LUND UNIVERSITY

A protein deep sequencing evaluation of metastatic melanoma tissues.

Welinder, Charlotte; Pawlowski, Krzysztof; Sugihara, Yutaka; Yakovleva, Maria; Jönsson, Göran B; Ingvar, Christian; Lundgren, Lotta; Baldetorp, Bo; Olsson, Håkan; Rezeli, Melinda; Jansson, Bo; Laurell, Thomas; Fehniger, Thomas; Döme, Balazs; Malm, Johan; Wieslander, Elisabet; Nishimura, Toshihide; Marko-Varga, György

Published in:
PLoS ONE

DOI:
10.1371/journal.pone.0123661

2015

Link to publication

Citation for published version (APA):
Welinder, C., Pawlowski, K., Sugihara, Y., Yakovleva, M., Jönsson, G. B., Ingvar, C., Lundgren, L., Baldetorp, B., Olsson, H., Rezeli, M., Jansson, B., Laurell, T., Fehniger, T., Döme, B., Malm, J., Wieslander, E., Nishimura, T., & Marko-Varga, G. (2015). A protein deep sequencing evaluation of metastatic melanoma tissues. *PLoS ONE*, *10*(4), Article e0123661. https://doi.org/10.1371/journal.pone.0123661

Total number of authors:
18

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 26. Feb. 2026

PLOS | ONE

RESEARCH ARTICLE

A Protein Deep Sequencing Evaluation of Metastatic Melanoma Tissues

Charlotte Welinder[1,2], Krzysztof Pawłowski[3], Yutaka Sugihara[1], Maria Yakovleva[4,5], Göran Jönsson[1], Christian Ingvar[6], Lotta Lundgren[1,7], Bo Baldetorp[1], Håkan Olsson[1,7,8], Melinda Rezeli[5], Bo Jansson[1], Thomas Laurell[2,5], Thomas Fehniger[1,2], Balazs Döme[4,9], Johan Malm[10], Elisabet Wieslander[1], Toshihide Nishimura[1,2,11], György Marko-Varga[2,5,11]*

1 Oncology and Pathology, Dept. of Clinical Sciences, Lund University, Lund, Sweden, **2** Centre of Excellence in Biological and Medical Mass Spectrometry "CEBMMS", Biomedical Centre D13, Lund University, Lund, Sweden, **3** Warsaw University of Life Sciences, Warsaw, Poland, **4** National Korányi Institute of Pulmonology, Budapest, Hungary, **5** Clinical Protein Science & Imaging, Biomedical Centre, Dept. of Biomedical Engineering, Lund University, Lund, Sweden, **6** Surgery, Dept. of Clinical Sciences, Lund University, Skåne University Hospital, Lund, Sweden, **7** Skåne University Hospital, Lund, Sweden, **8** Cancer Epidemiology, Dept. of Clinical Sciences, Lund University, Lund, Sweden, **9** Department of Thoracic Surgery, Medical University of Vienna, Vienna, Austria, **10** Section for Clinical Chemistry, Dept. of Laboratory Medicine, Lund University, Skåne University Hospital in Malmö, Malmö, Sweden, **11** First Dept. of Surgery, Tokyo Medical University, Tokyo, Japan

* Gyorgy.Marko-Varga@bme.lth.se

Abstract

Malignant melanoma has the highest increase of incidence of malignancies in the western world. In early stages, front line therapy is surgical excision of the primary tumor. Metastatic disease has very limited possibilities for cure. Recently, several protein kinase inhibitors and immune modifiers have shown promising clinical results but drug resistance in metastasized melanoma remains a major problem. The need for routine clinical biomarkers to follow disease progression and treatment efficacy is high. The aim of the present study was to build a protein sequence database in metastatic melanoma, searching for novel, relevant biomarkers. Ten lymph node metastases (South-Swedish Malignant Melanoma Biobank) were subjected to global protein expression analysis using two proteomics approaches (with/without orthogonal fractionation). Fractionation produced higher numbers of protein identifications (4284). Combining both methods, 5326 unique proteins were identified (2641 proteins overlapping). Deep mining proteomics may contribute to the discovery of novel biomarkers for metastatic melanoma, for example dividing the samples into two metastatic melanoma "genomic subtypes", ("pigmentation" and "high immune") revealed several proteins showing differential levels of expression. In conclusion, the present study provides an initial version of a metastatic melanoma protein sequence database producing a total of more than 5000 unique protein identifications. The raw data have been deposited to the ProteomeXchange with identifiers PXD001724 and PXD001725.

OPEN ACCESS

Citation: Welinder C, Pawłowski K, Sugihara Y, Yakovleva M, Jönsson G, Ingvar C, et al. (2015) A Protein Deep Sequencing Evaluation of Metastatic Melanoma Tissues. PLoS ONE 10(4): e0123661. doi:10.1371/journal.pone.0123661

Academic Editor: Yiqun G. Shellman, University of Colorado, School of Medicine, UNITED STATES

Received: October 29, 2014

Accepted: February 21, 2015

Published: April 13, 2015

Copyright: © 2015 Welinder et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The full proteomics dataset will be in the PRIDE database at (http://www.ebi.ac.uk/pride/archive). The data have been deposited to the ProteomeXchange with identifiers PXD001724 and PXD001725.

Funding: This work was supported by grants from Mrs. Berta Kamprad Foundation, Swedish Academy of Pharmaceutical Sciences, Swedish Research Council, the Swedish Foundation for Strategic Research, Vinnova, Ingabritt & Arne Lundbergs forskningsstiftelse, the Crafoord Foundation and European Research Council Advanced Grant ERC-2011-294576. The funders had no role in study

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Significance

The present study delivers an initial volume of a high-quality mass spectrometry-derived protein sequence database for metastatic melanoma. Complementary benefits of two alternative proteomics approaches are compared. The need for clinically proven biomarkers for application in the diagnosis, staging, and monitoring of treatment of melanoma is critical, which justifies deep mining proteomic analysis of metastatic tissues. Further development and validation of identified proteins aim to deliver markers of clinical utility.

Introduction

Malignant Melanoma (MM) is defined as cancer of the melanocyte, the cell that produces pigment (melanin) in the skin. Malignant melanoma in a disseminated state has a poor prognosis. According to World Health Organization, there were about 55000 deaths from melanoma occurring globally in 2012 (http://www.iarc.fr/en/publications/books/wcr/index.php).

In Sweden, MM is the sixth most common form of cancer and the one most rapidly increasing with an annual increase around 5% (http://www.Socialstyrelsen.Se/register/halsodataregister/cancerregistret/inenglish).

The majority of early cases of cutaneous melanoma are cured surgically; however some primary tumors will relapse and become metastatic. The American Joint Committee on Cancer staging of the tumors is based on tumor thickness, mitotic rate and ulceration as well as on regional and distant spread [1–3]. Metastatic melanoma has been inherently difficult to treat with a very low 5 year survival (<15%) [4]. Newly developed drugs allowing targeted therapy such as protein kinase inhibitors or drugs modulating the immune response provide more promise [5–10]. However, even with these newer treatments drug resistance may also develop [11]. With treatment options requiring individualized therapies, there is a great demand for validated biomarkers that can support both the primary diagnosis, understanding the progression of disease and response to the treatment of metastatic disease.

Several biochemical markers are already clinically used to monitor progression and relapse of melanoma, such as S100B, MART1 and PMEL [12] and S100A13 [13]. A large number of other markers have been investigated in melanoma, recently reviewed by Levine and Fisher [14]. However, their relevance to melanoma progression, clinical outcome and the selection of best treatment strategies still needs to be established. The search for novel, more accurate markers continues. Both genetic and genomic approaches have been employed in studying MM and specific gene profiles have been correlated to prognosis and survival [15–19]. Gene expression profiles can thus be useful, but identifying and understanding of the functional role of protein in disease development is necessary, for it is proteins rather than genes that are the targets of therapy. Directed protein identification strategies are themselves complex due to the heterogeneities in protein structural components. Proteins can be subjected to a wide variety, even as many as 200, of chemical modifications after translation [20]. These post-translational modifications, often critical to the protein function, may often be altered in disease. Obviously, such modified proteins are also highly important as drug targets.

Various immunological techniques like immunohistochemistry, ELISA, etc. have contributed to build protein expression knowledge and global protein analysis. Technologies such as 2D-PAGE and/or mass spectrometry (MS) have made the identification of an even higher number of proteins possible. A recent publication reports 1528 proteins identified from formalin fixed archival tissue samples of benign nevi, primary melanomas and metastatic melanomas, where 171 proteins differentiated significantly between the three groups [21].

In the present study frozen samples from well-characterized MM tissue, in the South-Swedish Malignant Melanoma Biobank, were utilized [22, 23]. Samples were processed and

subjected to chromatographic separation followed by mass spectrometric analysis (LC-MS/MS). The samples were fractionated using strong cation exchange ion chromatography (SCX) in order to increase the number of detected proteins [24]. Protein expression in unfractionated versus fractionated samples was then compared. Data was also analyzed by referring the samples to two melanoma subtype groups based on previously published genomic profiling, where lymph node metastases were grouped by their specific genomic profiles [15, 25]. In that study each "sub-type group" was characterized by differential expression of specific genes; immune-response genes for the "high-immune" group and genes involved in melanin synthesis and melanocyte differentiation for the "pigmentation" group. Also, the sub-type classification was earlier shown to have a prognostic role related to the clinical outcome of the patients [15, 25]

Half of the samples in the current study were obtained from a set defined as the "pigmentation" sub-type and half from samples defined as the "high-immune" subtype [26]. Various statistical techniques were applied to assess biological significance of the protein repertoire detected in the malignant melanoma samples. As detailed in the Methods and Results sections, DAVID analyses related various protein lists to biological processes and pathways, ANOVA analysis was used to look for protein features that might explain why some proteins were, surprisingly, detected only in the unfractionated proteomics approach, and Mann-Whitney U-test was applied to assess differences in protein detection frequencies between sample subsets.

Based on detected protein signatures and molecular classification using bioinformatics analysis, emerging biological relevance could be assigned to several marker proteins.

Materials and Methods

Clinical Samples

This study was approved by the Regional Ethical Review Board, Southern Sweden; approval number: DNR 191/2007 and 101/2013. All patients within the study provided a written informed consent. The tumor tissues used were lymph nodes metastases from 10 MM patients undergoing surgery at Lund University Hospital, Sweden. The fresh specimens were divided into two parts. One portion of the metastasis was fixed in formalin, paraffin embedded and sectioned for histopathological confirmation and the other part was snap frozen within minutes after removal and stored at -80°C in the Biobank. The frozen specimens were used as described below for both protein expression analysis and histological comparisons.

Histology of Tumors

Frozen tissue samples were sectioned on a cryostat into 6 μm thick sections, placed upon glass slides, dried at 37°C for 30 min and fixed with 100% methanol for 5 min. The sections were stained with HE [27, 28], where protein-rich cytoplasm stains dark pink while cytoplasm that is actively synthesizing protein stains rich purple and the nucleus stains blue.

Sample Preparation

Frozen tissue samples from each tumor were sliced into 10 x 10 μm thick sections using a cryotome. The sections were lysed in 200 μl of 50 mM ammonium bicarbonate and 6 M urea and sonicated with a Branson Sonifier 250 (output 4, 10% duty cycle) for 2 minutes followed by centrifugation at 10 000 g for 5 minutes. The amount of protein in the samples was determined by the BCA method (Pierce, Rockford, IL). A fixed amount (150 μg) of protein were reduced with 10 mM DTT (1 h at 37°C) and alkylated using 40 mM iodoacetamide (30 min, kept dark at room temperature) followed by buffer exchange with 50 mM ammonium bicarbonate buffer (pH 7.6). The samples were then digested with sequencing grade trypsin (Promega, Madison,

WI) in a ratio 1:120 w/w (trypsin:protein) overnight at 37°C. The digestion was stopped by adding 30 μL 1% formic acid. The samples were dried using a centrifugal evaporator and resuspended in 150 μL 1% formic acid and centrifuged for 5 min at 10 000 g. The supernatants were stored at -80°C until further use.

Sample Fractionation

Strong cation exchange chromatography was performed using Microspin columns (MA SEM HIL-SCX, 10–100μg capacity, The Nest group Inc., South Borough) according to the manufacturer's instruction. The amount of peptides loaded to the columns corresponded to 50μg protein. The peptides were eluted by stepwise salt gradient using 0, 20, 40, 60, 100 and 500 mM KCl in 10 mM potassium phosphate, 20% acetonitrile, pH 2.8. The eluted fractions were dried using a centrifugal evaporator and resuspended in 0.1% TFA in water. Before LC-MS/MS analysis the fractions were desalted using Ultra Microspin column Silica C18 (SUM SS18V, 3–30 μg capacity, The Nest group Inc., South Borough) according to the manufacturer's instruction. After elution with 50% ACN, 0.1% TFA in water, the fractions were dried using a centrifugal evaporator and each fraction was resuspended in 15 μL 1% formic acid.

LC-MS/MS Analysis of the Tumor Lysate Digest

The samples were first loaded onto a trapping column (150 mm x 20 μm, Thermo Scientific, San José, CA, USA). The samples were then separated using a column (50 cm x 75 μm, C18, 2 μm and 100Å, PN 164540, Thermo Scientific, San José, CA, USA) with a flow rate of 250 nL/min. A nonlinear gradient was used, using solvent A (0.1% formic acid) and solvent B (0.1% formic acid in ACN). The gradient started with 5% B and 20% B at 120 min, followed by 40% B at 180 min, increased to 90% at 185 min, which was maintained for 5 min.

Both unfractionated and fractionated tumor lysate digests were analyzed using a Q-Exactive, using top10 data-dependent approach. Full MS scans were acquired in the Orbitrap mass analyzer over m/z 350–1800 range with resolution 70,000 (at m/z200). The target value was 3.00E+06. The ten most intense peaks with charge state$\geq$2 were fragmented in the HCD collision cell with normalized collision energy of 30%, and tandem mass spectra were acquired in the Orbitrap mass analyzer with resolution 17,500 at m/z 200. The target value was 1.00E+05. The ion selection threshold was 3.30E+05 counts, and the maximum allowed ion accumulation times were 100 ms for full MS scans and 150 ms for tandem mass spectra. For all the experiments, dynamic exclusion was set to 90 s. The total protein amount of un-fractionated tumor lysate digest injected to the MS/MS platform was 1.25 μg and for the fractionated samples the protein amount was estimated to be 1 μg (50 μg of peptides were fractionated into 6 fractions, approximately 8.3 μg in each fraction). The fractions were dried and resuspended in 15 μL 1% formic acid which gave a concentration of 0.6 μg/μL and 2 μL was injected to the LC-MS/MS).

Data analysis

Raw data were analyzed with Proteome Discoverer v 1.3 (Thermo Scientific, San José, CA, USA) using both Sequest and Mascot search engines. Uniprot Human (release 06/03/2013) database was used. Precursor and fragment mass tolerances were set to 1 Da, where the maximum of missed cleavage sites was three and 1% false discovery rate was used. At least two unique peptides were necessary for protein identification, and at the parameter Peptide Confidence in the Proteome Discoverer was required to be at least "medium". The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium [29] via the PRIDE partner repository with the dataset identifiers PXD001724 and PDX001725.

Bioinformatics

Protein lists were analyzed using the DAVID tool set [30]. Enrichment of the lists in particular functional annotations was assessed in DAVID by Fisher's exact test with Benjamini-Hochberg correction for multiple tests. Generally, annotations were considered significant if the Fisher's exact test p-value after Benjamini-Hochberg correction was below 0.001. In this application of the Fisher's exact test, sample sizes were relatively large, specifically; the protein lists analyzed here numbered hundreds of proteins. In a single case, the protein list analyzed by the Fisher's test was small. This was an inherent result of comparison of occurrence frequencies between the melanoma subsets. Gene Ontology annotations, SwissProt keywords, and KEGG pathways were used as annotation terms for the DAVID enrichment analysis. For lists of proteins with significantly changed expression, the background protein sets consisted of all proteins detected by applying the given proteomics approach (fractionated or unfractionated). For the list of all proteins detected, the background protein set consisted of all human proteins. Proportional area Venn diagrams were created using the BioVenn program [31]. Other visualizations, list manipulations and charts were conducted in Spotfire (Somerville, MA, USA, www.tibco.com).

For statistical analysis of significant differences in protein occurrences between the "pigmentation" and "high immune" patient sample sets, Mann-Whitney U test was used as implemented in the in XLSTAT program. For the purpose of this test, for each protein, each sample was assigned an integer value between 0 and 3, corresponding to the number of times the protein was detected in replicates of the sample. Thus, for the unfractionated approach, for each protein the Mann-Whitney test compared two populations of five elements each. Differences in protein occurrences between the patient sample sets were considered significant if the Mann-Whitney U test p-value was below 0.05. In this application of the Mann-Whitney U test, sample sizes were relatively small, namely five patients in each group, as available in this pilot study. For the fractionated approach, sample numbers were too low to attempt a statistical evaluation of differences. This conservative approach to comparison of detection rates was selected to look for strong trends in differences in detection rates between the two sample sets. The ANOVA analysis was performed using the STATISTICA package (StatSoft) with default parameters, and the Least Significant Difference test. Differences in protein properties between sets of proteins detected using the different approaches (see Results) were considered significant if the ANOVA test p-value was below 0.001. In this application of the ANOVA test, sample sizes were relatively large; the groups that were compared numbered hundreds of proteins.

Results

The current investigation was performed utilizing a study workflow, as illustrated in Fig 1. The patient characteristics are summarized in Table 1. In the unfractionated approach, ten malignant melanoma metastasis samples were studied, each with three technical replicates. In the fractionated approach, only four out of those ten samples were studied, again each with three technical replicates.

Histological Characterization

Characterization of the tissue samples by histology of the lymph node tumors showed almost complete replacement of normal follicular architecture by metastatic melanoma cells, as shown in Fig 2 (2a and b represent the "pigmentation" subtype and 2c and d represent the "high-immune" subtype).

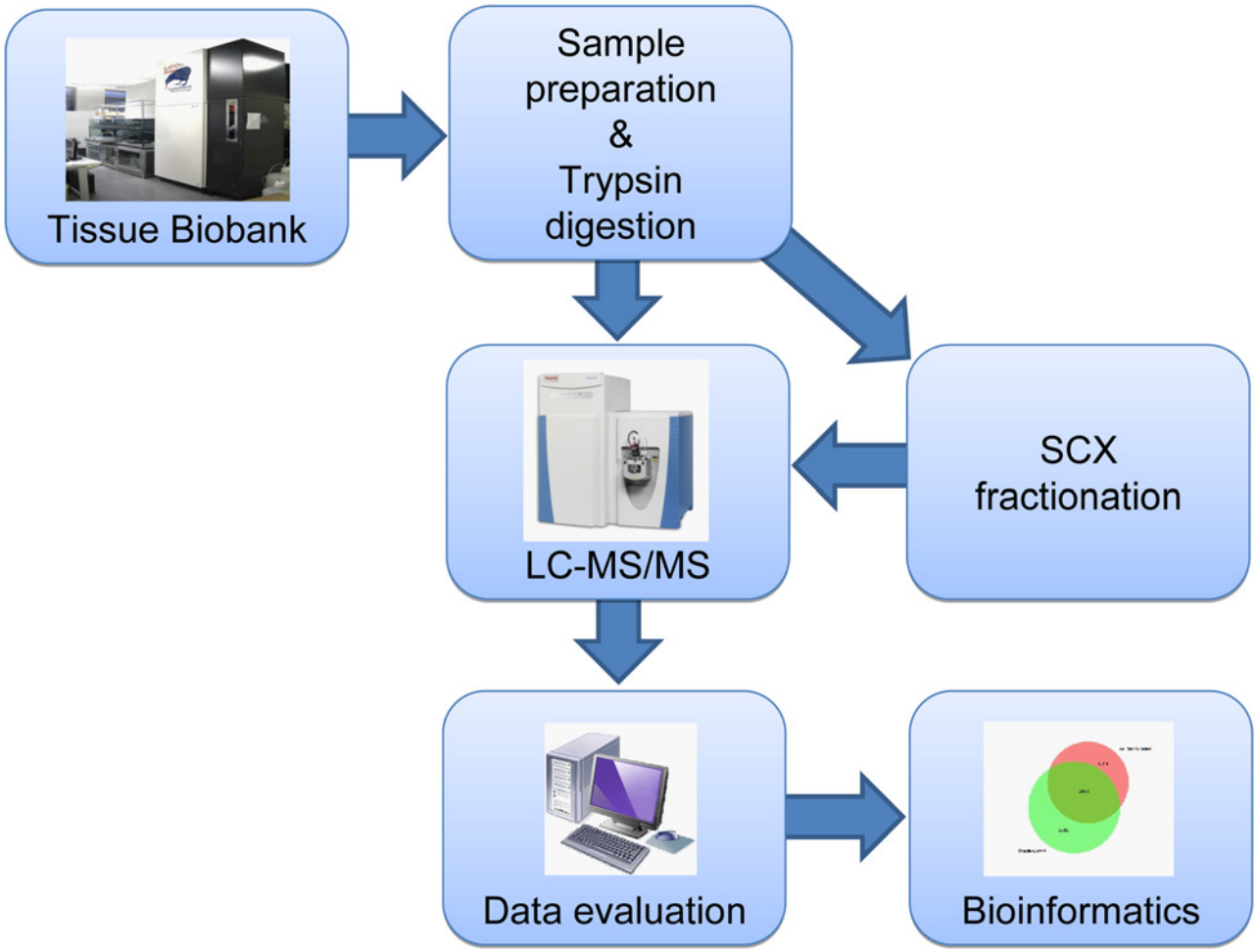

Fig 1. Illustration of study work flow.

doi:10.1371/journal.pone.0123661.g001

Table 1. Patient and tumor characteristics. Breslow tumor thickness and Clarks level of invasion refer to the primary melanoma.

Tumor	Gender	Age at primary melanoma	Years from primary diagnosis to diagnosed metastasis	Breslow class (T class)	Clark	Status
MM35	Male	54	1	3	4	Alive
MM98	Male	73	2	4	4	Dead
MM504	Male	NA				Dead
MM687	Male	72	2	1	2	Dead
MM787	Male	78	78	2	4	Dead
MM812	Male	NA				Alive
MM813	Female	54	0	2	3	Alive
MM825	Female	64	2	2	4	Alive
MM829	Male	49	6	1	2	Alive
MM835	Female	32	4	3	3	Alive

*NA—not available, primary tumor not diagnosed

doi:10.1371/journal.pone.0123661.t001

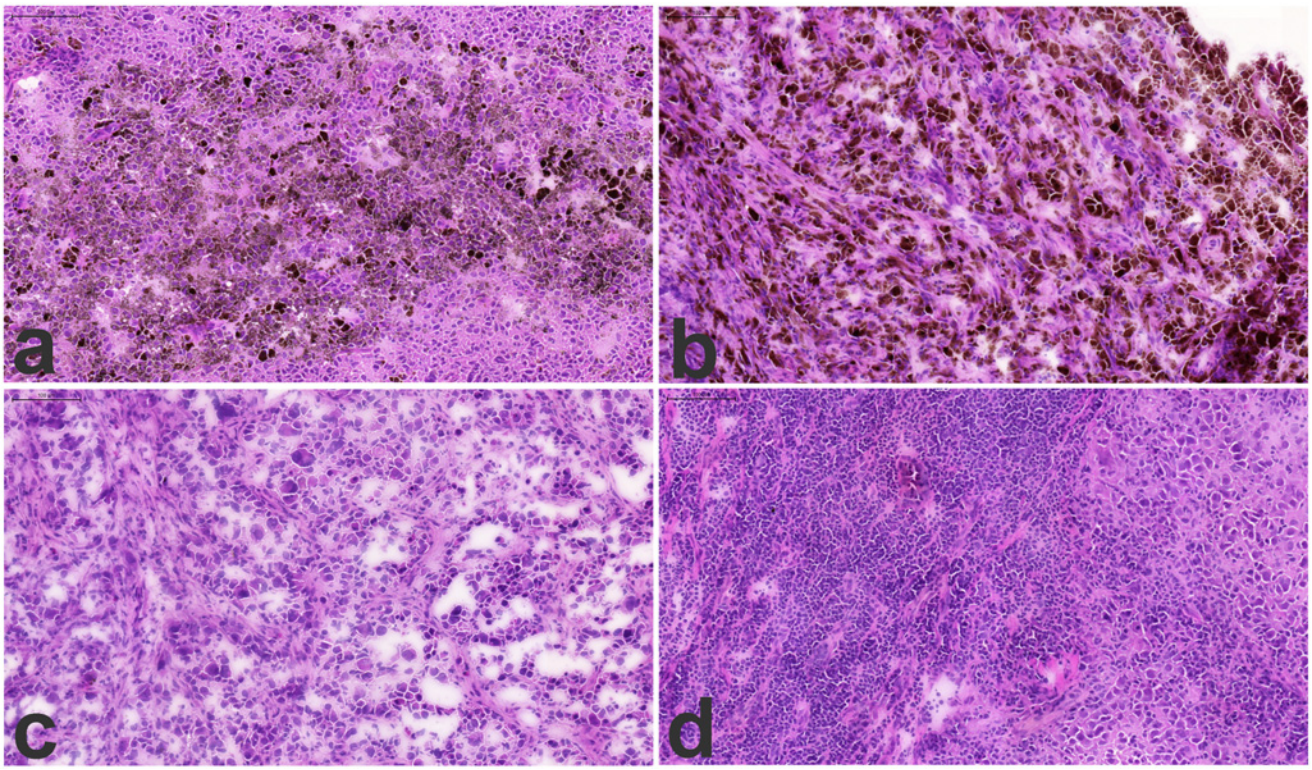

Fig 2. Histology images of lymph node metastasis, 2a and b the "pigmentation" subtype and 2c and d the "high-immune" subtype. Frozen tumor samples were cryosectioned and stained with HE. The nuclei stains blue and protein-rich cytoplasma stains dark pink while cytoplasma that is actively synthesizing proteins stains rich purple. The brown pigment seen scattered within clusters in 2a and b corresponds to focal hyper-expression of melanin by groups of melanoma cells.

doi:10.1371/journal.pone.0123661.g002

Deep mining—Protein Sequencing

Proteins detected using the two proteomics approaches. Since, rather unexpectedly, some proteins were only identified in the samples run without the SCX step (unfractionated) we decided to include a protein expression comparison that relates the peptide and protein annotations between unfractionated, and orthogonal- fractionated separation. The rationale here was to look for any factors that might contribute to the unexpected gain in protein identifications by combining the two proteomics approaches. Thus, the output of protein identifications can be related to the cycle time needed for a given work flow procedure. The most common bottom-up LC-MS platform used today in clinical protein expression studies is C18-hydrophobic separation. This methodology provides high resolution liquid phase nano-capillary separation of peptide sequences, where typically 1–4 hour organic solvent gradients are being applied. In order to increase the resolving power, an ion-exchange separation step can be introduced either on-line, or off-line as done in the present study [32–34].

The electrostatic charge properties will provide additional beneficial factors in the overall work flow for the separation efficiency in a proteomic study [35, 36].

The resulting outcome was evaluated and analyzed with pre-set criteria for protein sequence annotation confirmation (see Materials and Methods).

The total number of annotated proteins detected was 5326 using the conservative statistical criteria used (see Materials and Methods), combining both the fractionated and unfractionated approaches, and using the more than 17000 unique peptides identified. Among these, 4284

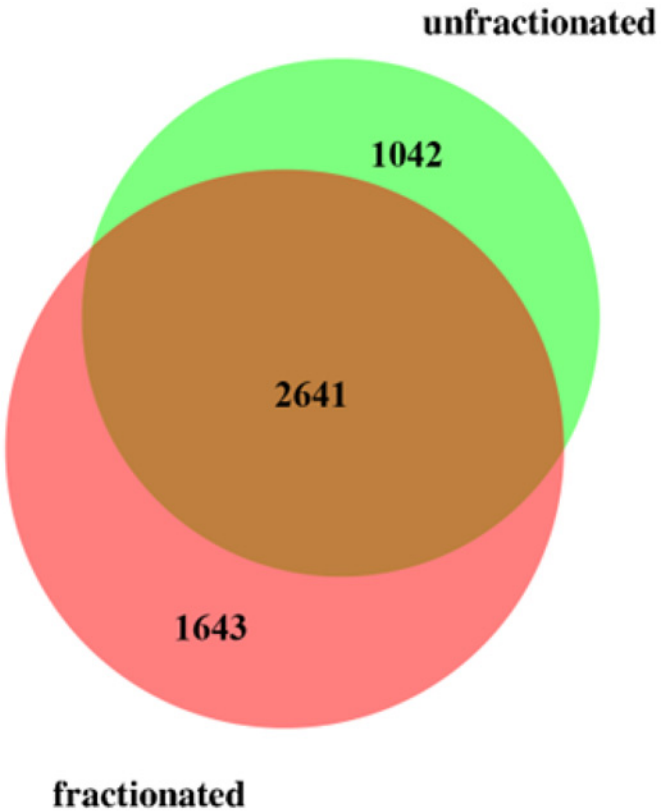

Fig 3. Venn diagram showing overlap between proteins detected applying the fractionated and the unfractionated approaches.

doi:10.1371/journal.pone.0123661.g003

proteins were annotated using more than 13600 peptides by applying the fractionated approach, while without fractionation 3683 proteins were annotated using more than 14700 peptides. Even though it is well known that pre-fractionation is favorable for analysis in order to maximize proteome coverage, our study showed that more than 1000 detected protein were only found using the unfractionated approach. The overlap between the two approaches was 2641 proteins, i.e. close to half of the total number of proteins detected (Fig 3 and S1 Table).

A substantial proportion of proteins were detected with few peptides only, e.g. in the unfractionated approach 2360 proteins (61% of all proteins detected) and in the fractionated approach 2375 proteins (58%) were detected with 2 or 3 peptides. Also, a substantial proportion of proteins were detected with relatively low sequence coverage, e.g. 1631 proteins (42%) had coverage below 10% in the unfractionated approach while 2169 (53%) had coverage below 10% in the fractionated approach. A substantial proportion of proteins were detected only in one or two sample, e.g. as many as 1443 proteins (37% of all proteins detected) were detected in 1 or 2 samples in the unfractionated approach. In the fractionated approach 2311 proteins (56%) were detected in 1 or 2 samples only. Proteins detected in almost every sample (all samples or all but one) were relatively few, 585 and 245 in the fractionated and unfractionated approaches, respectively.

Comparison between the fractionated and unfractionated approaches to proteomics. The proteins specific (unique) to each of the two proteomics approaches exhibited rather few special functional characteristics. Here, the DAVID tool set was used to look for biological processes and pathways that may be linked to or characteristic for the two protein sets. Thus, DAVID analyses of overrepresented functional annotations yielded following Gene Ontology

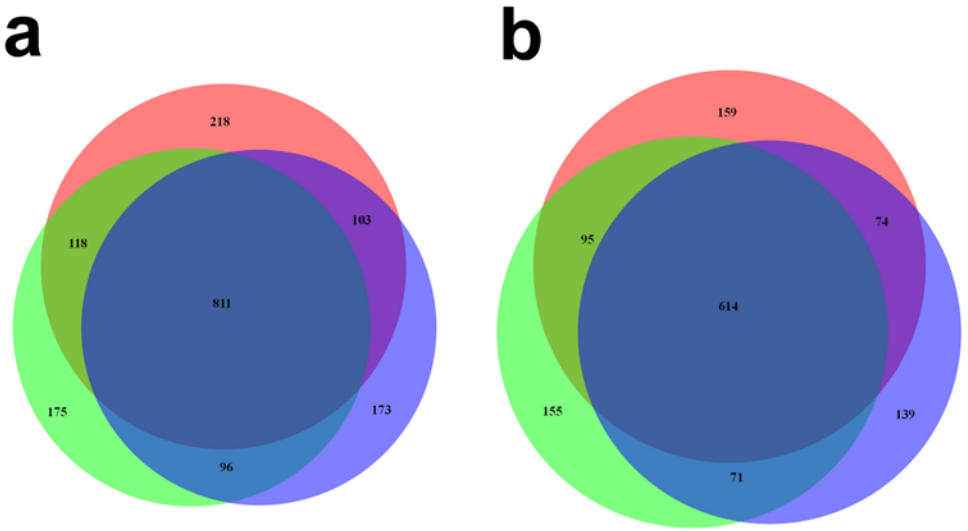

Fig 4. Venn diagram for unfractionated samples. Proteins seen in all three replicates *vs* those seen in a single replicate, for two typical patient samples (a and b).

doi:10.1371/journal.pone.0123661.g004

(GO) annotations: MHC complex, mitochondrial outer membrane and DNA binding for the unfractionated approach-unique protein set. For the fractionated approach-unique proteins, Pleckstrin homology domains, and Gene Ontology (GO) terms such as: GTPase regulator activity, zinc-finger, protein kinase activity were significantly over-represented.

Another DAVID analysis has been performed to look for biological processes and pathways that may be linked to or characteristic for the full set of proteins detected in the current experiment (combining the unfractionated and fractionated approaches). Thus, the full set when compared with human proteome in general was characterized by overrepresented functional annotations that could be expected for melanoma samples. Specifically, GO: "melanosome", GO:"pigment granule", GO:"cytoplasmic vesicle" annotations were found with p-values below 1E-15, and may correspond to melanocyte-specific protein groups. Then, the following terms may correspond to regulation of cancer-specific processes: KEGG: "Glycolysis", p-value approx. 1E-8, GO: Mitochondrion, p-value 1E-37, GO: RNA binding and RNA splicing, p-values below 1E-25.

Interestingly, separate analysis of technical replicates provided a substantial gain in number of proteins detected. As seen in Venn diagrams for two typical unfractionated patient samples (Fig 4a and 4b), approximately half of proteins detected in a sample were observed in all the three replicates while approximately a third of the proteins are observed in a single replicate only.

There is a major impact on n-numbers of identified peptides and resulting proteins, depending on the work flow chosen in melanoma patient tissues, in addition to the uniqueness of the protein sequences. The fractionated workflow provides an increase in number of detected proteins in relation to the unfractionated approach. One explanation of why we also find different proteins in unfractionated and fractionated LC-MS workflows is that hydrophobic separation mechanisms are being utilized for separation in unfractionated samples, while we have an electrostatic mechanism prior to the hydrophobic separation in the fractionated workflow. This means that the fractions being introduced into the 2nd dimension of separation, already have been fractionated based on the respective peptide polarity, given at the specific pH, and the increasing salt concentrations. These orthogonal properties will be different in combination

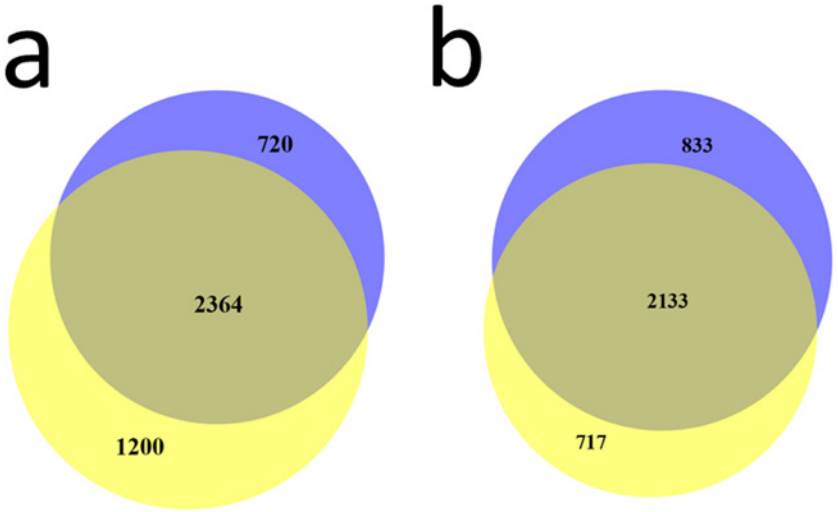

Fig 5. Venn diagrams for a) fractionated approach and b) unfractionated approach. Proteins seen in the "pigmentation" (blue color) sample subset *vs* those seen in the "high-immune" (yellow color) sample subset.

doi:10.1371/journal.pone.0123661.g005

(charge and hydrophobicity), in relation to not using fractionation, where only the apolar properties will impact the separation.

When physical properties of proteins detected a) by only using the fractionated approach, b) by only using the unfractionated approach and c) by using both approaches, were compared by the ANOVA technique, interesting patterns emerged. All three conditions were related to significantly different average molecular weight (79 kDa, 56 kDa and 68 kDa, respectively, p-value$<$0.004 for any of the three contrasts). As expected, similar differences were observed for the average lengths of the proteins detected. Thus, the fractionated approach-unique proteins were significantly larger, on average, than the others.

Among detection parameters, the average sequence coverage for proteins detected: a) using only the fractionated approach, b) using only the unfractionated approach and c) using both approaches, was also significantly different between the conditions (8.9%, 12.2% and 15.6%, respectively, p-value $<$0.0001 for any of the three contrasts).

Other detection parameters exhibited different variation patterns. Number of unique peptides per protein or number of peptide-spectrum matches (PSMs was significantly higher for proteins detected using both approaches (ANOVA p-value $<$0.0001), with borderline or no difference between proteins detected using solely one of the approaches.

This fact may suggest that proteins detected exclusively using a single approach (either without fractionation or with fractionation) are lower in abundance, although rigorous quantification was not performed here.

Comparison between "pigmentation" and "high-immune" subtypes of melanoma. Both proteomics approaches allowed comparison between two subtypes of melanoma identified previously in a genomics study, the "pigmentation" and the "high immune" subsets [15, 37]. Substantial differences could be observed in protein presence (detection) between the two subsets, with only about 50% of the proteins detected in both subsets (Fig 5a and 5b shows these differences for fractionated and unfractionated approaches, respectively).

However, more biologically relevant would be comparison of protein detection frequencies between the subsets. Due to the small numbers of patient samples in this pilot study, protein presence (detection) frequency could only be compared between the "pigmentation" and the "high-immune" melanomas for the unfractionated approach. This was done using Mann-Whitney U

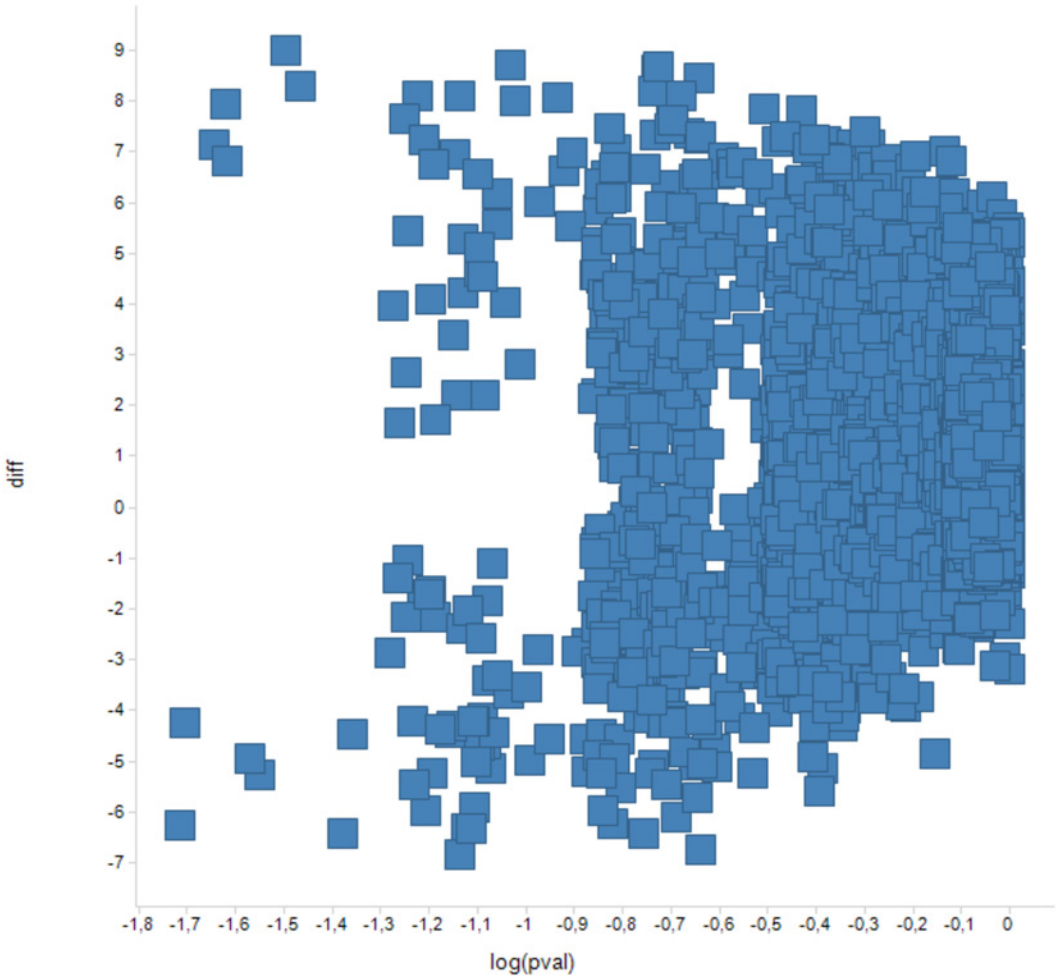

Fig 6. Scatter graph—Y axis: protein detection counts in the "pigmentation" subset minus protein detection counts in the "high-immune" subset. X axis: Mann Whitney test p-values for the differences.

doi:10.1371/journal.pone.0123661.g006

test for comparison of 5 versus 5 samples. For every sample, the detection frequency could assume values of 0, 1/3, 2/3 and 1. With the threshold for the p-value of the Mann-Whitney U test set at 0.05, there were 11 proteins differing in detection frequencies between the "pigmentation" and the "high-immune" subsets in the unfractionated approach (Fig 6 and Table 2).

The rigorous Mann Whitney test yielded a small number of proteins differing in detection frequency between the "high-immune" and "pigmentation" sets. Hence, analysis of overrepresented functional terms had low statistical power. Here, DAVID analysis has been performed to look for biological processes and pathways that may be linked to or characteristic for the proteins differing in detection frequency between the "high-immune" and "pigmentation" sample sets. The only term reaching significance in the DAVID analysis even without multiple test correction was "cellular stress response". This suggests that the differences between "high immune" and "pigmentation" sample sets may be related to these processes. One protein recently described to be expressed in melanoma is alpha-synuclein [38, 39]. Dysregulation of alpha-synuclein is observed in Parkinson's disease and the protein is implicated in the pathway of dopamine as well as melanin synthesis. In the present study, the differential expression of alpha-synuclein between the "pigmentation" and "high-immune" melanomas did not reach statistical

Table 2. Proteins differing in detection frequency between the "high-immune" and the "pigmentation" melanoma sample subsets. Detection counts include replicates.

Protein ID	Mann Whitney—p-value	Count in "pigmentation" subset	Count in "high-immune" subset	Description
P47756	0,019	8	0	F-actin-capping protein subunit beta (CapZ beta); CAPZB_HUMAN
Q9UJU6	0,019	14	5	Drebrin-like protein (Cervical SH3P7); DBNL_HUMAN
B7Z1I0	0,023	0	4	Integrin-linked protein kinase; ILK_HUMAN
E5RIW3	0,024	0	4	Tubulin-specific chaperone A; E5RIW3_HUMAN
P04264	0,024	0	7	Keratin, type II cytoskeletal 1 (67 kDa cytokeratin); K2C1_HUMAN
Q99733	0,025	15	7	Nucleosome assembly protein 1-like 4 (Nucleosome assembly protein 2); NP1L4_HUMAN
Q9NZM1	0,025	0	6	Myoferlin (Fer-1-like protein 3); MYOF_HUMAN
O95881	0,042	12	4	Thioredoxin domain-containing protein 12 (Endoplasmic reticulum resident protein 18); TXD12_HUMAN
Q8IZP2	0,042	14	5	Putative protein FAM10A4 (Suppression of tumorigenicity 13 pseudogene 4); ST134_HUMAN
Q9UH99	0,043	1	8	SUN domain-containing protein 2 (Protein unc-84 homolog B); SUN2_HUMAN
P31930	0,049	5	9	Cytochrome b-c1 complex subunit 1, mitochondrial; QCR1_HUMAN

doi:10.1371/journal.pone.0123661.t002

significance, however a trend was clear. Thus, alpha-synuclein was observed in 4 "pigmentation" tissue samples and 3 "high-immune" tissue samples using the unfractionated approach, and 2 "pigmentation" samples and 1 "high-immune" using the fractionated approach.

This trend is interesting remembering relatively poor prognosis of patients with pigmentation subtype of melanoma [15]. In our earlier study we found a strong correlation between protein and mRNA expression levels of alpha-synuclein in metastatic tumor lysate from MM [39].

A recent study by Byrum et al. utilized Formalin-Fixed, Paraffin—Embedded tissues samples and compared protein abundances between benign, primary and metastatic melanomas [21]. These melanoma categories do not translate directly into "pigmentation" and "high-immune" melanoma subsets. However, out of 171 proteins that they reported to be differentially expressed between the conditions studied, 147 proteins were detected in our study. The vast majority of the significantly differentiating proteins reported by Byrum et al. were detected in their metastatic melanoma samples, and approximately half of these were actually upregulated in metastatic melanoma vs primary or malignant samples.

Another recent study by Mactier and co-workers elucidated protein signatures related to survival outcome in melanoma patients with stage IIIc lymph node metastases [40]. Out of 84 proteins they reported to be differentially abundant between the prognostic groups, 76 proteins were detected in our study. Thus, for both studies, the majority of proteins differentiating between the respective clinically relevant conditions are detected in our study.

Among a number of melanoma markers being considered in the literature (https://www.aacc.org/SiteCollectionDocuments/NACB/LMPG/tumor/chp3L_melanoma.pdf Table 2b Tissue tumor markers), five proteins exhibited substantial expression in our experiment (presence in at least 40% of the samples in either approach): S100B, ICAM, NDKA (NM23), MUC18 (MCAM), PMEL (GP100). Further five proteins were seen in fewer samples (Table 3). Interestingly, MMP9 was seen in 50% of the fractionated samples while it was never seen in the unfractionated ones. Also, both AP-2 and MITF were only detected in a single fractionated sample.

Table 3. Present rates for selected melanoma markers proposed in the literature. Percentages of samples (including technical replicates) in which a protein was detected.

Protein	Unfractionated approach	Fractionated approach
S100B	70%	100%
ICAM	43,3%	100%
NDKA (NM23)	46,7%	83.3%
MUC18 (MCAM)	40%	100%
PMEL (GP100)	46,7%	58.3%
MMP-9	0%	50%
CD44	3,3%	25%
Tyrosinase	3,3%	16.7%
AP-2	0%	8.3%
MITF	0%	8.3%

doi:10.1371/journal.pone.0123661.t003

Discussion

In this study, applying conservative criteria, more than 5000 proteins were identified in metastatic melanoma samples at a 1% false discovery rate. Two approaches of sample preparation using either unfractionated whole tissue lysates or orthogonal fractionation were applied to test the conditions required for optimal MS proteomics annotations. Both the approaches were found to be complementary and each provided identifications of approximately half of the detected proteins. Deep mining the protein sequencing by nano-LC MS, has for over a decade been a challenging task, and still is.

Even in comparison, while studying the individual MS spectra derived from replicates of the same sample and analyzed on the same instrument platform, one has to keep in mind that the current nano-LC MS platforms provide an output that is to some extent variable. The probable factors contributing to these variances in detection precision are already known and include 1) the integral complexity of the protein composition within whole tissue extracts, 2) the power of the nano-chromatographic separation to provide single peptide separation, and rather the separation of multiple peptides that are interfaced into the same nano-ESI interface separation time space, 3) since the mass spectrometer acts as both a separator, as well as a detector, it is unable to generate MS and MS/MS sequence data on all individual peptides present in the sample [41].

The full set of proteins detected herein in metastatic melanoma samples was enriched in functional annotations well expected for melanocyte-containing tissues and in annotations related to regulation of cancer-specific processes. Interestingly, not only alternative proteomics approaches, but also technical replicates provided a very substantial gain in number of proteins detected.

The two protein sets specific (unique) to the two alternative proteomics approaches were enriched in some particular protein groups, e.g. MHC complex and mitochondrial outer membrane proteins (unfractionated approach) or zinc-finger proteins and protein kinases (fractionated approach).

The unique proteins identified by the fractionated approach were significantly larger in size, on the average, than the others. Reasons for this are not clear to us. Probably as a consequence, the average sequence coverage for proteins detected using only the fractionated approach was significantly lower than the coverage for proteins detected using both approaches and those unique to the unfractionated approach. Interestingly, no differences in protein pI were observed between the two approaches. Overall, significantly higher number of unique peptides

per protein and numbers of PSMs observed for proteins detected using both approaches when compared to approach-specific proteins may suggest that proteins detected using only a single approach (be it unfractionated or fractionated) are of lower abundance.

A comparison of protein detection frequencies was conducted between the biologically relevant "high-immune" and "pigmentation" sample sets. The few proteins significantly differing in detection frequencies were enriched in cellular stress response annotations.

In addition, our study indicates that a number of possible melanoma markers being considered in the literature can be readily detected in a substantial fraction of our metastatic melanoma samples, even if the experiment was not directed at quantifying the particular proteins.

Conclusion

We present here the initial stage of a protein sequence database for metastatic melanoma using deep mining high-quality mass spectrometry data. More than 5000 proteins were identified in metastatic tissues using our approach.

Further development of the metastatic melanoma proteomics database will involve larger sample numbers, collection of additional clinical parameters for the samples, development of validation assays (e.g. SRM) directed at particular protein molecules. Larger sample numbers will allow a deeper insight into the biological mechanisms related to metastatic melanoma development and potential differences between disease subtypes.

Increased knowledge about protein expression and regulation during melanoma disease progression and treatment will support development of validated biomarkers for melanoma patients.

Supporting Information

S1 Table. Protein identified using the two proteomics approaches. Protein identified using Uniprot.
(XLS)

Acknowledgments

We thank Dr Dariusz Gozdowski for helpful discussions and advice in statistical issues. We thank the PRIDE database team for their assistance.

Author Contributions

Conceived and designed the experiments: CW GMV EW KP. Performed the experiments: CW YS MY MR TF. Analyzed the data: KP. Contributed reagents/materials/analysis tools: GJ CI LL BB HO BJ BD JM TN GMV. Wrote the paper: CW KP TL TF EW GMV.

References

1. Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2009; 27(36):6199–206.
2. Nading MA, Balch CM, Sober AJ. Implications of the 2009 American Joint Committee on Cancer Melanoma Staging and Classification on dermatologists and their patients. Seminars in cutaneous medicine and surgery. 2010; 29(3):142–7. doi: 10.1016/j.sder.2010.06.004 PMID: 21051007
3. Breslow A. Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma. Annals of surgery. 1970; 172(5):902–8. PMID: 5477666
4. Tas F. Metastatic behavior in melanoma: timing, pattern, survival, and influencing factors. Journal of oncology. 2012; 2012:647684. doi: 10.1155/2012/647684 PMID: 22792102

5. Sugihara Y, Vegvari A, Welinder C, Jonsson G, Ingvar C, Lundgren L, et al. A new look at drugs targeting malignant melanoma-An application for mass spectrometry imaging. Proteomics. 2014; 14(17–18):1963–70. doi: 10.1002/pmic.201400075 PMID: 25073962
6. Ascierto PA, Kirkwood JM, Grob JJ, Simeone E, Grimaldi AM, Maio M, et al. The role of BRAF V600 mutation in melanoma. Journal of translational medicine. 2012; 10:85. doi: 10.1186/1479-5876-10-85 PMID: 22554099
7. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. The New England journal of medicine. 2011; 364(26):2507–16. doi: 10.1056/NEJMoa1103782 PMID: 21639808
8. Bollag G, Hirth P, Tsai J, Zhang J, Ibrahim PN, Cho H, et al. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. Nature. 2010; 467(7315):596–9. doi: 10.1038/nature09454 PMID: 20823850
9. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. The New England journal of medicine. 2010; 363(8):711–23. doi: 10.1056/NEJMoa1003466 PMID: 20525992
10. Rizos H, Menzies AM, Pupo GM, Carlino MS, Fung C, Hyman J, et al. BRAF inhibitor resistance mechanisms in metastatic melanoma: spectrum and clinical impact. Clinical cancer research: an official journal of the American Association for Cancer Research. 2014; 20(7):1965–77.
11. Sun C, Wang L, Huang S, Heynen GJ, Prahallad A, Robert C, et al. Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. Nature. 2014; 508(7494):118–22. doi: 10.1038/nature13121 PMID: 24670642
12. Sabel MS, Liu Y, Lubman DM. Proteomics in melanoma biomarker discovery: great potential, many obstacles. International journal of proteomics. 2011; 2011:181890. doi: 10.1155/2011/181890 PMID: 22084682
13. Azimi A, Pernemalm M, Frostvik Stolt M, Hansson J, Lehtio J, Egyhazi Brage S, et al. Proteomics analysis of melanoma metastases: association between S100A13 expression and chemotherapy resistance. British journal of cancer. 2014; 110(10):2489–95. doi: 10.1038/bjc.2014.169 PMID: 24722184
14. Levine D, Fisher DE. Current status of diagnostic and prognostic markers in melanoma. Methods in molecular biology (Clifton, NJ). 2014; 1102:177–97. doi: 10.1007/978-1-62703-727-3_11 PMID: 24258980
15. Jonsson G, Busch C, Knappskog S, Geisler J, Miletic H, Ringner M, et al. Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. Clinical cancer research: an official journal of the American Association for Cancer Research. 2010; 16(13):3356–67. doi: 10.1158/1078-0432.CCR-09-2509 PMID: 20460471
16. Hoek KS, Schlegel NC, Eichhoff OM, Widmer DS, Praetorius C, Einarsson SO, et al. Novel MITF targets identified using a two-step DNA microarray strategy. Pigment cell & melanoma research. 2008; 21(6):665–76.
17. Winnepenninckx V, Lazar V, Michiels S, Dessen P, Stas M, Alonso SR, et al. Gene expression profiling of primary cutaneous melanoma and clinical outcome. Journal of the National Cancer Institute. 2006; 98(7):472–82. PMID: 16595783
18. Haqq C, Nosrati M, Sudilovsky D, Crothers J, Khodabakhsh D, Pulliam BL, et al. The gene expression signatures of melanoma progression. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(17):6092–7. PMID: 15833814
19. Widmer DS, Cheng PF, Eichhoff OM, Belloni BC, Zipser MC, Schlegel NC, et al. Systematic classification of melanoma cells by phenotype-specific gene expression mapping. Pigment cell & melanoma research. 2012; 25(3):343–53.
20. Srinivas PR, Srivastava S, Hanash S, Wright GL Jr. Proteomics in early detection of cancer. Clinical chemistry. 2001; 47(10):1901–11. PMID: 11568117
21. Byrum SD, Larson SK, Avaritt NL, Moreland LE, Mackintosh SG, Cheung WL, et al. Quantitative Proteomics Identifies Activation of Hallmark Pathways of Cancer in Patient Melanoma. Journal of proteomics & bioinformatics. 2013; 6(3):43–50.
22. Baldetorp B. PROFILE: Malignant melanoma initiative in southern Sweden. PUBLIC SERVICE REVIEW. 2013(25).
23. Welinder C, Jonsson G, Ingvar C, Lundgren L, Olsson H, Breslin T, et al. Establishing a Southern Swedish Malignant Melanoma OMICS and biobank clinical capability. Clinical and translational medicine. 2013; 2(1):7. doi: 10.1186/2001-1326-2-7 PMID: 23445834
24. Wolters DA, Washburn MP, Yates JR 3rd. An automated multidimensional protein identification technology for shotgun proteomics. Analytical chemistry. 2001; 73(23):5683–90. PMID: 11774908

25. Harbst K, Lauss M, Cirenajwis H, Winter C, Howlin J, Torngren T, et al. Molecular and genetic diversity in the metastatic process of melanoma. The Journal of pathology. 2014; 233(1):39–50. doi: 10.1002/path.4318 PMID: 24399611

26. Cirenajwis H, Ekedahl H, Martin Lauss M, Harbst K, Carneiro A, Enoksson J, et al. Molecular stratification of metastatic melanoma using gene expression profiling—prediction of survival outcome and benefit from molecular targeted therapy. (submitted). 2015.

27. Lillie RD. Histopathologic technic and practical histochemistry. 3rd ed. New York: McGraw-Hill Book Co; 1965.

28. Avwioro G. Histochemical Uses Of Haematoxylin. J Pharmacy and Clinical Sciences. 2011; 1.

29. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. 2014; 32(3):223–6. doi: 10.1038/nbt.2839 PMID: 24727771

30. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 2009; 4(1):44–57. doi: 10.1038/nprot.2008.211 PMID: 19131956

31. Hulsen T, de Vlieg J, Alkema W. BioVenn—a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC genomics. 2008; 9:488. doi: 10.1186/1471-2164-9-488 PMID: 18925949

32. Machtejevas E, John H, Wagner K, Standker L, Marko-Varga G, Forssmann WG, et al. Automated multi-dimensional liquid chromatography: sample preparation and identification of peptides from human blood filtrate. Journal of chromatography B, Analytical technologies in the biomedical and life sciences. 2004; 803(1):121–30. PMID: 15026005

33. Machtejevas E, Marko-Varga G, Lindberg C, Lubda D, Hendriks R, Unger KK. Profiling of endogenous peptides by multidimensional liquid chromatography: On-line automated sample cleanup for biomarker discovery in human urine. Journal of separation science. 2009; 32(13):2223–32. doi: 10.1002/jssc.200900058 PMID: 19569102

34. Wagner K, Miliotis T, Marko-Varga G, Bischoff R, Unger KK. An automated on-line multidimensional HPLC system for protein and peptide mapping with integrated sample preparation. Analytical chemistry. 2002; 74(4):809–20. PMID: 11866061

35. Hartvig RA, van de Weert M, Ostergaard J, Jorgensen L, Jensen H. Protein adsorption at charged surfaces: the role of electrostatic interactions and interfacial charge regulation. Langmuir: the ACS journal of surfaces and colloids. 2011; 27(6):2634–43. doi: 10.1021/la104720n PMID: 21322572

36. Zarei M, Sprenger A, Gretzmeier C, Dengjel J. Combinatorial use of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) and strong cation exchange (SCX) chromatography for in-depth phosphoproteome analysis. Journal of proteome research. 2012; 11(8):4269–76. doi: 10.1021/pr300375d PMID: 22768876

37. Harbst K, Staaf J, Lauss M, Karlsson A, Masback A, Johansson I, et al. Molecular profiling reveals low- and high-grade forms of primary melanoma. Clinical cancer research: an official journal of the American Association for Cancer Research. 2012; 18(15):4026–36. doi: 10.1158/1078-0432.CCR-12-0343 PMID: 22675174

38. Matsuo Y, Kamitani T. Parkinson's disease-related protein, alpha-synuclein, in malignant melanoma. PloS one. 2010; 5(5):e10481. doi: 10.1371/journal.pone.0010481 PMID: 20463956

39. Welinder C, Jonsson GB, Ingvar C, Lundgren L, Baldetorp B, Olsson H, et al. Analysis of Alpha-Synuclein in Malignant Melanoma—Development of a SRM Quantification Assay. PloS one. 2014; 9(10):e110804. doi: 10.1371/journal.pone.0110804 PMID: 25333933

40. Mactier S, Kaufman KL, Wang P, Crossett B, Pupo GM, Kohnke PL, et al. Protein signatures correspond to survival outcomes of AJCC stage III melanoma patients. Pigment cell & melanoma research. 2014.

41. Tabb DL, Vega-Montoto L, Rudnick PA, Variyath AM, Ham AJ, Bunk DM, et al. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. Journal of proteome research. 2010; 9(2):761–76. doi: 10.1021/pr9006365 PMID: 19921851