol Rev* **75**, 50-83, doi:10.1128/MMBR.00031-10 (2011).
26 Hubbard, S. R. & Miller, W. T. Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr Opin Cell Biol* **19**, 117-123, doi:10.1016/j.ceb.2007.02.010 (2007).
27 Wagner, M. J., Stacey, M. M., Liu, B. A. & Pawson, T. Molecular mechanisms of SH2- and PTB-domain-containing proteins in receptor tyrosine kinase signaling. *Cold Spring Harb Perspect Biol* **5**, a008987, doi:10.1101/cshperspect.a008987 (2013).
28 Matallanas, D. *et al.* Raf family kinases: old dogs have learned new tricks. *Genes Cancer* **2**, 232-260, doi:10.1177/1947601911407323 (2011).
29 Hodis, E. *et al.* A landscape of driver mutations in melanoma. *Cell* **150**, 251-263, doi:10.1016/j.cell.2012.06.024 (2012).
30 Amaral, T. *et al.* MAPK pathway in melanoma part II-secondary and adaptive resistance mechanisms to BRAF inhibition. *Eur J Cancer* **73**, 93-101, doi:10.1016/j.ejca.2016.12.012 (2017).
31 Davies, H. *et al.* Mutations of the BRAF gene in human cancer. *Nature* **417**, 949-954, doi:10.1038/nature00766 (2002).
32 Flaherty, K. T. *et al.* Improved Survival with MEK Inhibition in BRAF-Mutated Melanoma. *New England Journal of Medicine* **367**, 107-114, doi:10.1056/NEJMoa1203421 (2012).
33 Allen, E. M. V. *et al.* The genetic landscape of clinical resistance to RAF inhibition in metastatic melnaoma. *Cancer Discov.* **4**, 94-109, doi:10.1158/2159-8290.CD-13-0617.The (2014).
34 Long, G. V. *et al.* Prognostic and clinicopathologic associations of oncogenic BRAF in metastatic melanoma. *J Clin Oncol* **29**, 1239-1246, doi:10.1200/JCO.2010.32.4327 (2011).
35 Brose, M. S. *et al.* BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Res* **62**, 6997-7000 (2002).
36 Chamcheu, J. C. *et al.* Role and Therapeutic Targeting of the PI3K/Akt/mTOR Signaling Pathway in Skin Cancer: A Review of Current Status and Future Trends on Natural and Synthetic Agents Therapy. *Cells* **8**, doi:10.3390/cells8080803 (2019).
37 Hemmings, B. A. & Restuccia, D. F. PI3K-PKB/Akt pathway. *Cold Spring Harb Perspect Biol* **4**, a011189, doi:10.1101/cshperspect.a011189 (2012).
38 Luo, J., Manning, B. D. & Cantley, L. C. Targeting the PI3K-Akt pathway in human cancer: rationale and promise. *Cancer Cell* **4**, 257-262, doi:10.1016/s1535-6108(03)00248-4 (2003).

39 Maehama, T. & Dixon, J. E. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem* **273**, 13375-13378, doi:10.1074/jbc.273.22.13375 (1998).

40 Hocker, T. & Tsao, H. Ultraviolet radiation and melanoma: a systematic review and analysis of reported sequence variants. *Hum Mutat* **28**, 578-588, doi:10.1002/humu.20481 (2007).

41 Hensin Tsao1, X. Z., Eric Benoit and Frank G Haluska. Identification of PTEN/MMAC1 alterations in uncultured melanomas and melanoma cell lines. *Oncogene* **16**, 3397-3402 (1998).

42 Goel, V. K., Lazar, A. J., Warneke, C. L., Redston, M. S. & Haluska, F. G. Examination of mutations in BRAF, NRAS, and PTEN in primary cutaneous melanoma. *J Invest Dermatol* **126**, 154-160, doi:10.1038/sj.jid.5700026 (2006).

43 Carpten, J. D. *et al.* A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. *Nature* **448**, 439-444, doi:10.1038/nature05933 (2007).

44 Davies, M. A. *et al.* A novel AKT3 mutation in melanoma tumours and cell lines. *Br J Cancer* **99**, 1265-1268, doi:10.1038/sj.bjc.6604637 (2008).

45 Stahl, J. M. *et al.* Deregulated Akt3 activity promotes development of malignant melanoma. *Cancer Res* **64**, 7002-7010, doi:10.1158/0008-5472.CAN-04-1399 (2004).

46 Curtin, J. A., Stark, M. S., Pinkel, D., Hayward, N. K. & Bastian, B. C. PI3-kinase subunits are infrequent somatic targets in melanoma. *J Invest Dermatol* **126**, 1660-1663, doi:10.1038/sj.jid.5700311 (2006).

47 Omholt, K., Krockel, D., Ringborg, U. & Hansson, J. Mutations of PIK3CA are rare in cutaneous melanoma. *Melanoma Res* **16**, 197-200, doi:10.1097/01.cmr.0000200488.77970.e3 (2006).

48 Gibney, G. T. & Smalley, K. S. An unholy alliance: cooperation between BRAF and NF1 in melanoma development and BRAF inhibitor resistance. *Cancer Discov* **3**, 260-263, doi:10.1158/2159-8290.CD-13-0017 (2013).

49 Maertens, O. *et al.* Elucidating distinct roles for NF1 in melanomagenesis. *Cancer Discov* **3**, 338-349, doi:10.1158/2159-8290.CD-12-0313 (2013).

50 Whittaker, S. R. *et al.* A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. *Cancer Discov* **3**, 350-362, doi:10.1158/2159-8290.CD-12-0470 (2013).

51 Nissan, M. H. *et al.* Loss of NF1 in cutaneous melanoma is associated with RAS activation and MEK dependence. *Cancer Res* **74**, 2340-2350, doi:10.1158/0008-5472.CAN-13-2625 (2014).

52 Krauthammer, M. *et al.* Exome sequencing identifies recurrent mutations in NF1 and RASopathy genes in sun-exposed melanomas. *Nat Genet* **47**, 996-1002, doi:10.1038/ng.3361 (2015).

53 Forbes, S. A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* **39**, D945-950, doi:10.1093/nar/gkq929 (2011).

54 Lyons, J. *et al.* Two G protein oncogenes in human endocrine tumors. *Science* **249**, 655-659, doi:10.1126/science.2116665 (1990).

55 Huang, J. L., Urtatiz, O. & Van Raamsdonk, C. D. Oncogenic G Protein GNAQ Induces Uveal Melanoma and Intravasation in Mice. *Cancer Res* **75**, 3384-3397, doi:10.1158/0008-5472.CAN-14-3229 (2015).

56 Van Raamsdonk, C. D. *et al.* Frequent somatic mutations of GNAQ in uveal melanoma and blue naevi. *Nature* **457**, 599-602, doi:10.1038/nature07586 (2009).

57 Gastonguay, A. *et al.* The role of Rac1 in the regulation of NF-kappaB activity, cell proliferation, and cell migration in non-small cell lung carcinoma. *Cancer Biol Ther* **13**, 647-656, doi:10.4161/cbt.20082 (2012).

58 Austin, E., Mamalis, A., Ho, D. & Jagdeo, J. Laser and light-based therapy for cutaneous and soft-tissue metastases of malignant melanoma: a systematic review. *Arch Dermatol Res* **309**, 229-242, doi:10.1007/s00403-017-1720-9 (2017).
59 Soengas, M. S. & Lowe, S. W. Apoptosis and melanoma chemoresistance. *Oncogene* **22**, 3138-3151, doi:10.1038/sj.onc.1206454 (2003).
60 Li, J. *et al.* Recent advances in targeted nanoparticles drug delivery to melanoma. *Nanomedicine* **11**, 769-794, doi:10.1016/j.nano.2014.11.006 (2015).
61 Kim, C. *et al.* Long-term survival in patients with metastatic melanoma treated with DTIC or temozolomide. *Oncologist* **15**, 765-771, doi:10.1634/theoncologist.2009-0237 (2010).
62 Boussios, S., Pentheroudakis, G., Katsanos, K. & Pavlidis, N. Systemic treatment-induced gastrointestinal toxicity: incidence, clinical presentation and management. *Ann Gastroenterol* **25**, 106-118 (2012).
63 Jiang, G., Li, R. H., Sun, C., Liu, Y. Q. & Zheng, J. N. Dacarbazine combined targeted therapy versus dacarbazine alone in patients with malignant melanoma: a meta-analysis. *PLoS One* **9**, e111920, doi:10.1371/journal.pone.0111920 (2014).
64 Thaxton, J. E. & Li, Z. To affinity and beyond: harnessing the T cell receptor for cancer immunotherapy. *Hum Vaccin Immunother* **10**, 3313-3321, doi:10.4161/21645515.2014.973314 (2014).
65 Kim, R., Emi, M. & Tanabe, K. Cancer immunoediting from immune surveillance to immune escape. *Immunology* **121**, 1-14, doi:10.1111/j.1365-2567.2007.02587.x (2007).
66 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
67 Schreiber, R. D., Old, L. J. & Smyth, M. J. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* **331**, 1565-1570, doi:10.1126/science.1203486 (2011).
68 Quandt, D., Hoff, H., Rudolph, M., Fillatreau, S. & Brunner-Weinzierl, M. C. A new role of CTLA-4 on B cells in thymus-dependent immune responses in vivo. *J Immunol* **179**, 7316-7324, doi:10.4049/jimmunol.179.11.7316 (2007).
69 Allison, J. P. & Krummel, M. F. The Yin and Yang of T cell costimulation. *Science* **270**, 932-933, doi:10.1126/science.270.5238.932 (1995).
70 Perkins, D. *et al.* Regulation of CTLA-4 expression during T cell activation. *J Immunol* **156**, 4154-4159 (1996).
71 Greene, J. L. *et al.* Covalent dimerization of CD28/CTLA-4 and oligomerization of CD80/CD86 regulate T cell costimulatory interactions. *J Biol Chem* **271**, 26762-26771, doi:10.1074/jbc.271.43.26762 (1996).
72 Simpson, T. R. *et al.* Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. *J Exp Med* **210**, 1695-1710, doi:10.1084/jem.20130579 (2013).
73 Hodi, F. S. *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* **363**, 711-723, doi:10.1056/NEJMoa1003466 (2010).
74 Good-Jacobson, K. L. *et al.* PD-1 regulates germinal center B cell survival and the formation and affinity of long-lived plasma cells. *Nat Immunol* **11**, 535-542, doi:10.1038/ni.1877 (2010).
75 Keir, M. E., Butte, M. J., Freeman, G. J. & Sharpe, A. H. PD-1 and its ligands in tolerance and immunity. *Annu Rev Immunol* **26**, 677-704, doi:10.1146/annurev.immunol.26.021607.090331 (2008).
76 Raedler, L. A. Opdivo (Nivolumab): Second PD-1 Inhibitor Receives FDA Approval for Unresectable or Metastatic Melanoma. *Am Health Drug Benefits* **8**, 180-183 (2015).
77 Wolchok, J. D. *et al.* Nivolumab plus ipilimumab in advanced melanoma. *N Engl J Med* **369**, 122-133, doi:10.1056/NEJMoa1302369 (2013).

78 Wolchok, J. D. *et al.* Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *N Engl J Med* **377**, 1345-1356, doi:10.1056/NEJMoa1709684 (2017).
79 Samlowski, W. E. *et al.* High frequency of brain metastases after adjuvant therapy for high-risk melanoma. *Cancer Med* **6**, 2576-2585, doi:10.1002/cam4.1223 (2017).
80 Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* **348**, 69-74, doi:10.1126/science.aaa4971 (2015).
81 Riaz, N. *et al.* The role of neoantigens in response to immune checkpoint blockade. *Int Immunol* **28**, 411-419, doi:10.1093/intimm/dxw019 (2016).
82 Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568-571, doi:10.1038/nature13954 (2014).
83 Dunn, G. P., Old, L. J. & Schreiber, R. D. The three Es of cancer immunoediting. *Annu Rev Immunol* **22**, 329-360, doi:10.1146/annurev.immunol.22.012703.104803 (2004).
84 Van Allen, E. M. *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207-211, doi:10.1126/science.aad0095 (2015).
85 Anagnostou, V. *et al.* Evolution of Neoantigen Landscape during Immune Checkpoint Blockade in Non-Small Cell Lung Cancer. *Cancer Discov* **7**, 264-276, doi:10.1158/2159-8290.CD-16-0828 (2017).
86 Zaretsky, J. M. *et al.* Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N Engl J Med* **375**, 819-829, doi:10.1056/NEJMoa1604958 (2016).
87 Zhao, F. *et al.* Melanoma Lesions Independently Acquire T-cell Resistance during Metastatic Latency. *Cancer Res* **76**, 4347-4358, doi:10.1158/0008-5472.CAN-16-0008 (2016).
88 Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48-61, doi:10.1016/j.cell.2014.12.033 (2015).
89 Roh, W. *et al.* Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci Transl Med* **9**, doi:10.1126/scitranslmed.aah3560 (2017).
90 Homet Moreno, B. *et al.* Response to Programmed Cell Death-1 Blockade in a Murine Melanoma Syngeneic Model Requires Costimulation, CD4, and CD8 T Cells. *Cancer Immunol Res* **4**, 845-857, doi:10.1158/2326-6066.CIR-16-0060 (2016).
91 Shin, D. S. *et al.* Primary Resistance to PD-1 Blockade Mediated by JAK1/2 Mutations. *Cancer Discov* **7**, 188-201, doi:10.1158/2159-8290.CD-16-1223 (2017).
92 Gao, J. *et al.* Loss of IFN-gamma Pathway Genes in Tumor Cells as a Mechanism of Resistance to Anti-CTLA-4 Therapy. *Cell* **167**, 397-404 e399, doi:10.1016/j.cell.2016.08.069 (2016).
93 Liu, C. *et al.* BRAF inhibition increases tumor infiltration by T cells and enhances the antitumor activity of adoptive immunotherapy in mice. *Clin Cancer Res* **19**, 393-403, doi:10.1158/1078-0432.CCR-12-1626 (2013).
94 Liu, L. *et al.* The BRAF and MEK Inhibitors Dabrafenib and Trametinib: Effects on Immune Function and in Combination with Immunomodulatory Antibodies Targeting PD-1, PD-L1, and CTLA-4. *Clin Cancer Res* **21**, 1639-1651, doi:10.1158/1078-0432.CCR-14-2339 (2015).
95 George, S. *et al.* Loss of PTEN Is Associated with Resistance to Anti-PD-1 Checkpoint Blockade Therapy in Metastatic Uterine Leiomyosarcoma. *Immunity* **46**, 197-204, doi:10.1016/j.immuni.2017.02.001 (2017).
96 Peng, W. *et al.* Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy. *Cancer Discov* **6**, 202-216, doi:10.1158/2159-8290.CD-15-0283 (2016).

97 O'Donnell, J. S., Massi, D., Teng, M. W. L. & Mandala, M. PI3K-AKT-mTOR inhibition in cancer immunotherapy, redux. *Semin Cancer Biol* **48**, 91-103, doi:10.1016/j.semcancer.2017.04.015 (2018).
98 Kaneda, M. M. *et al.* Corrigendum: PI3Kgamma is a molecular switch that controls immune suppression. *Nature* **542**, 124, doi:10.1038/nature21026 (2017).
99 Bollag, G. *et al.* Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **467**, 596-599, doi:10.1038/nature09454 (2010).
100 Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Eng J Med* **364**, 2507-2516, doi:10.1056/NEJMoa1103782.Improved (2011).
101 Sosman, J. A. *et al.* Survival in BRAF V600–Mutant Advanced Melanoma Treated with Vemurafenib. *N Eng J Med* **366**, 707-714, doi:10.1056/NEJMoa1112302.Survival (2012).
102 Shi, H., Hugo, W. & Kong, X. Acquired resistance and clonal evolution in melanoma during BRAF inhibior therapy. *Cancer Discov.* **144**, 724-732, doi:10.1038/jid.2014.371 (2014).
103 Wagle, N. *et al.* Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. *Journal of Clinical Oncology* **29**, 3085-3096, doi:10.1200/JCO.2010.33.2312 (2011).
104 Manzano, J. L. *et al.* Resistant mechanisms to BRAF inhibitors in melanoma. *Annals of translational medicine* **4** (2016).
105 Neuzillet, C. *et al.* MEK in cancer and cancer therapy. *Pharmacol Ther* **141**, 160-171, doi:10.1016/j.pharmthera.2013.10.001 (2014).
106 Lugowska, I., Kosela-Paterczyk, H., Kozak, K. & Rutkowski, P. Trametinib: a MEK inhibitor for management of metastatic melanoma. *Onco Targets Ther* **8**, 2251-2259, doi:10.2147/OTT.S72951 (2015).
107 Banks, M., Crowell, K., Proctor, A. & Jensen, B. C. Cardiovascular Effects of the MEK Inhibitor, Trametinib: A Case Report, Literature Review, and Consideration of Mechanism. *Cardiovasc Toxicol* **17**, 487-493, doi:10.1007/s12012-017-9425-z (2017).
108 Ribas, A. *et al.* Combination of vemurafenib and cobimetinib in patients with advanced BRAF(V600)-mutated melanoma: a phase 1b study. *Lancet Oncol* **15**, 954-965, doi:10.1016/S1470-2045(14)70301-8 (2014).
109 Larkin, J. *et al.* Combined Vemurafenib and Cobimetinib in BRAF-Mutated Melanoma. *The New England journal of medicine* **371**, 1867-1876, doi:10.1056/NEJMoa1408868 (2014).
110 Ascierto, P. A. *et al.* Cobimetinib combined with vemurafenib in advanced BRAF(V600)-mutant melanoma (coBRIM): updated efficacy results from a randomised, double-blind, phase 3 trial. *Lancet Oncol* **17**, 1248-1260, doi:10.1016/S1470-2045(16)30122-X (2016).
111 Sun, C. *et al.* Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. *Nature* **508**, 118-122, doi:10.1038/nature13121 (2014).
112 Paraiso, K. H. *et al.* PTEN loss confers BRAF inhibitor resistance to melanoma cells through the suppression of BIM expression. *Cancer Res* **71**, 2750-2760, doi:10.1158/0008-5472.CAN-10-2954 (2011).
113 Xing, F. *et al.* Concurrent loss of the PTEN and RB1 tumor suppressors attenuates RAF dependence in melanomas harboring (V600E)BRAF. *Oncogene* **31**, 446-457, doi:10.1038/onc.2011.250 (2012).
114 Straussman, R. *et al.* Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* **487**, 500-504, doi:10.1038/nature11183 (2012).

115 Nazarian, R. *et al.* Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. *Nature* **468**, 973-977, doi:10.1038/nature09626 (2010).
116 Heidorn, S. J. *et al.* Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. *Cell* **140**, 209-221, doi:10.1016/j.cell.2009.12.040 (2010).
117 Su, F. *et al.* Resistance to selective BRAF inhibition can be mediated by modest upstream pathway activation. *Cancer Res* **72**, 969-978, doi:10.1158/0008-5472.CAN-11-1875 (2012).
118 Johnson, D. B. *et al.* Acquired BRAF inhibitor resistance: A multicenter meta-analysis of the spectrum and frequencies, clinical behaviour, and phenotypic associations of resistance mechanisms. *Eur J Cancer* **51**, 2792-2799, doi:10.1016/j.ejca.2015.08.022 (2015).
119 Poulikakos, P. I. *et al.* RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). *Nature* **480**, 387-390, doi:10.1038/nature10662 (2011).
120 Rizos, H. *et al.* BRAF inhibitor resistance mechanisms in metastatic melanoma: spectrum and clinical impact. *Clin Cancer Res* **20**, 1965-1977, doi:10.1158/1078-0432.CCR-13-3122 (2014).
121 Johannessen, C. M. *et al.* COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. *Nature* **468**, 968-972, doi:10.1038/nature09627 (2010).
122 Villanueva, J. *et al.* Acquired resistance to BRAF inhibitors mediated by a RAF kinase switch in melanoma can be overcome by cotargeting MEK and IGF-1R/PI3K. *Cancer Cell* **18**, 683-695, doi:10.1016/j.ccr.2010.11.023 (2010).
123 Trunzer, K. *et al.* Pharmacodynamic effects and mechanisms of resistance to vemurafenib in patients with metastatic melanoma. *Journal of Clinical Oncology*, doi:10.1200/JCO.2012.44.7888 (2013).
124 Carlino, M. S. *et al.* Preexisting MEK1P124 mutations diminish response to BRAF inhibitors in metastatic melanoma patients. *Clin Cancer Res* **21**, 98-105, doi:10.1158/1078-0432.CCR-14-0759 (2015).
125 Muller, J. *et al.* Low MITF/AXL ratio predicts early resistance to multiple targeted drugs in melanoma. *Nat Commun* **5**, 5712, doi:10.1038/ncomms6712 (2014).
126 Martz, C. A. *et al.* Systematic identification of signaling pathways with potential to confer anticancer drug resistance. *Sci Signal* **7**, ra121, doi:10.1126/scisignal.aaa1877 (2014).
127 Karoulia, Z., Gavathiotis, E. & Poulikakos, P. I. New perspectives for targeting RAF kinase in human cancer. *Nat Rev Cancer* **17**, 676-691, doi:10.1038/nrc.2017.79 (2017).
128 Eroglu, Z. & Ribas, A. Combination therapy with BRAF and MEK inhibitors for melanoma: latest evidence and place in therapy. *Ther Adv Med Oncol* **8**, 48-56, doi:10.1177/1758834015616934 (2016).
129 Welsh, S. J. & Corrie, P. G. Management of BRAF and MEK inhibitor toxicities in patients with metastatic melanoma. *Ther Adv Med Oncol* **7**, 122-136, doi:10.1177/1758834014566428 (2015).
130 Goetz, E. M., Ghandi, M., Treacy, D. J., Wagle, N. & Garraway, L. A. ERK mutations confer resistance to mitogen-activated protein kinase pathway inhibitors. *Cancer Res* **74**, 7079-7089, doi:10.1158/0008-5472.CAN-14-2073 (2014).
131 Moriceau, G. *et al.* Tunable-combinatorial mechanisms of acquired resistance limit the efficacy of BRAF/MEK cotargeting but result in melanoma drug addiction. *Cancer Cell* **27**, 240-256, doi:10.1016/j.ccell.2014.11.018 (2015).
132 Kwong, L. N. & Davies, M. A. Navigating the therapeutic complexity of PI3K pathway inhibition in melanoma. *Clin Cancer Res* **19**, 5310-5319, doi:10.1158/1078-0432.CCR-13-0142 (2013).

133 Penna, I. *et al.* Primary cross-resistance to BRAFV600E-, MEK1/2- and PI3K/mTOR-specific inhibitors in BRAF-mutant melanoma cells counteracted by dual pathway blockade. *Oncotarget* **7**, 3947-3965, doi:10.18632/oncotarget.6600 (2016).
134 Long, G. V. *et al.* Increased MAPK reactivation in early resistance to dabrafenib/trametinib combination therapy of BRAF-mutant metastatic melanoma. *Nat Commun* **5**, 5694, doi:10.1038/ncomms6694 (2014).
135 Rajasagi, M. *et al.* Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood* **124**, 453-462, doi:10.1182/blood-2014-04-567933 (2014).
136 Dong, L. *et al.* Clinical Next Generation Sequencing for Precision Medicine in Cancer. *Curr Genomics* **16**, 253-263, doi:10.2174/1389202915666150511205313 (2015).
137 Welinder, C. *et al.* Correlation of histopathologic characteristics to protein expression and function in malignant melanoma. *PLoS One* **12**, e0176167, doi:10.1371/journal.pone.0176167 (2017).
138 Bystrom, S. *et al.* Affinity Proteomics Exploration of Melanoma Identifies Proteins in Serum with Associations to T-Stage and Recurrence. *Transl Oncol* **10**, 385-395, doi:10.1016/j.tranon.2017.03.002 (2017).
139 Klein, W. M. *et al.* Increased expression of stem cell markers in malignant melanoma. *Mod Pathol* **20**, 102-107, doi:10.1038/modpathol.3800720 (2007).
140 Sellheyer, K. & Krahl, D. Spatiotemporal expression pattern of neuroepithelial stem cell marker nestin suggests a role in dermal homeostasis, neovasculogenesis, and tumor stroma development: A study on embryonic and adult human skin. *Journal of American Dermatology* **63**, 93-113, doi:10.1016/j.jaad.2009.07.013 (2009).
141 Qendro, V. *et al.* Large-scale proteomic characterization of melanoma expressed proteins reveals nestin and vimentin as biomarkers that can potentially distinguish melanoma subtypes. *Journal of proteome research* **13**, 5031-5040, doi:10.1021/pr5006789 (2014).
142 Neradil, J. & Veselska, R. Nestin as a marker of cancer stem cells. *Cancer Sci* **106**, doi:10.1111/cas.12691 (2015).
143 Bernal, A. & Arranz, L. Nestin-expressing progenitor cells: function, identity and therapeutic implications. *Cell Mol Life Sci* **75**, 2177-2195, doi:10.1007/s00018-018-2794-z (2018).
144 Michalczyk, K. & Ziman, M. Nestin structure and predicted function in cellular cytoskeletal organisation. *Histol Histopathol* **20**, 665-671, doi:10.14670/HH-20.665 (2005).
145 Ishiwata, T., Teduka, K. & Yamamoto, T. Neuroepithelial stem cell marker nestin regulates the migration , invasion and growth of human gliomas. *Oncology reports*, 91-99, doi:10.3892/or.2011.1267 (2011).
146 Brychtova, S., Fiuraskova, M., Hlobilkova, A., Brychta, T. & Hirnak, J. Nestin expression in cutaneous melanomas and melanocytic nevi. *J Cutan Pathol* **34**, 370-375, doi:10.1111/j.1600-0560.2006.00627.x (2007).
147 Matsuda, Y., Ishiwata, T., Yoshimura, H., Yamahatsu, K. & Minamoto, T. Nestin phosphorylatin at threonine 315 and 1299 correlates with proliferation and metastasis of human pancreatic cancer. *Cancer Sci* **108**, 354-361, doi:10.1111/cas.13139 (2016).
148 Hyder, C. L., Lazaro, G., Pylva, J. W., Qvarnstro, S. M. & Eriksson, J. E. Nestin regulates prostate cancer cell invasion by influencing the localisation and functions of FAK and integrins. *Journal of cell science* **127**, 2161-2173, doi:10.1242/jcs.125062 (2014).
149 Ladstein, R. G., Bachmann, I. M., Straume, O. & Akslen, L. A. Nestin expression is associated with aggressive cutaneous melanoma of the nodular type. *Mod Pathol* **27**, 396-401, doi:10.1038/modpathol.2013.151 (2014).

150 Ishiwata, T. *et al.* Nestin in gastrointestinal and other cancers : Effects on cells and tumor angiogenesis. *World J Gastoenterol* **17**, 409-418, doi:10.3748/wjg.v17.i4.409 (2011).
151 Piras, F. *et al.* The stem cell marker nestin predicts poor prognosis in human melanoma The stem cell marker nestin predicts poor prognosis in human melanoma. *Oncology reports* **23**, 17-24, doi:10.3892/or (2010).
152 Akiyama, M., Matsuda, Y., Ishiwata, T., Naito, Z. & Kawana, S. Inhibition of the Stem Cell Marker Nestin Reduces Tumor Growth and Invasion of Malignant Melanoma. *Journal of investigative Dermatology* **133**, 1384-1387, doi:10.1038/jid.2012.508 (2013).
153 Narita, K. *et al.* Nestin regulates proliferation, migration, invasion and stemness of lung adenocarcinoma. *Int J Oncol* **44**, 1118-1130, doi:10.3892/ijo.2014.2278 (2014).
154 Yamahatsu, K., Matsuda, Y., Ishiwata, T., Uchida, E. & Naito, Z. Nestin as a novel therapeutic target for pancreatic cancer via tumor angiogenesis. *Int J Oncol* **40**, 1345-1357, doi:10.3892/ijo.2012.1333 (2012).
155 Lee, C.-W. *et al.* Nestin depletion induces melanoma matrix metalloproteinases and invasion. *Lab Invest* **94**, 1382-1395, doi:10.1083/jcb.200504124 (2014).
156 Otto, F., Lübbert, M. & Stock, M. Upstream and downstream targets of RUNX proteins. *Journal of cellular biochemistry* **89**, 9-18 (2003).
157 Levanon, D. & Groner, Y. Structure and regulated expression of mammalian RUNX genes. *Oncogene* **23**, 4211 (2004).
158 Westendorf, J. J. & Hiebert, S. W. Mammalian runt-domain proteins and their roles in hematopoiesis, osteogenesis, and leukemia. *Journal of cellular biochemistry* **75**, 51-58 (1999).
159 Ito, Y. & Miyazono, K. RUNX transcription factors as key targets of TGF-β superfamily signaling. *Current opinion in genetics & development* **13**, 43-47 (2003).
160 Goyama, S. & Mulloy, J. C. Molecular pathogenesis of core binding factor leukemia: current knowledge and future prospects. *International journal of hematology* **94**, 126-133 (2011).
161 Bartfeld, D. *et al.* DNA recognition by the RUNX1 transcription factor is mediated by an allosteric transition in the RUNT domain and by DNA bending. *Structure* **10**, 1395-1407 (2002).
162 Hyde, R. K. *et al.* Cbfb/Runx1 repression-independent blockage of differentiation and accumulation of Csf2rb-expressing cells by Cbfb-MYH11. *Blood* **115**, 1433-1443, doi:10.1182/blood-2009-06-227413 (2010).
163 Harada, H. *et al.* High incidence of somatic mutations in the AML1/RUNX1 gene in myelodysplastic syndrome and low blast percentage myeloid leukemia with myelodysplasia. *Blood* **103**, 2316-2324 (2004).
164 Imai, Y. *et al.* Mutations of the AML1 gene in myelodysplastic syndrome and their functional implications in leukemogenesis. *Blood* **96**, 3154-3160 (2000).
165 Consortium, A. P. G. AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer Discov* **7**, 818-831, doi:10.1158/2159-8290.CD-17-0151 (2017).
166 Miyoshi, H. *et al.* t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proc Natl Acad Sci U S A* **88**, 10431-10434, doi:10.1073/pnas.88.23.10431 (1991).
167 Sun, X. J. *et al.* A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature* **500**, 93-97, doi:10.1038/nature12287 (2013).
168 Chimge, N. & Frenkel, B. The RUNX family in breast cancer: relationships with estrogen signaling. *Oncogene* **32**, 2121 (2013).

169 Scheitz, C. J. F., Lee, T. S., McDermitt, D. J. & Tumbar, T. Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer. *The EMBO journal* **31**, 4124-4139 (2012).
170 Ramaswamy, S., Ross, K. N., Lander, E. S. & Golub, T. R. A molecular signature of metastasis in primary solid tumors. *Nature genetics* **33**, 49 (2002).
171 Giricz, O. *et al.* The RUNX1/IL-34/CSF-1R axis is an autocrinally regulated modulator of resistance to BRAF-V600E inhibition in melanoma. *JCI insight* **3** (2018).
172 Goyama, S., Huang, G., Kurokawa, M. & Mulloy, J. C. Posttranslational modifications of RUNX1 as potential anticancer targets. *Oncogene* **34**, 3483-3492, doi:10.1038/onc.2014.305 (2015).
173 Zhang, L., Fried, F. B., Guo, H. & Friedman, A. D. Cyclin-dependent kinase phosphorylation of RUNX1/AML1 on 3 sites increases transactivation potency and stimulates cell proliferation. *Blood* **111**, 1193-1200 (2008).
174 Biggs, J. R., Peterson, L. F., Zhang, Y., Kraft, A. S. & Zhang, D.-E. AML1/RUNX1 phosphorylation by cyclin-dependent kinases regulates the degradation of AML1/RUNX1 by the anaphase-promoting complex. *Molecular and cellular biology* **26**, 7420-7429 (2006).
175 Tanaka, T. *et al.* The extracellular signal-regulated kinase pathway phosphorylates AML1, an acute myeloid leukemia gene product, and potentially regulates its transactivation ability. *Molecular and cellular biology* **16**, 3967-3979 (1996).
176 Imai, Y. *et al.* The corepressor mSin3A regulates phosphorylation-induced activation, intranuclear location, and stability of AML1. *Mol Cell Biol* **24**, 1033-1043, doi:10.1128/mcb.24.3.1033-1043.2004 (2004).
177 Jin, Y. H. *et al.* Transforming growth factor-beta stimulates p300-dependent RUNX3 acetylation, which inhibits ubiquitination-mediated degradation. *J Biol Chem* **279**, 29409-29417, doi:10.1074/jbc.M313120200 (2004).
178 Wee, H. J., Voon, D. C., Bae, S. C. & Ito, Y. PEBP2-beta/CBF-beta-dependent phosphorylation of RUNX1 and p300 by HIPK2: implications for leukemogenesis. *Blood* **112**, 3777-3787, doi:10.1182/blood-2008-01-134122 (2008).
179 Yamaguchi, Y. *et al.* AML1 is functionally regulated through p300-mediated acetylation on specific lysine residues. *Journal of Biological Chemistry* **279**, 15630-15638 (2004).
180 Lazer, G. & Katzav, S. Guanine nucleotide exchange factors for RhoGTPases: good therapeutic targets for cancer therapy? *Cell Signal* **23**, 969-979, doi:10.1016/j.cellsig.2010.10.022 (2011).
181 Cote, J. F. & Vuori, K. Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. *J Cell Sci* **115**, 4901-4913, doi:10.1242/jcs.00219 (2002).
182 Cote, J. F. & Vuori, K. In vitro guanine nucleotide exchange activity of DHR-2/DOCKER/CZH2 domains. *Methods Enzymol* **406**, 41-57, doi:10.1016/S0076-6879(06)06004-6 (2006).
183 Kobayashi, S. *et al.* Membrane recruitment of DOCK180 by binding to PtdIns(3,4,5)P3. *Biochem J* **354**, 73-78, doi:10.1042/0264-6021:3540073 (2001).
184 Nishikimi, A. *et al.* Sequential regulation of DOCK2 dynamics by two phospholipids during neutrophil chemotaxis. *Science* **324**, 384-387, doi:10.1126/science.1170179 (2009).
185 Sanematsu, F. *et al.* Phosphatidic acid-dependent recruitment and function of the Rac activator DOCK1 during dorsal ruffle formation. *J Biol Chem* **288**, 8092-8100, doi:10.1074/jbc.M112.410423 (2013).

186 Komander, D. *et al.* An alpha-helical extension of the ELMO1 pleckstrin homology domain mediates direct interaction to DOCK180 and is critical in Rac signaling. *Mol Biol Cell* **19**, 4837-4851, doi:10.1091/mbc.E08-04-0345 (2008).
187 Lu, M. *et al.* A Steric-inhibition model for regulation of nucleotide exchange via the Dock180 family of GEFs. *Curr Biol* **15**, 371-377, doi:10.1016/j.cub.2005.01.050 (2005).
188 Patel, M., Pelletier, A. & Cote, J. F. Opening up on ELMO regulation: New insights into the control of Rac signaling by the DOCK180/ELMO complex. *Small GTPases* **2**, 268-275, doi:10.4161/sgtp.2.5.17716 (2011).
189 Patel, M. *et al.* An evolutionarily conserved autoinhibitory molecular switch in ELMO proteins regulates Rac signaling. *Curr Biol* **20**, 2021-2027, doi:10.1016/j.cub.2010.10.028 (2010).
190 Makino, Y. *et al.* Elmo1 inhibits ubiquitylation of Dock180. *J Cell Sci* **119**, 923-932, doi:10.1242/jcs.02797 (2006).
191 Feng, H. *et al.* EGFRvIII stimulates glioma growth and invasion through PKA-dependent serine phosphorylation of Dock180. *Oncogene* **33**, 2504-2512, doi:10.1038/onc.2013.198 (2014).
192 Jarzynka, M. J. *et al.* ELMO1 and Dock180, a bipartite Rac1 guanine nucleotide exchange factor, promote human glioma cell invasion. *Cancer Res* **67**, 7203-7211, doi:10.1158/0008-5472.CAN-07-0473 (2007).
193 Feng, H. *et al.* Activation of Rac1 by Src-dependent phosphorylation of Dock180(Y1811) mediates PDGFRalpha-stimulated glioma tumorigenesis in mice and humans. *J Clin Invest* **121**, 4670-4684, doi:10.1172/JCI58559 (2011).
194 Laurin, M. *et al.* Rac-specific guanine nucleotide exchange factor DOCK1 is a critical regulator of HER2-mediated breast cancer metastasis. *Proc Natl Acad Sci U S A* **110**, 7434-7439, doi:10.1073/pnas.1213050110 (2013).
195 Feng, H. *et al.* Phosphorylation of dedicator of cytokinesis 1 (Dock180) at tyrosine residue Y722 by Src family kinases mediates EGFRvIII-driven glioblastoma tumorigenesis. *Proc Natl Acad Sci U S A* **109**, 3018-3023, doi:10.1073/pnas.1121457109 (2012).
196 Kim, M. *et al.* Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. *Cell* **125**, 1269-1281, doi:10.1016/j.cell.2006.06.008 (2006).
197 Tomino, T. *et al.* DOCK1 inhibition suppresses cancer cell invasion and macropinocytosis induced by self-activating Rac1(P29S) mutation. *Biochem Biophys Res Commun* **497**, 298-304, doi:10.1016/j.bbrc.2018.02.073 (2018).
198 Tajiri, H. *et al.* Targeting Ras-Driven Cancer Cell Survival and Invasion through Selective Inhibition of DOCK1. *Cell Rep* **19**, 969-980, doi:10.1016/j.celrep.2017.04.016 (2017).
199 Watson, I. R. *et al.* The RAC1 P29S hotspot mutation in melanoma confers resistance to pharmacological inhibition of RAF. *Cancer Res* **74**, 4845-4852, doi:10.1158/0008-5472.CAN-14-1232-T (2014).
200 Vu, H. L., Rosenbaum, S., Purwin, T. J., Davies, M. A. & Aplin, A. E. RAC1 P29S regulates PD-L1 expression in melanoma. *Pigment Cell Melanoma Res* **28**, 590-598, doi:10.1111/pcmr.12392 (2015).
201 Jansen, R., Embden, J. D., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565-1575, doi:10.1046/j.1365-2958.2002.02839.x (2002).
202 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).

203 Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579-2586, doi:10.1073/pnas.1208507109 (2012).

204 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).

205 Jinek, M. *et al.* RNA-programmed genome editing in human cells. *Elife* **2**, e00471, doi:10.7554/eLife.00471 (2013).

206 Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* **45**, 273-297, doi:10.1146/annurev-genet-110410-132430 (2011).

207 Wang, H., La Russa, M. & Qi, L. S. CRISPR/Cas9 in Genome Editing and Beyond. *Annual Review of Biochemistry* **85**, 227-264, doi:10.1146/annurev-biochem-060815-014607 (2016).

208 Chylinski, K., Le Rhun, A. & Charpentier, E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol* **10**, 726-737, doi:10.4161/rna.24321 (2013).

209 Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096, doi:10.1126/science.1258096 (2014).

210 Nishimasu, H. *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935-949, doi:10.1016/j.cell.2014.02.001 (2014).

211 Rodriguez-Rodriguez, D. R., Ramirez-Solis, R., Garza-Elizondo, M. A., Garza-Rodriguez, M. L. & Barrera-Saldana, H. A. Genome editing: A perspective on the application of CRISPR/Cas9 to study human diseases (Review). *Int J Mol Med* **43**, 1559-1574, doi:10.3892/ijmm.2019.4112 (2019).

212 Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636-646, doi:10.1038/nrg2842 (2010).

213 Christian, M. *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757-761, doi:10.1534/genetics.110.120717 (2010).

214 Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* **14**, 49-55, doi:10.1038/nrm3486 (2013).

215 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281-2308, doi:10.1038/nprot.2013.143.Genome (2013).

216 Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278, doi:10.1016/j.cell.2014.05.010 (2014).

217 Tian, X. *et al.* CRISPR/Cas9 - An evolving biological tool kit for cancer biology and oncology. *NPJ Precis Oncol* **3**, 8, doi:10.1038/s41698-019-0080-7 (2019).

218 Liu, M. *et al.* Methodologies for Improving HDR Efficiency. *Front Genet* **9**, 691, doi:10.3389/fgene.2018.00691 (2018).

219 Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol* **34**, 339-344, doi:10.1038/nbt.3481 (2016).

220 Ferreira da Silva, J. *et al.* Genome-scale CRISPR screens are efficient in non-homologous end-joining deficient cells. *Sci Rep* **9**, 15751, doi:10.1038/s41598-019-52078-9 (2019).

221 He, X. *et al.* Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res* **44**, e85, doi:10.1093/nar/gkw064 (2016).

222 Lino, C. A., Harper, J. C., Carney, J. P. & Timlin, J. A. Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv* **25**, 1234-1257, doi:10.1080/10717544.2018.1474964 (2018).
223 Sharma, A. *et al.* CRISPR/Cas9-Mediated Fluorescent Tagging of Endogenous Proteins in Human Pluripotent Stem Cells. *Curr Protoc Hum Genet* **96**, 21 11 21-21 11 20, doi:10.1002/cphg.52 (2018).
224 Cebrian-Serrano, A. & Davies, B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm Genome* **28**, 247-261, doi:10.1007/s00335-017-9697-4 (2017).
225 Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57-63, doi:10.1038/nature26155 (2018).
226 Hajiahmadi, Z. *et al.* Strategies to Increase On-Target and Reduce Off-Target Effects of the CRISPR/Cas9 System in Plants. *Int J Mol Sci* **20**, doi:10.3390/ijms20153719 (2019).
227 Shen, B. *et al.* Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods* **11**, 399-402, doi:10.1038/nmeth.2857 (2014).
228 Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol* **17**, 5-15, doi:10.1038/nrm.2015.2 (2016).
229 Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* **8**, 2180-2196, doi:10.1038/nprot.2013.132 (2013).
230 Chen, B., Zou, W., Xu, H., Liang, Y. & Huang, B. Efficient labeling and imaging of protein-coding genes in living cells using CRISPR-Tag. *Nat Commun* **9**, 5065, doi:10.1038/s41467-018-07498-y (2018).
231 Zhan, T., Rindtorff, N., Betge, J., Ebert, M. P. & Boutros, M. CRISPR/Cas9 for cancer research and therapy. *Semin Cancer Biol* **55**, 106-119, doi:10.1016/j.semcancer.2018.04.001 (2019).
232 White, M. K. & Khalili, K. CRISPR/Cas9 and cancer targets: future possibilities and present challenges. *Oncotarget* **7**, 12305-12317, doi:10.18632/oncotarget.7104 (2016).
233 Saunderson, E. A. *et al.* Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat Commun* **8**, 1450, doi:10.1038/s41467-017-01078-2 (2017).
234 Wang, H. & Sun, W. CRISPR-mediated targeting of HER2 inhibits cell proliferation through a dominant negative mutation. *Cancer Lett* **385**, 137-143, doi:10.1016/j.canlet.2016.10.033 (2017).
235 Heckl, D. *et al.* Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat Biotechnol* **32**, 941-946, doi:10.1038/nbt.2951 (2014).
236 Han, K. *et al.* Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat Biotechnol* **35**, 463-474, doi:10.1038/nbt.3834 (2017).
237 Bester, A. C. *et al.* An Integrated Genome-wide CRISPRa Approach to Functionalize lncRNAs in Drug Resistance. *Cell* **173**, 649-664 e620, doi:10.1016/j.cell.2018.03.052 (2018).
238 Chen, S. *et al.* Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* **160**, 1246-1260, doi:10.1016/j.cell.2015.02.038 (2015).
239 Shi, J. *et al.* Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat Biotechnol* **33**, 661-667, doi:10.1038/nbt.3235 (2015).
240 Gootenberg, J. S. *et al.* Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* **356**, 438-442, doi:10.1126/science.aam9321 (2017).

241 Abudayyeh, O. O. *et al.* RNA targeting with CRISPR-Cas13. *Nature* **550**, 280-284, doi:10.1038/nature24049 (2017).
242 Chen, J. S. *et al.* CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* **360**, 436-439, doi:10.1126/science.aar6245 (2018).
243 Ayoubi, T. A. & Van De Ven, W. J. Regulation of gene expression by alternative promoters. *FASEB J* **10**, 453-460 (1996).
244 Schneider, M. V. & Orchard, S. Omics technologies, data and bioinformatics principles. *Methods Mol Biol* **719**, 3-30, doi:10.1007/978-1-61779-027-0_1 (2011).
245 Virág, D. *et al.* Current Trends in the Analysis of Post-translational Modifications. *Chromatographia* **83**, 1-10, doi:10.1007/s10337-019-03796-9 (2019).
246 Larance, M. & Lamond, A. I. Multidimensional proteomics for cell biology. *Nat Rev Mol Cell Biol* **16**, 269-280, doi:10.1038/nrm3970 (2015).
247 Kim, M. S. *et al.* A draft map of the human proteome. *Nature* **509**, 575-581, doi:10.1038/nature13302 (2014).
248 Wilhelm, M. *et al.* Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582-587, doi:10.1038/nature13319 (2014).
249 Duan, G. & Walther, D. The roles of post-translational modifications in the context of protein interaction networks. *PLoS Comput Biol* **11**, e1004049, doi:10.1371/journal.pcbi.1004049 (2015).
250 Minguez, P. *et al.* Deciphering a global network of functionally associated post-translational modifications. *Mol Syst Biol* **8**, 599, doi:10.1038/msb.2012.31 (2012).
251 Beltrao, P., Bork, P., Krogan, N. J. & van Noort, V. Evolution and functional cross-talk of protein post-translational modifications. *Mol Syst Biol* **9**, 714, doi:10.1002/msb.201304521 (2013).
252 Sharma, K. *et al.* Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep* **8**, 1583-1594, doi:10.1016/j.celrep.2014.07.036 (2014).
253 Olsen, J. V. *et al.* Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal* **3**, ra3, doi:10.1126/scisignal.2000475 (2010).
254 Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat Struct Mol Biol* **17**, 666-672, doi:10.1038/nsmb.1842 (2010).
255 Allfrey, V. G. & Mirsky, A. E. Structural Modifications of Histones and their Possible Role in the Regulation of RNA Synthesis. *Science* **144**, 559, doi:10.1126/science.144.3618.559 (1964).
256 Pazin, M. J. & Kadonaga, J. T. What's up and down with histone deacetylation and transcription? *Cell* **89**, 325-328, doi:10.1016/s0092-8674(00)80211-1 (1997).
257 Kim, S. C. *et al.* Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell* **23**, 607-618, doi:10.1016/j.molcel.2006.06.026 (2006).
258 Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840, doi:10.1126/science.1175371 (2009).
259 Ciechanover, A. Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nat Rev Mol Cell Biol* **6**, 79-87, doi:10.1038/nrm1552 (2005).
260 Hunter, T. The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Molecular cell* **28**, 730-738, doi:10.1016/j.molcel.2007.11.019 (2007).
261 Minguez, P., Letunic, I., Parca, L. & Bork, P. PTMcode: a database of known and predicted functional associations between post-translational modifications in proteins. *Nucleic Acids Res* **41**, D306-311, doi:10.1093/nar/gks1230 (2013).
262 Swaney, D. L. *et al.* Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nat Methods* **10**, 676-682, doi:10.1038/nmeth.2519 (2013).

263 Yao, Q., Li, H., Liu, B. Q., Huang, X. Y. & Guo, L. SUMOylation-regulated protein phosphorylation, evidence from quantitative phosphoproteomics analyses. *J Biol Chem* **286**, 27342-27349, doi:10.1074/jbc.M111.220848 (2011).
264 http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1922/aston-lecture.html. , N. o. F. W. A.-N. L. M. S. a. I. N. M. A.
265 Finehout, E. J. & Lee, K. H. An introduction to mass spectrometry applications in biological research. *Biochem Mol Biol Educ* **32**, 93-100, doi:10.1002/bmb.2004.494032020331 (2004).
266 Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198-207 (2003).
267 Pappireddi, N., Martin, L. & Wühr, M. A review on quantitative multiplexed proteomics. *ChemBioChem* **20**, 1210-1224, doi:10.1002/ (2019).
268 Starostin, K. V. *et al.* Identification of Bacillus strains by MALDI TOF MS using geometric approach. *Sci Rep* **5**, 16989, doi:10.1038/srep16989 (2015).
269 Mourino-Alvarez, L. *et al.* MALDI-Imaging Mass Spectrometry: a step forward in the anatomopathological characterization of stenotic aortic valve tissue. *Sci Rep* **6**, 27106, doi:10.1038/srep27106 (2016).
270 Ho, C. S. *et al.* Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical Biochemist Reviews* **24**, 3 (2003).
271 Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nature reviews. Molecular cell biology* **5**, 699-711, doi:10.1038/nrm1468 (2004).
272 Scigelova, M., Hornshaw, M., Giannakopulos, A. & Makarov, A. Fourier transform mass spectrometry. *Mol Cell Proteomics* **10**, M111 009431, doi:10.1074/mcp.M111.009431 (2011).
273 Han, X., Aslanian, A. & Yates, J. R., 3rd. Mass spectrometry for proteomics. *Curr Opin Chem Biol* **12**, 483-490, doi:10.1016/j.cbpa.2008.07.024 (2008).
274 Scheltema, R. A. *et al.* The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. *Mol Cell Proteomics* **13**, 3698-3708, doi:10.1074/mcp.M114.043489 (2014).
275 Schwartz, J. C., Senko, M. W. & Syka, J. E. A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom* **13**, 659-669, doi:10.1016/S1044-0305(02)00384-7 (2002).
276 Second, T. P. *et al.* Dual-pressure linear ion trap mass spectrometer improving the analysis of complex protein mixtures. *Anal Chem* **81**, 7757-7765, doi:10.1021/ac901278y (2009).
277 Hu, Q. *et al.* The Orbitrap: a new mass spectrometer. *J Mass Spectrom* **40**, 430-443, doi:10.1002/jms.856 (2005).
278 Makarov, A., Denisov, E., Lange, O. & Horning, S. Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. *J Am Soc Mass Spectrom* **17**, 977-982, doi:10.1016/j.jasms.2006.03.006 (2006).
279 Zubarev, R. A. & Makarov, A. Orbitrap mass spectrometry. *Anal Chem* **85**, 5288-5296, doi:10.1021/ac4001223 (2013).
280 Shen, Y., Tolic, N., Purvine, S. O. & Smith, R. D. Improving collision induced dissociation (CID), high energy collision dissociation (HCD), and electron transfer dissociation (ETD) fourier transform MS/MS degradome-peptidome identifications using high accuracy mass information. *Journal of proteome research* **11**, 668-677, doi:10.1021/pr200597j (2012).
281 Potel, C. M., Lemeer, S. & Heck, A. J. R. Phosphopeptide Fragmentation and Site Localization by Mass Spectrometry: An Update. *Anal Chem* **91**, 126-141, doi:10.1021/acs.analchem.8b04746 (2019).

282 Liu, J. & McLuckey, S. A. Electron Transfer Dissociation: Effects of Cation Charge State on Product Partitioning in Ion/Ion Electron Transfer to Multiply Protonated Polypeptides. *Int J Mass Spectrom* **330-332**, 174-181, doi:10.1016/j.ijms.2012.07.013 (2012).

283 Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-3567, doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2 (1999).

284 Diament, B. J. & Noble, W. S. Faster SEQUEST searching for peptide identification from tandem mass spectra. *Journal of proteome research* **10**, 3871-3879, doi:10.1021/pr101196n (2011).

285 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26**, 1367-1372, doi:10.1038/nbt.1511 (2008).

286 Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* **10**, 1794-1805, doi:10.1021/pr101065j (2011).

287 Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* **5**, 976-989, doi:10.1016/1044-0305(94)80016-2 (1994).

288 Gillet, L. C., Leitner, A. & Aebersold, R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annu Rev Anal Chem (Palo Alto Calif)* **9**, 449-472, doi:10.1146/annurev-anchem-071015-041535 (2016).

289 Burkhart, J. M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A. & Zahedi, R. P. Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. *J Proteomics* **75**, 1454-1462, doi:10.1016/j.jprot.2011.11.016 (2012).

290 Liigand, P., Kaupmees, K. & Kruve, A. Influence of the amino acid composition on the ionization efficiencies of small peptides. *J Mass Spectrom* **54**, 481-487, doi:10.1002/jms.4348 (2019).

291 Giansanti, P., Tsiatsiani, L., Low, T. Y. & Heck, A. J. Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat Protoc* **11**, 993-1006, doi:10.1038/nprot.2016.057 (2016).

292 Gaspari, M. & Cuda, G. Nano LC-MS/MS: a robust setup for proteomic analysis. *Methods Mol Biol* **790**, 115-126, doi:10.1007/978-1-61779-319-6_9 (2011).

293 Batth, T. S. & Olsen, J. V. in *Phospho-Proteomics: Methods and Protocols* (ed Louise von Stechow) 179-192 (Springer New York, 2016).

294 Zubarev, R. A. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* **13**, 723-726, doi:10.1002/pmic.201200451 (2013).

295 Manadas, B., Mendes, V. M., English, J. & Dunn, M. J. Peptide fractionation in proteomics approaches. *Expert Rev Proteomics* **7**, 655-663, doi:10.1586/epr.10.46 (2010).

296 Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* **2**, 1896-1906, doi:10.1038/nprot.2007.261 (2007).

297 Zhao, Y. & Jensen, O. N. Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* **9**, 4632-4641, doi:10.1002/pmic.200900398 (2009).

298 Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal Biochem* **203**, 173-179, doi:10.1016/0003-2697(92)90061-b (1992).
299 Wilm, M. *et al.* Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379**, 466-469, doi:10.1038/379466a0 (1996).
300 Edelmann, M. J. Strong cation exchange chromatography in analysis of posttranslational modifications: innovations and perspectives. *J Biomed Biotechnol* **2011**, 936508, doi:10.1155/2011/936508 (2011).
301 Doll, S. & Burlingame, A. L. Mass spectrometry-based detection and assignment of protein posttranslational modifications. *ACS Chem Biol* **10**, 63-71, doi:10.1021/cb500904b (2015).
302 Mann, M., Jensen, O. N. . Proteomic analysis of post-translational modifications. *Nature Biotechnology* **21**, 255–261, doi:10.1038/nbt0303-255 (2003).
303 Riley, N. M. & Coon, J. J. Phosphoproteomics in the Age of Rapid and Deep Proteome Profiling. *Anal Chem* **88**, 74-94, doi:10.1021/acs.analchem.5b04123 (2016).
304 Cho, K. C., Chen, L., Hu, Y., Schnaubelt, M. & Zhang, H. Developing Workflow for Simultaneous Analyses of Phosphopeptides and Glycopeptides. *ACS Chem Biol* **14**, 58-66, doi:10.1021/acschembio.8b00902 (2019).
305 Mann, M. *et al.* Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends in Biotechnology* **20**, 261-268, doi:10.1016/S0167-7799(02)01944-3 (2002).
306 Bian, Y. *et al.* Ultra-deep tyrosine phosphoproteomics enabled by a phosphotyrosine superbinder. *Nat Chem Biol* **12**, 959-966, doi:10.1038/nchembio.2178 (2016).
307 Li, Q. R., Ning, Z. B., Tang, J. S., Nie, S. & Zeng, R. Effect of peptide-to-TiO2 beads ratio on phosphopeptide enrichment selectivity. *Journal of proteome research* **8**, 5375-5381, doi:10.1021/pr900659n (2009).
308 Zarei, M., Sprenger, A., Rackiewicz, M. & Dengjel, J. Fast and easy phosphopeptide fractionation by combinatorial ERLIC-SCX solid-phase extraction for in-depth phosphoproteome analysis. *Nat Protoc* **11**, 37-45, doi:10.1038/nprot.2015.134 (2016).
309 Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212-217, doi:10.1126/science.1124619 (2006).
310 Macek, B., Mann, M. & Olsen, J. V. Global and site-specific quantitative phosphoproteomics: principles and applications. *Annual review of pharmacology and toxicology* **49**, 199-221, doi:10.1146/annurev.pharmtox.011008.145606 (2009).
311 Ong, S. E. & Mann, M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* **1**, 252-262, doi:10.1038/nchembio736 (2005).
312 Lindemann, C. *et al.* Strategies in relative and absolute quantitative mass spectrometry based proteomics. *Biol Chem* **398**, 687-699, doi:10.1515/hsz-2017-0104 (2017).
313 Nahnsen, S., Bielow, C., Reinert, K. & Kohlbacher, O. Tools for label-free peptide quantification. *Mol Cell Proteomics* **12**, 549-556, doi:10.1074/mcp.R112.025163 (2013).
314 Zhou, J. Y. *et al.* Improved LC-MS/MS spectral counting statistics by recovering low-scoring spectra matched to confidently identified peptide sequences. *Journal of proteome research* **9**, 5698-5704, doi:10.1021/pr100508p (2010).
315 Karpievitch, Y. V., Dabney, A. R. & Smith, R. D. Normalization and missing value imputation for label-free LC-MS analysis. *BMC Bioinformatics* **13 Suppl 16**, S5, doi:10.1186/1471-2105-13-S16-S5 (2012).
316 Ong, S. E., Blagoev, B. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics*, doi:10.1074/ (2002).
317 Mann, M. Functional and quantitative proteomics using SILAC. *Nature* **7** (2006).

318 Zhang, G., Fenyo, D. & Neubert, T. A. Evaluation of the variation in sample preparation for comparative proteomics using stable isotope labeling by amino acids in cell culture. *Journal of proteome research* **8**, 1285-1292, doi:10.1021/pr8006107 (2009).
319 Liu, T., Belov, M. E., Jaitly, N., Qian, W. J. & Smith, R. D. Accurate mass measurements in proteomics. *Chem Rev* **107**, 3621-3653, doi:10.1021/cr068288j (2007).
320 Kirchner, M. & Selbach, M. In vivo quantitative proteome profiling: planning and evaluation of SILAC experiments. *Methods Mol Biol* **893**, 175-199, doi:10.1007/978-1-61779-885-6_13 (2012).
321 Jiang, H. & English, A. M. Quantitative analysis of the yeast proteome by incorporation of isotopically labeled leucine. *Journal of proteome research* **1**, 345-350, doi:10.1021/pr025523f (2002).
322 de Godoy, L. M. *et al.* Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251-1254, doi:10.1038/nature07341 (2008).
323 Kruger, M. *et al.* SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* **134**, 353-364, doi:10.1016/j.cell.2008.05.033 (2008).
324 Soufi, B. *et al.* Stable isotope labeling by amino acids in cell culture (SILAC) applied to quantitative proteomics of Bacillus subtilis. *Journal of proteome research* **9**, 3638-3646, doi:10.1021/pr100150w (2010).
325 Bantscheff, M., Lemeer, S., Savitski, M. M. & Kuster, B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Analytical and bioanalytical chemistry* **404**, 939-965 (2012).
326 Gygi, S. P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994-999, doi:10.1038/13690 (1999).
327 Karp, N. A. *et al.* Addressing accuracy and precision issues in iTRAQ quantitation. *Mol Cell Proteomics* **9**, 1885-1897, doi:10.1074/mcp.M900628-MCP200 (2010).
328 Thompson, A. *et al.* Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **75**, 1895-1904, doi:10.1021/ac0262560 (2003).
329 Shiio, Y. & Aebersold, R. Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. *Nature protocols* **1**, 139 (2006).
330 Rauniyar, N. & Yates, J. R., 3rd. Isobaric labeling-based relative quantification in shotgun proteomics. *Journal of proteome research* **13**, 5293-5309, doi:10.1021/pr500880b (2014).
331 Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* **4**, 484-494, doi:10.1038/nprot.2009.21 (2009).
332 Sheynkman, G. M., Shortreed, M. R., Cesnik, A. J. & Smith, L. M. Proteogenomics: Integrating Next-Generation Sequencing and Mass Spectrometry to Characterize Human Proteomic Variation. *Annu Rev Anal Chem (Palo Alto Calif)* **9**, 521-545, doi:10.1146/annurev-anchem-071015-041722 (2016).
333 Krug, K., Nahnsen, S. & Macek, B. Mass spectrometry at the interface of proteomics and genomics. *Mol Biosyst* **7**, 284-291, doi:10.1039/c0mb00168f (2011).
334 Battle, A. *et al.* Genomic variation. Impact of regulatory variation from RNA to protein. *Science* **347**, 664-667, doi:10.1126/science.1260793 (2015).
335 Foss, E. J. *et al.* Genetic variation shapes protein networks mainly through non-transcriptional mechanisms. *PLoS Biol* **9**, e1001144, doi:10.1371/journal.pbio.1001144 (2011).

336 Jaffe, J. D., Berg, H. C. & Church, G. M. Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* **4**, 59-77, doi:10.1002/pmic.200300511 (2004).
337 Brent, M. R. Steady progress and recent breakthroughs in the accuracy of automated genome annotation. *Nat Rev Genet* **9**, 62-73, doi:10.1038/nrg2220 (2008).
338 Mann, M., Kulak, N. A., Nagaraj, N. & Cox, J. The coming age of complete, accurate, and ubiquitous proteomes. *Mol Cell* **49**, 583-590, doi:10.1016/j.molcel.2013.01.029 (2013).
339 Lappalainen, T. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**, 506-511, doi:10.1038/nature12531 (2013).
340 Creixell, P. *et al.* Kinome-wide Decoding of Network-Attacking Mutations Rewiring Cancer Signaling. *Cell* **163**, 202-217, doi:10.1016/j.cell.2015.08.056 (2015).
341 Choi, M. *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* **106**, 19096-19101, doi:10.1073/pnas.0910672106 (2009).
342 Hunt, S. E. *et al.* Ensembl variation resources. *Database (Oxford)* **2018**, doi:10.1093/database/bay119 (2018).
343 O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44**, D733-745, doi:10.1093/nar/gkv1189 (2016).
344 UniProt Consortium, T. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **46**, 2699, doi:10.1093/nar/gky092 (2018).
345 Ma, B. & Johnson, R. De novo sequencing and homology searching. *Mol Cell Proteomics* **11**, O111 014902, doi:10.1074/mcp.O111.014902 (2012).
346 Ang, M. Y. *et al.* Proteogenomics: From next-generation sequencing (NGS) and mass spectrometry-based proteomics to precision medicine. *Clin Chim Acta* **498**, 38-46, doi:10.1016/j.cca.2019.08.010 (2019).
347 Trapnell, C. & Salzberg, S. L. How to map billions of short reads onto genomes. *Nat Biotechnol* **27**, 455-457, doi:10.1038/nbt0509-455 (2009).
348 Nesvizhskii, A. I. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. *J Proteomics* **73**, 2092-2123, doi:10.1016/j.jprot.2010.08.009 (2010).
349 Krug, K. *et al.* Deep coverage of the Escherichia coli proteome enables the assessment of false discovery rates in simple proteogenomic experiments. *Mol Cell Proteomics* **12**, 3420-3430, doi:10.1074/mcp.M113.029165 (2013).
350 Zhang, K. *et al.* A note on the false discovery rate of novel peptides in proteogenomics. *Bioinformatics* **31**, 3249-3253, doi:10.1093/bioinformatics/btv340 (2015).
351 Wang, X. & Zhang, B. customProDB: an R package to generate customized protein databases from RNA-Seq data for proteomics search. *Bioinformatics* **29**, 3235-3237, doi:10.1093/bioinformatics/btt543 (2013).
352 Sheynkman, G. M. *et al.* Using Galaxy-P to leverage RNA-Seq for the discovery of novel protein variations. *BMC Genomics* **15**, 703, doi:10.1186/1471-2164-15-703 (2014).
353 Krasnov, G. S. *et al.* PPLine: An Automated Pipeline for SNP, SAP, and Splice Variant Detection in the Context of Proteogenomics. *Journal of proteome research* **14**, 3729-3737, doi:10.1021/acs.jproteome.5b00490 (2015).
354 Wen, B. *et al.* PGA: an R/Bioconductor package for identification of novel peptides using a customized database derived from RNA-Seq. *BMC Bioinformatics* **17**, 244, doi:10.1186/s12859-016-1133-3 (2016).
355 Ruggles, K. V. *et al.* Methods, Tools and Current Perspectives in Proteogenomics. *Mol Cell Proteomics* **16**, 959-981, doi:10.1074/mcp.MR117.000024 (2017).

356 Ruggles, K. V. *et al.* An Analysis of the Sensitivity of Proteogenomic Mapping of Somatic Mutations and Novel Splicing Events in Cancer. *Mol Cell Proteomics* **15**, 1060-1071, doi:10.1074/mcp.M115.056226 (2016).
357 Woo, S. *et al.* Proteogenomic database construction driven from large scale RNA-seq data. *Journal of proteome research* **13**, 21-28, doi:10.1021/pr400294c (2014).
358 Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308-311, doi:10.1093/nar/29.1.308 (2001).
359 Alfaro, J. A., Sinha, A., Kislinger, T. & Boutros, P. C. Onco-proteogenomics: cancer proteomics joins forces with genomics. *Nat Methods* **11**, 1107-1113, doi:10.1038/nmeth.3138 (2014).
360 Rivers, R. C. *et al.* Linking cancer genome to proteome: NCI's investment into proteogenomics. *Proteomics* **14**, 2633-2636, doi:10.1002/pmic.201400193 (2014).
361 Zhang, B. *et al.* Proteogenomic characterization of human colon and rectal cancer. *Nature* **513**, 382-387, doi:10.1038/nature13438 (2014).
362 Mertins, P. *et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**, 55-62, doi:10.1038/nature18003 (2016).
363 Woo, S. *et al.* Proteogenomic strategies for identification of aberrant cancer peptides using large-scale next-generation sequencing data. *Proteomics* **14**, 2719-2730, doi:10.1002/pmic.201400206 (2014).
364 Alfaro, J. A. *et al.* Detecting protein variants by mass spectrometry: a comprehensive study in cancer cell-lines. *Genome Med* **9**, 62, doi:10.1186/s13073-017-0454-9 (2017).
365 Sun, H. *et al.* Identification of gene fusions from human lung cancer mass spectrometry data. *BMC Genomics* **14 Suppl 8**, S5, doi:10.1186/1471-2164-14-S8-S5 (2013).
366 Reimand, J. & Bader, G. D. Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. *Mol Syst Biol* **9**, 637, doi:10.1038/msb.2012.68 (2013).
367 Gentile, S. *et al.* The human ERG1 channel polymorphism, K897T, creates a phosphorylation site that inhibits channel activity. *Proc Natl Acad Sci U S A* **105**, 14704-14708, doi:10.1073/pnas.0802250105 (2008).
368 Creixell, P. *et al.* Unmasking determinants of specificity in the human kinome. *Cell* **163**, 187-201, doi:10.1016/j.cell.2015.08.057 (2015).
369 Ryu, G. M. *et al.* Genome-wide analysis to predict protein sequence variations that change phosphorylation sites or their corresponding kinases. *Nucleic Acids Res* **37**, 1297-1307, doi:10.1093/nar/gkn1008 (2009).
370 Keegan, S., Cortens, J. P., Beavis, R. C. & Fenyo, D. g2pDB: A Database Mapping Protein Post-Translational Modifications to Genomic Coordinates. *Journal of proteome research* **15**, 983-990, doi:10.1021/acs.jproteome.5b01018 (2016).
371 Yang, C. Y. *et al.* PhosphoPOINT: a comprehensive human kinase interactome and phospho-protein database. *Bioinformatics* **24**, i14-20, doi:10.1093/bioinformatics/btn297 (2008).
372 Zhang, H. *et al.* Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell* **166**, 755-765, doi:10.1016/j.cell.2016.05.069 (2016).
373 Nishimura, T. & Nakamura, H. Developments for Personalized Medicine of Lung Cancer Subtypes: Mass Spectrometry-Based Clinical Proteogenomic Analysis of Oncogenic Mutations. *Adv Exp Med Biol* **926**, 115-137, doi:10.1007/978-3-319-42316-6_8 (2016).
374 Kondo, T. Proteogenomics for the Study of Gastrointestinal Stromal Tumors. *Adv Exp Med Biol* **926**, 139-151, doi:10.1007/978-3-319-42316-6_9 (2016).
375 Mardis, E. R. A decade's perspective on DNA sequencing technology. *Nature* **470**, 198-203, doi:10.1038/nature09796 (2011).

376 Geiger, T., Wehner, A., Schaab, C., Cox, J. & Mann, M. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Molecular & cellular proteomics : MCP* **11**, M111 014050, doi:10.1074/mcp.M111.014050 (2012).
377 Mardamshina, M. & Geiger, T. Next-Generation Proteomics and Its Application to Clinical Breast Cancer Research. *Am J Pathol* **187**, 2175-2184, doi:10.1016/j.ajpath.2017.07.003 (2017).
378 Olsen, J. V. & Mann, M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Molecular & cellular proteomics : MCP* **12**, 3444-3452, doi:10.1074/mcp.O113.034181 (2013).
379 von Stechow, L., Francavilla, C. & Olsen, J. V. Recent findings and technological advances in phosphoproteomics for cells and tissues. *Expert Rev Proteomics* **12**, 469-487, doi:10.1586/14789450.2015.1078730 (2015).
380 Huang, K. L. *et al.* Proteogenomic integration reveals therapeutic targets in breast cancer xenografts. *Nat Commun* **8**, 14864, doi:10.1038/ncomms14864 (2017).
381 Ren, J. *et al.* PhosSNP for systematic analysis of genetic polymorphisms that influence protein phosphorylation. *Molecular & cellular proteomics : MCP* **9**, 623-634, doi:10.1074/mcp.M900273-MCP200 (2010).
382 Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* **43**, D512-520, doi:10.1093/nar/gku1267 (2015).
383 Granger, B. R. *et al.* Visualization of Metabolic Interaction Networks in Microbial Communities Using VisANT 5.0. *PLoS Comput Biol* **12**, e1004875, doi:10.1371/journal.pcbi.1004875 (2016).
384 Shi, Z., Wang, J. & Zhang, B. NetGestalt: integrating multidimensional omics data over biological networks. *Nature methods* **10**, 597-598, doi:10.1038/nmeth.2517 (2013).
385 Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* **46**, W537-W544, doi:10.1093/nar/gky379 (2018).
386 Leitner, A. *et al.* Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proceedings of the National Academy of Sciences* **111**, 9455-9460, doi:10.1073/pnas.1320298111 (2014).
387 Sinnberg, T. *et al.* A Nexus Consisting of Beta-Catenin and Stat3 Attenuates BRAF Inhibitor Efficacy and Mediates Acquired Resistance to Vemurafenib. *EBioMedicine* **8**, 132-149, doi:10.1016/j.ebiom.2016.04.037 (2016).
388 shiny: Web Application Framework for R (2019).
389 Rossi, S. *et al.* TNF-alpha and metalloproteases as key players in melanoma cells aggressiveness. *J Exp Clin Cancer Res* **37**, 326, doi:10.1186/s13046-018-0982-1 (2018).
390 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* **17**, 2994-3004, doi:10.1093/emboj/17.11.2994 (1998).
391 Leong, W. Y. *et al.* Runx1 Phosphorylation by Src Increases Trans-activation via Augmented Stability, Reduced Histone Deacetylase (HDAC) Binding, and Increased DNA Affinity, and Activated Runx1 Favors Granulopoiesis. *J Biol Chem* **291**, 826-836, doi:10.1074/jbc.M115.674234 (2016).
392 Hu, H., Bliss, J. M., Wang, Y. & Colicelli, J. RIN1 is an ABL tyrosine kinase activator and a regulator of epithelial-cell adhesion and migration. *Curr Biol* **15**, 815-823, doi:10.1016/j.cub.2005.03.049 (2005).
393 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* **105**, 6308-6313, doi:10.1073/pnas.0707601105 (2008).

394 Nesvizhskii, A. I. Proteogenomics: concepts, applications and computational strategies. *Nature methods* **11**, 1114-1125, doi:10.1038/nmeth.3144 (2014).

395 Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632 (2012).

396 Pfeifer, G. P., You, Y. H. & Besaratinia, A. Mutations induced by ultraviolet light. *Mutat Res* **571**, 19-31, doi:10.1016/j.mrfmmm.2004.06.057 (2005).

397 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-404, doi:10.1158/2159-8290.CD-12-0095 (2012).

398 Jayachandran, A. *et al.* Thrombospondin 1 promotes an aggressive phenotype through epithelial-to-mesenchymal transition in human melanoma. *Oncotarget* **5**, 5782-5797, doi:10.18632/oncotarget.2164 (2014).

399 Liu, B. *et al.* Genomic landscape and mutational impacts of recurrently mutated genes in cancers. *Mol Genet Genomic Med* **6**, 910-923, doi:10.1002/mgg3.458 (2018).

400 Kato, M. *et al.* Transgenic mouse model for skin malignant melanoma. *Oncogene* **17**, 1885-1888, doi:10.1038/sj.onc.1202077 (1998).

401 Kren, N. P., Zagon, I. S. & McLaughlin, P. J. Mutations in the opioid growth factor receptor in human cancers alter receptor function. *Int J Mol Med* **36**, 289-293, doi:10.3892/ijmm.2015.2221 (2015).

402 Duffy, D. L. *et al.* Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nat Commun* **9**, 4774, doi:10.1038/s41467-018-06649-5 (2018).

403 Gumaste, P. V. *et al.* Skin cancer risk in BRCA1/2 mutation carriers. *Br J Dermatol* **172**, 1498-1506, doi:10.1111/bjd.13626 (2015).

404 Felder, M. *et al.* MUC16 suppresses human and murine innate immune responses. *Gynecol Oncol* **152**, 618-628, doi:10.1016/j.ygyno.2018.12.023 (2019).

405 Jonckheere, N. & Van Seuningen, I. Integrative analysis of the cancer genome atlas and cancer cell lines encyclopedia large-scale genomic databases: MUC4/MUC16/MUC20 signature is associated with poor survival in human carcinomas. *J Transl Med* **16**, 259, doi:10.1186/s12967-018-1632-2 (2018).

406 Riker, A. I. *et al.* The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. *BMC Med Genomics* **1**, 13, doi:10.1186/1755-8794-1-13 (2008).

407 Wang, D. *et al.* A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol Syst Biol* **15**, e8503, doi:10.15252/msb.20188503 (2019).

408 Wang, X. *et al.* Protein identification using customized protein sequence databases derived from RNA-Seq data. *Journal of proteome research* **11**, 1009-1017, doi:10.1021/pr200766z (2012).

409 Amaral, T. *et al.* The mitogen-activated protein kinase pathway in melanoma part I - Activation and primary resistance mechanisms to BRAF inhibition. *Eur J Cancer* **73**, 85-92, doi:10.1016/j.ejca.2016.12.010 (2017).

410 Cohen-Solal, K. A., Kaufman, H. L. & Lasfar, A. Transcription factors as critical players in melanoma invasiveness, drug resistance, and opportunities for therapeutic drug development. *Pigment Cell Melanoma Res* **31**, 241-252, doi:10.1111/pcmr.12666 (2018).

411 Han, S. *et al.* ERK-mediated phosphorylation regulates SOX10 sumoylation and targets expression in mutant BRAF melanoma. *Nat Commun* **9**, 28, doi:10.1038/s41467-017-02354-x (2018).

412 Lipinska, N. *et al.* Telomerase and drug resistance in cancer. *Cell Mol Life Sci* **74**, 4121-4132, doi:10.1007/s00018-017-2573-2 (2017).
413 Zhang, G. *et al.* Induction of Telomere Dysfunction Prolongs Disease Control of Therapy-Resistant Melanoma. *Clin Cancer Res* **24**, 4771-4784, doi:10.1158/1078-0432.CCR-17-2773 (2018).
414 Guo, H. & Friedman, A. D. Phosphorylation of RUNX1 by cyclin-dependent kinase reduces direct interaction with HDAC1 and HDAC3. *J Biol Chem* **286**, 208-215, doi:10.1074/jbc.M110.149013 (2011).
415 Schmitt, M. *et al.* Quantitative proteomics links the intermediate filament nestin to resistance to targeted BRAF inhibition in melanoma cells. *Molecular & cellular proteomics : MCP*, doi:10.1074/mcp.RA119.001302 (2019).
416 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-11 10 33, doi:10.1002/0471250953.bi1110s43 (2013).
417 R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2018).
418 Hu, J. *et al.* PhosphoNetworks: a database for human phosphorylation networks. *Bioinformatics* **30**, 141-142, doi:10.1093/bioinformatics/btt627 (2014).
419 Diella, F. *et al.* Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* **5**, 79, doi:10.1186/1471-2105-5-79 (2004).
420 Bamford, S. *et al.* The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* **91**, 355-358, doi:10.1038/sj.bjc.6601894 (2004).
421 Liu, Y., Sun, J. & Zhao, M. ONGene: A literature-based database for human oncogenes. *J Genet Genomics* **44**, 119-121, doi:10.1016/j.jgg.2016.12.004 (2017).
422 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308, doi:10.1038/nprot.2013.143 (2013).
423 Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* **10**, 1794-1805, doi:10.1021/pr101065j (2011).
424 Oughtred, R. *et al.* The BioGRID interaction database: 2019 update. *Nucleic Acids Res* **47**, D529-D541, doi:10.1093/nar/gky1079 (2019).
425 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).
426 Coordinators, N. R. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **46**, D8-D13, doi:10.1093/nar/gkx1095 (2018).
427 Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450, doi:10.1093/nar/gky1106 (2019).
428 Whole-exome sequencing pipeline v. 1.0 (Zenodo, 2020).
429 Leach, D. R., Krummel, M. F. & Allison, J. P. Enhancement of antitumor immunity by CTLA-4 blockade. *Science* **271**, 1734-1736, doi:10.1126/science.271.5256.1734 (1996).
430 Hirano, F. *et al.* Blockade of B7-H1 and PD-1 by monoclonal antibodies potentiates cancer therapeutic immunity. *Cancer Res* **65**, 1089-1096 (2005).
431 Ott, P. A., Hodi, F. S. & Robert, C. CTLA-4 and PD-1/PD-L1 blockade: new immunotherapeutic modalities with durable clinical benefit in melanoma patients. *Clin Cancer Res* **19**, 5300-5309, doi:10.1158/1078-0432.CCR-13-0143 (2013).
432 Robert, C. *et al.* Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N Engl J Med* **364**, 2517-2526, doi:10.1056/NEJMoa1104621 (2011).

433 Topalian, S. L. *et al.* Survival, durable tumor remission, and long-term safety in patients with advanced melanoma receiving nivolumab. *J Clin Oncol* **32**, 1020-1030, doi:10.1200/JCO.2013.53.0105 (2014).
434 Jenkins, R. W., Barbie, D. A. & Flaherty, K. T. Mechanisms of resistance to immune checkpoint inhibitors. *Br J Cancer* **118**, 9-16, doi:10.1038/bjc.2017.434 (2018).
435 Jamieson, N. B. & Maker, A. V. Gene-expression profiling to predict responsiveness to immunotherapy. *Cancer Gene Ther* **24**, 134-140, doi:10.1038/cgt.2016.63 (2017).
436 Harel, M. *et al.* Proteomics of Melanoma Response to Immunotherapy Reveals Mitochondrial Dependence. *Cell* **179**, 236-250 e218, doi:10.1016/j.cell.2019.08.012 (2019).
437 Kemper, K. *et al.* Intra- and inter-tumor heterogeneity in a vemurafenib-resistant melanoma patient and derived xenografts. *EMBO molecular medicine* **7**, e201404914-e201404914, doi:10.15252/emmm.201404914 (2015).
438 Creixell, P. *et al.* Kinome-wide decoding of netwok-attacking mutations rewiring cancer signaling. *Cell* **163**, 202-217 (2015).
439 Curtin, J. a. *et al.* Distinct sets of genetic alterations in melanoma. *The New England journal of medicine* **353**, 2135-2147, doi:10.1056/NEJMoa050092 (2005).
440 Roesch, A. Tumor heterogeneity and plasticity as elusive drivers for resistance to MAPK pathway inhibition in melanoma. *Oncogene* **34**, 2951-2957, doi:10.1038/onc.2014.249 (2015).
441 Laurin, M. & Cote, J. F. Insights into the biological functions of Dock family guanine nucleotide exchange factors. *Genes Dev* **28**, 533-547, doi:10.1101/gad.236349.113 (2014).
442 Feng, H. *et al.* Protein kinase A-dependent phosphorylation of Dock180 at serine residue 1250 is important for glioma growth and invasion stimulated by platelet derived-growth factor receptor alpha. *Neuro Oncol* **17**, 832-842, doi:10.1093/neuonc/nou323 (2015).
443 Valles, A. M., Beuvin, M. & Boyer, B. Activation of Rac1 by paxillin-Crk-DOCK180 signaling complex is antagonized by Rap1 in migrating NBT-II cells. *J Biol Chem* **279**, 44490-44496, doi:10.1074/jbc.M405144200 (2004).
444 Toret, C. P., Collins, C. & Nelson, W. J. An Elmo-Dock complex locally controls Rho GTPases and actin remodeling during cadherin-mediated adhesion. *J Cell Biol* **207**, 577-587, doi:10.1083/jcb.201406135 (2014).
445 Snyder, A. *et al.* Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* **371**, 2189-2199, doi:10.1056/NEJMoa1406498 (2014).
446 Berger, M. F. *et al.* Integrative analysis of the melanoma transcriptome. *Genome Res* **20**, 413-427, doi:10.1101/gr.103697.109 (2010).
447 Hugo, W. *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **165**, 35-44, doi:10.1016/j.cell.2016.02.065 (2016).
448 Lobas, A. A. *et al.* Proteogenomics of Malignant Melanoma Cell Lines: The Effect of Stringency of Exome Data Filtering on Variant Peptide Identification in Shotgun Proteomics. *Journal of proteome research* **17**, 1801-1811, doi:10.1021/acs.jproteome.7b00841 (2018).
449 Fukumura, D., Kloepper, J., Amoozgar, Z., Duda, D. G. & Jain, R. K. Enhancing cancer immunotherapy using antiangiogenics: opportunities and challenges. *Nat Rev Clin Oncol* **15**, 325-340, doi:10.1038/nrclinonc.2018.29 (2018).
450 Pinto, S. M. *et al.* Quantitative phosphoproteomic analysis of IL-33-mediated signaling. *Proteomics* **15**, 532-544, doi:10.1002/pmic.201400303 (2015).
451 Horak, P. *et al.* Precision oncology based on omics data: The NCT Heidelberg experience. *Int J Cancer* **141**, 877-886, doi:10.1002/ijc.30828 (2017).

452 FastQC A Quality Control tool for High Throughput Sequence Data (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, 2010).

453 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

454 Grozav, A. G. *et al.* Casein kinase I delta/epsilon phosphorylates topoisomerase IIalpha at serine-1106 and modulates DNA cleavage activity. *Nucleic acids research* **37**, 382-392, doi:10.1093/nar/gkn934 (2009).

455 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80-92, doi:10.4161/fly.19695 (2012).

456 Zhang, T., Dutton-Regester, K., Brown, K. M. & Hayward, N. K. The genomic landscape of cutaneous melanoma. *Pigment Cell Melanoma Res* **29**, 266-283, doi:10.1111/pcmr.12459 (2016).

457 Thakur, R. *et al.* Transcriptomic Analysis Reveals Prognostic Molecular Signatures of Stage I Melanoma. *Clin Cancer Res* **25**, 7424-7435, doi:10.1158/1078-0432.CCR-18-3659 (2019).

458 Matsuda, Y. *et al.* Nestin is a novel target for suppressing pancreatic cancer cell migration , invasion and metastasis. *Cancer Biology & Therapy* **11**, 512-523, doi:10.4161/cbt.11.5.14673 (2011).

459 Griffin M, D. S., Debra H. Josephs, Silvia Mele, Silvia Crescioli, Heather J. Bax, Giulia Pellizzari, Matthew D. Wynne, Mano Nakamura, Ricarda M. Hoffmann, Kristina M. Ilieva, Anthony Cheung, James F. Spicer, Sophie Papa, Katie E. Lacy and Sophia N. Karagiannis. BRAF inhibitors: resistance and the promise of combination treatments for melanoma. *Oncotarget* **8**, 78174-75192 (2017).

8 Acknowledgements

An dieser Stelle möchte ich mich bei allen Personen, die zum Gelingen der Dissertation beigetragen haben, bedanken.

Prof. Dr. Boris Macek möchte ich für die hervorragende Betreuung, seine stetige Diskussionsbereitschaft, das in mich gesetzte Vertrauen und das tolle Arbeitsumfeld herzlich danken.

Mein Dank gilt ebenfalls Prof. Dr. Alfred Nordheim für seine Bereitschaft, diese Arbeit zu betreuen.

Bei meinem TAC Komitee bestehend aus Prof. Dr. Boris Macek, Prof. Dr. Alfred Nordheim und Prof. Dr. Ana J. Garcia-Saez, sowie dem ganzen IMPRS Team möchte ich mich recht herzlich für die Unterstützung, Diskussionsbereitschaft und Betreuung meiner Arbeit bedanken.

Des Weiterem bedanke ich mich herzlich bei Nicolas C. Nalpas für seine ständige Hilfe bei allen Fragestellungen, sowie Tobias Sinnberg und Heike Niessner von der Hautklink in Tübingen für die Hilfe bei klinischen und biologischen Fragen.

Allen gegenwärtigen und ehemaligen Mitarbeitern des Proteome Center Tübingens danke ich für das freundliche Arbeitsklima, insbesondere danke ich Uli, Silke, Ana, Johannes, Irina und Mirita. Für die motivierenden Worte, die schönen Momente und die Unterstützung möchte ich mich bei Christoph, Nicolas, Maja, Katharina, Katrin, Phil und Tariq bedanken. Ihr wurdet von Arbeitskollegen zu echten Freunden.

Meinen Schwestern Kathrin, Anne und Sophie mit den Kindern Mia, Levi, Zoe und Lian möchte ich mich von ganzen Herzen für die moralische Unterstützung danken. Ihr habt immer für eine Abwechslung gesorgt und mich immer wieder zum Lachen gebracht.

Meinem Freund Julian möchte ich für seine stetige Motivation, das Aushalten aller Launen und seinen Humor in allen Lebenslagen danken.

Mein ganz besonderer Dank gilt meinen Eltern, Jürgen und Gerda, für ihre unendliche Liebe und Vertrauen, ihre große Unterstützung in den vergangenen Jahren. Danke dass ihr mich immer ermutigt habt meinen eigenen Weg zu gehen. Ohne euch wäre dies nicht möglich gewesen.