Tumor Proteomics by Multivariate Analysis on Individual Pathway Data for Characterization of Vulvar Cancer Phenotypes

AnnSofi Sandberg‡, Gunnel Lindell§, Brita Nordström Källström¶, Rui Mamede Branca‡, Kristina Gemzell Danielsson§, Mats Dahlberg‖, Barbro Larson§, Jenny Forshed‡, and Janne Lehtio‡**

Vulvar squamous cell carcinoma (VSCC) is the fourth most common gynecological cancer. Based on etiology VSCC is divided into two subtypes; one related to high-risk human papilloma virus (HPV) and one HPV negative. The two subtypes are proposed to develop via separate intracellular signaling pathways. We investigated a suggested link between HPV infection and relapse risk in VSCC through in-depth protein profiling of 14 VSCC tumor specimens. The tumor proteomes were analyzed by liquid-chromatography tandem mass spectrometry. Relative protein quantification was performed by 8-plex isobaric tags for relative and absolute quantification. Labeled peptides were fractionated by high-resolution isoelectric focusing prior to liquid-chromatography tandem mass spectrometry to reduce sample complexity. In total, 1579 proteins were regarded as accurately quantified and analyzed further. For classification of clinical groups, data analysis was performed by comparing protein level differences between tumors defined by HPV and/or relapse status. Further, we performed a biological analysis on individual tumor proteomes by matching data to known biological pathways. We here present a novel analysis approach that combines pathway alteration data on individual tumor level with multivariate statistics for HPV and relapse status comparisons. Four proteins (signal transducer and activator of transcription-1, myxovirus resistance protein 1, proteasome subunit alpha type-5 and legumain) identified as main classifiers of relapse status were validated by immunohistochemistry (IHC). Two of the proteins are interferon-regulated and on mRNA level known to be repressed by HPV. By both liquid-chromatography tandem mass spectrometry and immunohistochemistry data we could single out a subgroup of HPV negative/relapse-associated tumors. The pathway level data analysis confirmed three of the proteins, and further identified the ubiquitin-proteasome pathway as altered in the high risk subgroup. We show that pathway fingerprinting with resolution on individual tumor level adds biological information that strengthens a generalized protein analysis. Molecular & Cellular Proteomics 11: 10.1074/mcp.M112.016998, 1–14, 2012.

Vulvar squamous cell carcinoma (VSCC) is a rare female genital skin tumor, and the fourth most common gynecological cancer. Based on etiology VSCC is divided into two subtypes; one related to infection with human papilloma virus (HPV) and one tumor group that is HPV negative (1). The proportion of vulvar cancer linked to HPV infection varies between different studies, reported percentages range between 9–70% (2–9). In a recent study based on the same cohort as the present study, Lindell et al. detected HPV in 31% of VSCC, which agrees well with the prevalence observed in other Nordic studies (22–52% HPV positive) (3, 7, 8, 10).

Based on clinical and histopathological features, the two VSCC subtypes (HPV positive/negative) are postulated to develop via separate intracellular signaling pathways preceded by their own type of premalignant lesion (1). On a molecular level, HPV positive VSCC is associated with increased expression of proteins p16INK4A and p14ARF (11–14). HPV negative VSCC is on the other hand often associated with low expression of p16INK4A and p14ARF and a higher frequency of p53 mutations or deletions (14–16). The patient age distribution differs between the two VSCC types. HPV related vulvar carcinoma is linked to younger age compared with HPV neg-

1 The abbreviations used are: VSCC, vulvar squamous cell carcinoma; CV-ANOVA, cross-validation analysis of variance; FDR, false discovery rate; HiRIEF, high-resolution isoelectric focusing; HPV, human papilloma virus; iTRAQ, isobaric tag for relative and absolute quantitation; OPLS, orthogonal partial least squares; PCA, principal component analysis; PQPQ, protein quantification by peptide quality control; VIP, variable importance in the projection.
ative carcinoma (median age approx. 65 and 75 years, respectively) (10, 14, 17).

While the pathways underlying the pathological alterations in HPV negative VSCC are largely unknown, the oncogenic pathways induced by HPV are more studied; for a review see (18). The HPV is a sexually transmitted double stranded DNA virus with a tropism for squamous epithelium. The infection is in most cases transient, but a subgroup of HPV strains (high risk, hr-HPV) may cause cancer. Hr-HPV are recognized as the cause of cervical carcinoma (99% of these cancers are HPV positive) and also play a role in other malignancies (19, 20). Although much information on HPV induced oncogenic pathways exist on mRNA level, few in-depth proteomic studies combining clinical samples, extensive fractionation and high resolution mass spectrometry have been performed and no such studies on VSCC.

The treatment of VSCC is mainly surgery, but depending on stage of disease, radiotherapy and chemotherapy may be given (21). Several studies indicate a positive correlation between HPV positive VSCC and favorable prognosis (7, 10, 11, 22). Poor disease specific survival has been linked to HPV negative tumors with high p53 levels and low p14ARF (7). There are however also studies showing no prognostic importance of HPV status in vulvar carcinoma, making the relation between HPV status and relapse uncertain (5, 14, 23). The current study aims at identifying patients with low relapse risk for whom less radical surgery could be an option, thus improving quality of life postsurgery. To gain this information on the molecular phenotype related to relapse risk, we performed mass spectrometry based quantitative proteomics on 14 vulvar carcinoma tumor specimens.

Predicting tumor development and extracting information on tumor-driving molecular changes is one of the biggest challenges for clinical oncology. Tumor proteomics provides information on phenotype level which combines genetic alterations and environmental effects and therefore gives valuable information guiding therapy selection. However, comparative proteomic studies on clinical material suffer from large inter-individual variation. To extract as much information as possible from the samples in this study, we introduced a novel strategy to analyze the data. The strategy includes pathway analysis on individual tumor level, which is evaluated on group level by multivariate analysis. The underlying hypothesis is that individual molecular networks drive the growth of each tumor. By analyzing the proteomics data on individual level we can also use all quantified proteins in each sample and are not limited to the overlapping protein identities. In combination with the data analysis on individual level, a more standard analysis on group level was made. In addition to identify a tumor phenotype associated with relapse, we aim here to provide the first in-depth data on the molecular pathways underlying HPV driven and HPV negative VSCC.

We believe that analyzing protein profiles on both individual tumor level and on a group level is a rational strategy that facilitates detection of potential patient subgroups, and may assist future individualized therapy.

**EXPERIMENTAL PROCEDURES**

**Patient Selection and Tumor Samples**—Clinical tumor samples were obtained from the Karolinska University Hospital in Solna, Sweden. The samples were surgical specimens collected between 2003 and 2006, in accordance to the Helsinki Declaration and with approval from the regional ethics committee in Stockholm 2003 (d.nr 02–222). The samples were snap frozen in liquid nitrogen and stored in an internal biobank. Representative tumor samples from 14 patients diagnosed with VSCC (7 HPV positive and 7 HPV negative tumors) were selected from a cohort of 37 samples and matched in sample pairs based on the 14 clinical variables listed in supplemental Table S1. Sample selection was performed by a pathologist, independently verified by a second pathologist and inspected regarding percentage of tumor cell content to minimize intra-sample heterogeneity. Classification of tumors as HPV negative or positive was performed based on PCR test results. General and type specific primers for HPV DNA were used on DNA extracts from selected parts of the tumor as previously described (10). Six of the HPV positive tumors were positive for HPV-16, whereas one tumor was positive for the HPV-53 variant. The samples were selected so that relapse status was an independent clinical factor, which made it possible to classify the samples according to relapse status independently of HPV status. Tumors from patients who remained disease free for 3 years were defined as relapse free. Clinicopathological characteristics of patients included in the study are shown in Table I.

Paraffin-embedded formalin-fixed tumor tissues from nine of the 14 samples analyzed by MS; plus three additional tumor samples; were used for immunohistochemistry (IHC) validation. The added samples were selected from the cohort of 37 samples to create matched pairs based on the same clinical variables as in the original sample cohort. In total, 12 samples were analyzed by IHC.

**Proteomics Analysis: Sample Preparation and Protein Extraction**—Selected tumor samples (wet weight ranging 0.013–0.093 g/sample) were homogenized using a Mixer mill (MM200, Retsch) with a 1 cm Teflon coated tungsten ball (Retsch) precooled in liquid nitrogen. The homogenization was performed in at least three steps of 2 min, between which the samples were put in liquid nitrogen for at least 2 min. Homogenized tissue was solubilized in sample buffer (250–400 μl depending on sample size) consisting of 4% SDS, 25 mM HEPES pH 7.6, 1 mM dithiothreitol, and stored at −80 °C until protein extraction. Protein extraction was performed by cell lysis at 90 °C for 10 min, followed by sonication at room temperature to shear DNA and reduce the viscosity. The lysates were cleared by centrifugation (16000 × g for 10 min), and the supernatant containing the soluble
protein fraction was transferred to a new tube. Protein concentration was determined using Bio-Rad DCC. The protein yield for the samples ranged 1.2–6.4 mg. For each sample, 500 μg of protein was then precipitated in four volumes of ice cold acetone to remove lysis buffer. The samples were kept on ice for one hour and centrifuged for 10 min at 4 °C and 12,000 × g. The supernatant was discarded and the protein pellet was allowed to air dry. The pellets were dissolved in 0.2% SDS and protein concentration was determined again.

**Protein Digestion and iTRAQ Labeling**—We used isobaric labels, isobaric tags for relative and absolute quantitation (iTRAQ) for mass spectrometry based peptide quantification. Three hundred micrograms of protein from each tumor sample was reduced and alkylated with iodoacetamide, and then digested with trypsin according to the manufacturers (Applied Biosystems, Foster City, CA) instructions. 100 μg of peptides per sample were then labeled with 8-plex iTRAQ as outlined in supplemental Table S2. For standardization between the two sample sets, an internal standard (IS) consisting of a mixture of all 14 samples was created from 14.3 μg of each individual tumor sample. The IS was labeled and included in both iTRAQ sets. Excess iTRAQ reagent and detergents were removed by strong cation exchange solid phase extraction (Strata X-C 33 μm polymeric SCX, Phenomenex). Purified samples were freeze dried in a SpeedVac system and placed in −80 °C overnight.

**Prefractionation by Peptide Isoelectric Focusing**—To increase the proteome coverage we used high resolution isoelectric focusing (HRIEF). Peptide samples were dissolved in 225 μl rehydration solution containing 8 M urea, and applied to the gel bridge. For reswelling of the IPG strip, 1% IPG pharmalyte pH 2.5–5.0 (GE Healthcare) was used. Acidic (pI 3.7–4.9) narrow range, 24 cm linear gradient IPG strips (GE Healthcare) were incubated overnight. Samples were applied to the IPG strips by the gel bridge (pH 3.7) at the cathode end and run as described in (24). After focusing, the peptides were passively eluted into 72 contiguous fractions with MilliQ water using an in-house constructed IPG extractor robotics (GE Healthcare Bio-Sciences AB, prototype instrument). Peptides in pI interval 3.7–4.9 were used for identification and quantification. This pI interval contains peptides from 96% of all proteins. The gain of the isoelectric focusing is a sample of greatly reduced complexity (24). The method is compatible with iTRAQ labeling (25). The resulting fractions were freeze dried and kept at −20 °C.

**LC-MS/MS Analysis**—Prior to injection, each peptide IPG fraction was re-dissolved in 8 μl of 3% acetonitrile 0.1% formic acid by the autosampler of an Agilent HPLC 1200 system (Agilent Technologies, Santa Clara, CA) using the injection program mode. A volume of 3 μl was injected to online HPLC-MS (LTQ-Orbitrap Velos mass spectrometer, Thermo Fisher Scientific, San Jose, CA). The HPLC 1200 system provided the gradient for online reversed-phase nano-LC at a flow of 0.4 μl/min. Solvent A was 97% water, 3% acetonitrile, 0.1% formic acid; and solvent B was 5% water, 95% acetonitrile, 0.1% formic acid. The curved gradient went from 2% B up to 40% B in 45min, followed by a steep increase to 100% B in 5 min. The sample was injected into a C-18 guard desalting column (Agilent Technologies, Santa Clara, CA) prior to a 15 cm long C-18 picofrit column (100 μm internal diameter, 5 μm bead size, Nikkyo Technos Co., Tokyo, Japan) installed on the nano electro-spray ionization (NSI) source of the Orbitrap Velos instrument. Acquisition proceeded in −3.5s scan cycles, starting by a single full scan MS at 30,000 resolution (profile mode), followed by two stages of data-dependent tandem MS (centroid mode); the top 5 ions from the full scan MS were selected firstly for collision induced dissociation (CID, at 35% energy) with MS/MS detection in the ion trap, and finally for higher energy collision dissociation (HCD, at 45% energy) with MS/MS detection in the Orbitrap. Precursors were isolated with a 2 m/z width and dynamic exclusion was used with 60 s duration.

**Protein Identification and Quantification**—All MS/MS spectra were searched by Mascot (26) 2.2 (Matrix Science Limited, London, UK) under the software platform Proteome Discoverer 1.1 (Thermo) against the NCBI non redundant human target-decoy protein database (20090118, 32118 sequences). The reversed sequences were used as decoy. Searches against a human protein database with added protein sequences of HPV-16 and HPV-53 gave no identifications of viral peptides. Subsequent searches were performed against a database composed of only human sequences. Prior to searching the data, we used a threshold of signal-to-noise ratio 1.5 on MS1 data. Charge states 2–7 were considered. Precursor mass tolerance of 10 ppm and product mass tolerances of 0.02 Da for HCD-FTMS and 0.8 Da for CID-ITMS were used. Additional settings were: trypsin with 1 missed cleavage; carboxyamidomethylation on cysteine and iTRAQ-8plex on lysine and N-terminal as fixed modifications; and oxidation of methionine as variable modification. Quantification of iTRAQ-8plex reporter ions was performed by Proteome Discoverer (version 1.2) on HCD-FTMS tandem mass spectra using an integration window tolerance of 20 ppm. The quantification was performed based on the peak intensities of the reporter ions in the MS/MS spectra. The relative abundances among samples between the sample sets were calculated by normalizing each peptide signal to the corresponding peptide signal of the pooled internal standard.

**Data Curation**—By MAYU algorithm (27) we identified the peptide score limits (34/33 for sample set 1/2) for the protein identification confidence of 5% false discovery rate (FDR). Protein identifications were based on at least one unique peptide above the score. The protein quantities from Proteome Discoverer were further curated using the in-house developed software POPO (28). POPO checks all the peptides matched to a protein by analyzing the quantitative pattern over samples. Based on the assumption that peptides originating from the same protein should show a correlating pattern, POPO can eliminate outlier peptides and detect possible protein variants. The peptide data was also normalized so that the median peptide intensities were equal between samples. Further, all proteins quantified by POPO had an identification based on at least two unique peptides, with at least one of the peptides being above the score limit. No noise limits were defined for reporter ion intensities.

**Protein Annotation and Pathway Analysis**—For general characterization of the properties of the proteins in the data set, Gene ontology annotation (cellular component, molecular function and biological process) was performed using ProteinCenter (Proxeon Biosystems, now Thermo Fisher Scientific, San José, CA). To further examine protein quantity alterations, matching to pathways was performed using the web based software from Ingenuity Systems; Mountain View, CA (Ingenuity Pathway Analysis, IPA, www.ingenuity.com). In brief, matching of proteins from the data set was performed against the Ingenuity pathway database of well known (canonical) pathways. The degree of matching is in IPA ranked by -log(p), where p is a measure of the probability that the pathway is associated with the data set by random chance (Fisher’s exact test, right-tailed). The -log(p) is from now denoted association rank. The higher the association rank, the stronger is the canonical pathway-data set matching. The analysis was performed both using a data subset consisting of only proteins shared between the two iTRAQ sets, as well as using individual tumor protein profiles as described in the “Results” section. Fig. 1 gives an overview of the data analysis workflow.

**Immunohistochemistry**—Sections of formalin-fixed 4 μm thick paraffin-embedded specimens from 12 VSCC patients were stained with antibodies generated by the Human Protein Atlas Project (29). The sections were deparaffinized in xylene, rehydrated in 70% ethanol after which antigen retrieval was carried out by heating (20 min, 98 °C in citrate buffer; pH 6.0) and rinsed in water. Tissue peroxidase activity was quenched by 0.5% hydrogen peroxide for 30 min, fol-
lowed by rinsing with Tris-buffered saline (TBS) and protein blocking (5% dry milk in TBS, 30 min). Sections were incubated with primary antibodies against STAT1 (1:500), PSMA5 (1:700), MX1 (1:350), and LGMN (1:25). For STAT1 and MX1, a second step of protein blocking was performed (2% dry milk in TBS, 20 min). The specific binding was followed by incubation with goat-anti-rabbit antibodies conjugated to biotin, followed by staining using avidin-coupled peroxidase. Counterstaining was performed with hematoxylin.

The immunohistochemistry (IHC) analyses were performed using an Olympus BX40 microscope. In a blinded evaluation, the slides were independently reviewed by two of the authors. Any differences between the two independent evaluations resulted in re-evaluation and consensus decision. The results are given in percentages of stained tumor cells per total area of tumor cells in the sample preparation. Staining intensity was graded using a scale from 0–3; with 0 representing negative/no staining and 3 strong staining. Negative and positive controls for the respective antibodies are shown in supplementary file S3. For illustration purposes a confocal microscope (Zeiss LSM 78) was used.

Statistical Analysis—For univariate comparison of patient groups, Student’s t test was performed on log 2 -transformed data using SAM (Significance Analysis of Microarrays) for excel version 3.09 and the t test function in excel (2 sided, 2 sample test) (30). SAM performs t test with permutation based correction for multiple testing. Although originally designed for array data, SAM has been shown to be valid also for LC-MS/MS data (31). Associations between protein expression measured by IHC and MS methods were analyzed by Pearson correlation and plotted using MATLAB (The MathWorks Inc., Natick, MA, 2011).

Cluster Analysis—To visualize clustering of groups, a two-way (by protein and tumor ID) hierarchical clustering was performed on log 2 -transformed data using the freeware Genesis, version 1.7.6 (32). Further multivariate statistics and modeling was performed with SIMCA (SIMCA-p + 12.0, Umetrics, Sweden) (33). The analysis was performed on mean centered, unit variance scaled data, assuming equal importance of each protein regardless of relative abundance and magnitude of variance between samples. Principal component analysis (PCA) (34, 35) was performed to get an overview of the data, detect clustering of the data and pick out outliers if any. The PCA summarizes the variation of the data matrix (i.e. protein ratios) and shows the relationship between the observations. For classification and identification of proteins separating HPV positive from HPV negative and relapse from relapse free tumors/patients, we used orthogonal partial least square analysis, OPLS (36). The OPLS analysis detects the protein expression data that covaries with the defined clinical groups. For optimization of the OPLS models, we used the VIP (Variable Importance in the projection) value to judge protein influence (including prediction performance) on the model. The OPLS models were validated by sevenfold full cross validation. Proteins with high VIP throughout the cross validation of the model (95% confidence interval) were selected for the optimized model. We used the plots of the scores predicted in the cross validation and analysis of variance (CV-ANOVA) to judge the model validity.

RESULTS

Proteomics Analysis: Protein Selection—After FDR assessment (5%) using MAYU, 2084 protein entries were considered accurate in terms of identification. The identified proteins are listed in supplementary file S2, and single peptide identifications annotated spectra are shown in supplementary file S3. 1579 proteins were represented by at least two unique peptides and estimated by PQPQ to be quantitatively accurate. All unique peptides with quantitative information are listed in supplementary file S4. Of the 1579 proteins, 449 were present in both iTRAQ sample sets and used for comparison of clinical sample groups defined by HPV and/or relapse status. This group level analysis is referred to as global analysis. In the individual tumor level analysis, all accurately quantified proteins were included if present in the internal standard (IS) for that particular iTRAQ sample set. The quantitative value for each protein is related to the IS in its sample set and the value 1 therefore indicates no quantitative change relative to the sample mean, as represented by the IS. In the individual analysis we were interested in proteins differentially expressed relative to the sample mean, in other terms protein level changes characterizing each tumor from the mean. We thus based the pathway analysis on proteins with levels altered (with a 99% confidence interval) compared with the sample mean (IS). The estimation of experimental noise was based on previous experimental data (supplementary file S1).

Based on this estimation, only proteins with ratio greater than 1.18, or less than 0.85 relative the IS were included in the individual tumor analysis. These proteins were differing from sample mean with more than what could be expected by the experimental noise. On average 546 (range: 313–895) protein entries from each tumor sample fulfilled this criterion. We have created a downloadable and searchable database (http://tools.scilifelab.se/VSCCtpd) of all 1566 quantified proteins used in the analyses, which can be used to obtain expression profiles for proteins of interest over the individual tumor samples. The database layout is exemplified in supplemental Fig. S1. A schematic outline of the global and individual analyses is shown in Fig. 1.

The two proteins sets used in the global analysis (449) and in the individual analysis (totally 1566) were functionally classified according to Gene Ontology and compared using ProteinCenter™ (supplemental Fig. S2). The comparison revealed no significant differences in enrichment of molecular functions or cellular components between the two protein sets.

Global Analysis

Unsupervised Statistical Analysis by PCA—One sample (HPV 42) was detected as outlier in the PCA scores plot (Fig. 2). This tumor was found to be of different histology (adenocystic carcinoma and not squamous cell carcinoma like the other tumors in the study); it was hence removed from the subsequent analyses. The PCA also showed that the pooled internal standards were centered. This is expected as they consist of equal aliquots of each of the 14 samples, thus representing a sample mean. The samples did however not cluster in terms of HPV or relapse status, showing that these clinical characteristics were not the source of the largest variation in the data set.

Classification Based on HPV Status—We further investigated if we could detect differences between HPV positive
and negative tumors in terms of protein levels in VSCC. As determined by student’s t test in SAM, the four statistically most significant proteins at an FDR of 22% were: Collagen type I alpha 2 and alpha 1 (COL1A1 and 2); periostin osteoblast specific factor (POSTN) and fibrillin 1 (FBN1). These proteins were up-regulated in HPV positive tumors. The high FDR level could be explained by the high between-individual variation of the clinical samples.

The cross-validated OPLS analysis (Fig. 3A) for classifying HPV status confirmed these four proteins. The OPLS model, initially based on all 449 shared proteins, was optimized by stepwise removal of proteins with small VIP (Variable Importance in the projection) value. This was performed until the model did not improve anymore as judged by the CV-ANOVA p value, indicative of the probability that the model is the result of chance alone. The optimized OPLS model included 43 proteins (p = 0.015), which are listed in supplemental Table S4 with VIP, fold change, and t test derived p values for each protein.

We performed a pathway analysis on the 43 proteins using Ingenuity Pathway analysis (IPA). In IPA, the 43 proteins were matched against a database consisting of known protein signaling pathways. This gave 28 significant pathways (p < 0.05). The top four were α-adrenergic signaling (p = 0.00007), Protein Kinase A signaling (p = 0.0001), caveolar mediated endocytosis signaling (p = 0.0009) and breast cancer regulation by stathmin (p = 0.001). In supplemental Fig. S3 a network generated by IPA using the OPLS loadings shows the relative protein up- or down-regulation on group level.

Classification Based on Relapse Status—We further sought the protein pattern that could discriminate between relapse and nonrelapse tumor groups, regardless of HPV status. The
The top four significant proteins, as determined by student's t test with correction for multiple testing, were up-regulated in the relapse group (n = 5) compared with the nonrelapse patient group (n = 8). These proteins were tryptophanyl-tRNA synthetase (WARS), signal transducer and activator of transcription 191kDa (STAT1), 2'-5'-oligoadenylate cyclase (OAS3) and myxovirus resistance 1 interferon-inducible protein p78 (MX1). As in the previous OPLS with HPV as classifier, the FDR was high (27%), most likely explained by large intertumor variation. An OPLS-model with relapse status as response variable was optimized in the same way as described in the previous section. The final OPLS model based on 16 proteins (p = 0.007) confirmed WARS, STAT1, 2'-5'-OAS and MX1 as classifiers of relapse status (Fig. 3B). The 16 proteins are listed in supplemental Table S5 with VIP, fold change and t test derived p values for each protein. There were no overlapping proteins between the two OPLS models based on HPV status and relapse status.

The proteins from the OPLS model were used for protein network generation and matched to canonical pathways using IPA as described above for the HPV query (supplemental Fig. S4). The analysis revealed 11 significantly altered pathways (p < 0.05). Those with the strongest significance were the IFN signaling pathway (p = 0.0005), followed by ERK/MAPK signaling, IL-22 signaling and the protein ubiquitination pathway (all with p = 0.03). The IFN signaling pathway with protein quantities indicated is shown in supplemental Fig. S5.

To see whether we had proteins that were unique to any of the clinical groups, we also checked the distribution of the quantified proteins over the clinical sample groups. We saw 23 proteins that were present in most relapse tumors, whereas absent in most relapse negative tumors (supplemental Fig. S6).

**Individual Tumor Analysis**

Pathway analysis was also performed on each of the 14 tumor samples separately. By this we obtain a fingerprint of affected pathways for each tumor. As described in Fig. 1, all proteins with a level differing from the mean protein level among all tumors were included in the individual tumor analysis; in total 1566 proteins. The relative intensities of these tumor specific proteins were mapped to the canonical pathways in the IPA database. The analysis identified the pathways that were most significant in each of the 14 individual tumor data sets, measured by Fisher's exact test. We used the association ranks (described under “Experimental Procedures”) as input variables to the multivariate analysis for sample comparison, thus based on pathway enrichment.

**Unsupervised Analysis on Pathway Association Data**—To investigate if the proteomics data of the individual VSCC tumors...
tumor samples revealed any pathway alteration based on subgroupings of tumors, we performed hierarchal clustering on the pathway association rank value. We included 13 samples in the analysis, as one sample had been previously identified as an outlier. The result is shown in supplemental Fig. S7. The many nonoverlapping proteins between the two sample sets affected also the pathways identified, and consequently the main clustering was according to sample set belonging. A PCA showed the same pattern as the hierarchal analysis (supplemental Fig. S8). These analyses demonstrates that the largest variation in the pathway association rank data set is not because of HPV or relapse, and that we require supervised analysis methods such as t test and OPLS to detect any proteome biology differences related to these questions.

Connecting Pathway Alterations to HPV Status and Relapse Risk—To detect pathway alterations connected to HPV status and relapse risk, we performed an OPLS analysis on the pathway association data. The OPLS model was optimized as described in the group analysis. We performed stepwise removal of variables (i.e. pathways) with less influence on the separation performance of the model until the model did not improve anymore. The model was judged by separation performance, robustness (cross-validation plot) and associated p value. The final model for classification of HPV status was based on 16 variables (pathways) with p value 0.003. The OPLS loadings in this analysis reflect how much the pathways are affected given the HPV status. Interpreting the OPLS loadings from this model, the most significant pathway in the HPV positive tumors was integrin linked kinase signaling. The top significant pathways for HPV negative tumors were glucocorticoid signaling, neuregulin signaling and purine metabolism. Fig. 4A shows the OPLS model and the top ranked significant pathways.

Corresponding pathway analysis on individual tumor data and subsequent OPLS analysis as described above was performed using relapse as class-identifier. The optimized model was based on 36 variables (p 0.002) and is shown in Fig. 4B, together with the most significant pathways. The top-ranked pathway in relapse cases was “RIG-1 like receptors in antiviral innate immunity” and in nonrelapse cases “B-cell development” and “Rac signaling pathway”.

Immunohistochemistry Validation

Staining of Selected Proteins—Four proteins, STAT1, PSMA5, MX1, and LGMN, were selected for immunohistochemistry (IHC) validation. The selection was based on the results from the t test and OPLS model for the global level classification of relapse status and further strengthened by the individual tumor analysis of both relapse and HPV status. STAT1 and MX1 are part of the interferon signaling pathway altered in the HPV negative tumors, and PSMA5 belongs to the protein ubiquitination pathway altered in relapse-associated tumors. The IHC results are listed with the MS results (iTRAQ ratio) in supplemental Table S6. PSMA5, MX1, and LGMN stained predominantly tumor cells in the VSCC samples, whereas STAT1 stained both tumor and lymphoid cells (lymphocytes). The correlation between the MS quantitative data and the IHC data for each protein is shown in Fig. 5A. The tumor samples are visualized as four groups based on relapse and HPV status. We note that tumor 42, which was detected as outlier in the PCA (global analysis) and excluded from all further data analysis, did not stain for any of the proteins in the IHC analysis. The outlier status of tumor 42 is hence confirmed once again. Representative staining of LGMN, which is the protein with the strongest correlation between MS and IHC data, is shown in Fig. 5B. As two of the proteins (LGMN and PSMA5) were selected as part of a protein panel in the OPLS model on the MS data, we also evaluated the four proteins together as a panel in an OPLS model based on the IHC data. As in previous OPLS models, model performance was evaluated by cross validation. The panel classified nine samples out of 12 correctly in terms of relapse status (supplemental Fig. S9).

DISCUSSION

We have performed a proteomics analysis on VSCC tumors using high resolution isoelectric focusing and LC-MS/MS. We used a novel data analysis approach by combining biological pathway mapping on individual tumor level with multivariate analysis to compare HPV and relapse status groups. This analysis allows detection of subgroups of patients with specific pathways contributing to high risk; potentially important information when aiming to select targeted cancer therapy combinations for future clinical trials. We demonstrate the validity of this approach by detecting known molecular alterations in HPV related carcinoma, some detected on protein level for the first time, and several novel observations of pathways linked to high risk subgroup of VSCC with potential therapeutic implications.

Sample Selection—The aim of this study was to investigate the link between HPV induced VSCC and relapse. Existing knowledge on the prognostic impact of HPV status is contradictory. We have addressed the question by analyzing pathway activation differences in 14 VSCC tumor samples. Although sample selection was made so that HPV and relapse status were independent factors, we detected an overlap between activated pathways in HPV positive tumors and nonrelapse, and between HPV negative tumors and relapse. Prognostic factors known to influence the relapse risk, tumor stage and metastasis, were also considered in the sample selection and were not confounders in the selected samples. Possible confounders with HPV status groups that we could not correct for were age and smoking status (supplemental Table S2). However, there were no correlations between relapse and age or smoking.
Sample Analysis—The isobaric labeling (iTRAQ) used for quantification provided good data on relative protein levels for samples analyzed within a sample set. However, because of high patient-to-patient tumor proteome variability and possibly also intratumor heterogeneity combined with high sample complexity; the overlap of identities between the two sample sets was poor. The highly variable protein composition also resulted in a relatively low number of significant proteins, and high FDR in the statistical analysis. In this work, we addressed the poor overlap between the sample sets by combining two lines of tumor protein data analysis: a) global comparison of altered protein levels between predetermined patient groups and b) individual tumor pathway profiling based on altered protein levels, followed by comparison of pathway association ranks in a supervised multivariate analysis. By this approach we included quantitative information from 1566 proteins in the individual analysis in addition to the 449 proteins overlapping between iTRAQ sample sets that were used in the global analysis.

Data Analysis and Validation—Based on the global analysis we selected 4 deregulated proteins associated with relapse (independent of HPV status); PSMA5, STAT1, MX1, and LGMN, for immunohistochemistry based validation of the mass spectrometry results. The selection of PSMA5, STAT1, and MX1 was supported by data from the individual tumor analysis associated with relapse.

By the bivariate visualization of MS and IHC data shown in Fig. 5A, we were able to single out a patient subgroup of HPV negative and relapse (HPV-/Rel+). It is notable that although the correlation between IHC and iTRAQ data was strong for proteins LGMN and MX1 (R² 0.95 and 0.87, respectively), the corresponding correlation for STAT1 and PSMA5 (R² 0.56 and 0.66, respectively) was more modest. There are at least two plausible explanations for the discrepancy between the measured IHC and MS quantities. First, although we controlled for intrasample heterogeneity by selecting tumor specimens of high homogeneity, each tumor still contains several cell types. Whether the protein is expressed in a tumor cell or an immune cell, it will be quantified in the MS analysis as part of the tumor lysate. On the other hand, in the IHC analysis, we specifically scored protein staining on tumor cells. In the case of STAT1, we could by the IHC see that also some tumor infiltrated lymphocytes stained. This could explain the poor correlation observed. Second, it has been shown that iTRAQ ratios compress toward 1, leading to an underestimation of ratios. Although the advantages of iTRAQ are high-throughput relative protein quantification suitable for hypothesis generation and finding proposed protein markers, more precise quantitative methods such as targeted proteomics using absolute

![Image](attachment:image_url)
A. Correlation between mass spectrometry (iTRAQ ratio) and immunohistochemistry (staining percentage) data for proteins STAT1, PSMA5, MX1, and LGMN. The strongest correlation ($R^2 = 0.95$) between IHC and MS data was observed for the protein LGMN. We visualized patient subgroup classification based on IHC and MS data by plotting sample subgroup mean and adding measured protein level differences within the patient groups (one standard deviation). Thus, the horizontal bars show within subgroup variation for the MS measurements, whereas the vertical bars show the within subgroup variation for the IHC measurements. The HPV negative tumors that have relapsed (Relapse+/HPV- group, three cases) separates from the other sample groups. The HPV+/Relapse- group included 2 samples, the HPV+/Relapse+ group included 5 samples and the HPV-/Relapse- 4 samples. B. Cellular localization and staining pattern of LGMN protein in four VSCC tumor samples of different relapse and HPV status. The HPV positive cases showed a decreased percentage of stained cells and as well as weaker staining intensity compared with the HPV negative cases.

Fig. 5. A, Correlation between mass spectrometry (iTRAQ ratio) and immunohistochemistry (staining percentage) data for proteins STAT1, PSMA5, MX1, and LGMN. The strongest correlation ($R^2 = 0.95$) between IHC and MS data was observed for the protein LGMN. We visualized patient subgroup classification based on IHC and MS data by plotting sample subgroup mean and adding measured protein level differences within the patient groups (one standard deviation). Thus, the horizontal bars show within subgroup variation for the MS measurements, whereas the vertical bars show the within subgroup variation for the IHC measurements. The HPV negative tumors that have relapsed (Relapse+/HPV- group, three cases) separates from the other sample groups. The HPV+/Relapse- group included 2 samples, the HPV+/Relapse+ group included 5 samples and the HPV-/Relapse- 4 samples. B, Cellular localization and staining pattern of LGMN protein in four VSCC tumor samples of different relapse and HPV status. The HPV positive cases showed a decreased percentage of stained cells and as well as weaker staining intensity compared with the HPV negative cases.
Individual VSCC Tumor Pathway Data for Group Comparison

quantification or fluorescent labeling with improved dynamic range of quantification on tissue sections, are future options. In addition, conventional IHC quantification depends on multiple factors, such as antibody specificity, antigen retrieval, fixation of tissue etc. However, even taking these factors into consideration, the correlation between the two methods for LGMN was very good ($R^2 = 0.95$). This indicates that for at least LGMN and MX1, a larger clinical validation could be performed without the development of more accurate methods for quantification in the first step. Using both methods, our results show that the patient subgroup of HPV-/Rel+ discriminates from the others based on the levels of the selected proteins. These patients may therefore be detected before treatment and might benefit from being treated as a separate group also in terms of clinical therapy. The biological role of the selected proteins is discussed below.

Biological Interpretation—Toward Individualized VSCC Therapy

Relapse Linked to Alterations of the “Ubiquitin-Proteasome” and of the “Interferon” Signaling Pathways—We focus our discussion on two pathways: the STAT1/interferon signaling pathway and the ubiquitin-proteasome signaling pathway. Both pathways are supported by results from the global level and individual tumor level analysis. In addition, targeted therapies against these pathways are in clinical use, although not against VSCC. Hence potential future studies based on these results may be performed using already approved drugs although on a new target disease.

Increased levels of proteasome subunit alpha type-5, (PSMA5) were linked to relapse in the global analysis (Fig. 3B), and part of the protein ubiquitination pathway was over-represented in relapse cases in the individual tumor analysis (Fig. 4). Interestingly, the protein ubiquitination pathway ranked first for all samples in the patient subgroup with HPV-negative and relapse (samples 40, 45, 49) in the individual analysis, but was ranked first for only one other tumor. The ubiquitin-proteasome pathway degrades intracellular proteins, and is frequently up-regulated in cancer cells. Hampered proteasome activity causes accumulation of proteins which may lead to cell death, hence the use of proteasome inhibitors to induce apoptosis in cancer therapy (39, 40). Proteasome inhibitors are also under evaluation for treatment against squamous cell carcinoma tumors (39, 41, 42), confirming the relevance of the proteasome in these cancers. In our IHC analysis, high levels of PSMA5 were associated with HPV negative tumors ($p < 0.04$, supplemental Table S6). It is possible that HPV infection down-regulates the proteasome pathway, which is responsible for viral antigen peptide presentation on MHC-1 molecules via PSMA5, thereby causing the observed difference between the two tumor groups (43). The clear protein level variation of PSMA5 between tumors and its relation to disease outcome in our study justify further research on the possible use of PSMA5 level as nominator of a VSCC patient subgroup that would benefit from proteasome inhibitors.

Among the proteins identified as classifiers of relapse status in the global analysis, three were proteins of the interferon signaling pathway: signal transducer and activator of transcription 1 (STAT1), interferon-inducible dynamine-like myxovirus resistance protein 1 (MX1) and 2’-5’ oligoadenylate synthetase (OAS), shown in Fig. 6 and in supplemental Fig. S5. STAT1, MX1, and OAS are IFN-induced and known from mRNA level studies to be repressed by hr-HPV proteins E6 and E7 (44). In line with this, we observed lower protein levels of STAT1 in HPV positive tumors in our IHC assay (Fig. 5A, $p < 0.05$). In the clinic, the IFN pathway is activated by administrating IFN-α to treat HPV-induced warts, although not all patients respond to treatment. Nonresponders are shown to have higher pretreatment E7 mRNA levels than responders (45), indicating that viral protein E7 levels are reflected on the IFN pathway.

The cytoplasmic protein MX1 is induced by interferon α/β and further requires STAT1 signaling (46, 47). As a mediator of innate immunity against viruses it redistributes to sites of viral replication where it promotes missorting of viral constituents (48, 49). Both MX1 and OAS3 are located downstream of STAT1 in the signaling cascade (Fig. 6), and like STAT1 they are down-regulated by hr-HPV, which is also seen in our protein level data. The list of significant pathways contains some overlapping proteins. Therefore, the altered STAT1 protein levels and downstream effects are also reflected in several other pathways in the individual pathway analysis (Fig. 4), notably the signaling pathways of the glucocorticoid receptor, thrombopoietin, interferon, oncostatin M and EGFR.

In agreement with the STAT1 protein level changes in our study, the STAT1 induced signaling pathway “RIG-1 (retinoid-inducible negative growth regulator) like receptors in antiviral innate immunity” was enriched among relapse cases (Fig. 4). MX1 and OAS3 require the induction of RIG (50). RIG-1 pro-apoptotic effects are mediated via the extrinsic apoptotic pathway, which includes the Fas receptor and adaptor protein FADD (51). We observed increased Fas levels in relapse cases (supplemental Fig. S5). STAT1 involvement in cell death regulation (52), and up-regulation of death receptors and ligands such as Fas and FasL is known (53). In our data, we observed a positive correlation between STAT1 and Fas protein levels in the relapse group (Spearman rank correlation $p < 0.05$, supplemental Fig. S10). Viral repression of STAT1 may contribute to cancer formation in HPV positive tumors by hindering elimination of (cancer) cells via the extrinsic apoptosis pathway. In HPV negative tumors, STAT1 signaling is not repressed and the RIG-1/IFN pathways and apoptosis via Fas signaling can be induced. Thus in HPV negative tumors, STAT1 is not likely to play a major role in cancer formation.

Increased Levels of Legumain are Linked to High Risk Subgroup—Legumain (LGMN) was up-regulated in relapse cases in the global MS analysis. The protein is not mapped to any canonical pathway and therefore not identified in the individ-
ual tumor analysis. Legumain is a cysteine protease activated by low pH, involved in the lysosomal and endosomal processing of bacterial peptides and endogenous proteins (54, 55). A positive correlation between LGMN and tumor invasion and metastasis has been observed in a mouse colon cancer model (56), and LGMN is proposed as a prognostic marker in breast cancer, with positive vesicular IHC staining correlated to poor outcome (57). Further, LGMN is suggested as a marker of squamous cell nasopharyngeal carcinoma (58). In our IHC analysis, LGMN staining exhibited a granular cytoplasmic pattern, which was significantly stronger and covering a higher percentage of tumor cells in HPV negative and relapse positive tumors. There was a strong correlation between MS and IHC levels of LGMN (R^2 = 0.95, Fig. 5). Recently a connection between LGMN and EGFR turnover was shown in a mouse model (59). Western blot analysis of EGFR

Fig. 6. Schematic showing observed protein level alterations in VSCC proteome in the context of HPV infection. The viral proteins E5, E6, and E7 (not detected on protein level in the study), responsible for many of the oncogenic features of the human papilloma virus (HPV), are indicated in yellow. By inactivation and degradation of pRb and p53, E6, and E7 cause proliferation and apoptosis inhibition. Inhibition of pRb releases the transcription factor E2F, a key regulator of cell cycle genes, which is also under the influence of PA2G4 (ErbB3-binding protein Ebp1) as indicated in the figure. We observed reduced levels of the p53-regulated protein 14–3-3-z, an inhibitor of cell cycle progression (68), among the majority of the HPV+ tumors in our study. Without investigating the connection to HPV, elevated cytoplasmic levels of 14–3-3-z have been linked to advanced disease in VSCC by immunohistochemistry analysis (69). Further, the relative protein levels of two endosomal proteins, Bap31 and Rab11, were decreased in HPV infected tumors. Viral protein E5 is proposed to reduce MHC I levels via Bap31 interaction (18, 70), and decreased levels of Rab11 in cytomegalovirus infected cells has been linked to decreased levels of MHC 1 and impaired degradation of EGFR (71). The immune response is further suppressed by the down-regulation of STAT1 and downstream proteins. We also inserted the IHC validated proteins LGMN and PSMA5 in the figure, although there is no known link to HPV infection. Protein levels up-regulated relative to the sample mean are indicated with red bars; green bars indicate down-regulation. The figure was created using IPA and based on a figure by Moody (18).
in kidneys from LGMN deficient mice revealed increased EGFR levels, interpreted as LGMN involvement in EGFR down-regulation. In our study, we hypothesized whether there is a link to HPV infection via the viral protein E5, which targets EGFR turnover by acidifying endosomes (60), but found no correlation between EGFR and LGMN levels in the MS data.

**Influence of hr-HPV on Host Proteome—**Most HPV positive cases determined by PCR represent transient infections rather than clinically relevant lesions (61). The potential value of a protein marker is therefore not showing the presence of HPV but where HPV infection has an impact on the tumor phenotype. As shown in Fig. 4, the canonical pathway with the highest influence on the OPLS model for HPV classification was the integrin linked kinase (ILK) pathway. The ILK pathway (supplemental Fig. S11) controls cell proliferation, survival and extracellular matrix breakdown. Pathway de-regulation has been linked to HPV (62, 63), although precisely how this contributes to pathogenesis is unclear. Integrins control endosomal trafficking of other receptor tyrosine kinases such as EGFR. However, much of the ILK pathway regulation is mediated via phosphorylations, therefore detailed conclusions of de-regulation could not be drawn based on the results from this study and would justify a separate phospho-proteomics study.

Fig. 6 summarizes the detected HPV related protein changes in VSCC observed in our study. We searched the individual tumor data for known targets of HPV oncogenic proteins: cell cycle regulatory proteins p53, pRb and E2F target genes, p16 <sup>ink4a</sup>, cyclinD1, EGFR. We detected p16 <sup>ink4a</sup> in sample set 2 in 7 tumor samples (3 HPV positive), but HPV status and p16 <sup>ink4a</sup> levels did not correlate. We could see 3 cases (R_HPV-23, R_NHPV-45, and NR_NHPV-52) with clear up-regulation of EGFR protein levels. An increased EGFR expression is observed in 47–67% of vulvar carcinomas, and linked to lymph node metastasis (64–67). We saw two tumors (R_NHPV-45, NR_NHPV-52) with relative EGFR up-regulation that were linked to metastasis. The observed proteome level effects of HPV are further described in the figure legend.

**CONCLUSION**

The analysis approach with multivariate statistics in combination with tumor pathway fingerprinting gave insight on pathways related to VSCC clinical groups with resolution on individual tumor level. The approach benefits from the possibility of detecting subpopulations of patients. Thus, analysis within the context of altered pathways can help to relate proteome data to future targeted cancer therapies.

This exploratory study contributes to the molecular understanding of VSCC and provides a number of potential proteins and pathways. We could distinguish a patient subgroup of relapse HPV- tumors based on the expression of four proteins, and our data suggest that this subgroup is characterized by an altered ubiquitin-proteasome signaling pathway. This is the prognostic group that would benefit most from new therapy options. A molecular definition of this subgroup like the one we have found here is important for treatment and further studies.

Acknowledgments—We thank technicians Birgitta Sundelin, Birgitta Byström and Eva Andersson for help with collecting and preparing the sample, and Inger Bodin for conducting the immunohistochemistry staining. We also thank Otto Manneberg, the ALM Bioimaging platform at the Science for Life Laboratory, Stockholm, Sweden, for help with taking the IHC photos.

* This work was supported by Karolinska Institutet (KID-funding), Stockholm county council (SLI), Swedish research council, the Cancer Society in Stockholm, GlycoHit FP 7 EU project and Karolinska University Hospital (ALF).

** To whom correspondence should be addressed: Clinical Proteomics Mass Spectrometry, Department of Oncology-Pathology, Science for Life Laboratory and Karolinska Institutet, Stockholm, Sweden. Tel.: +46 (0) 852481416; E-mail: janne.lehtio@ki.se.

REFERENCES

1. McCullage, W. G. (2009) Recent developments in vulvovaginal pathology. *Histopathology* 54, 156–173
2. Brandenberger, A. W., Rüdlinger, R., Hänggi, W., Bersinger, N. A., and Dreher, E. (1992) Detection of human papillomavirus in vulvar carcinoma. A study by in situ hybridisation. *Arch. Gynecol.* 252, 31–35
3. Iwasawa, A., Nieminen, P., Lehlinen, M., and Paavonen, J. (1997) Human papillomavirus in squamous cell carcinoma of the vulva by polymerase chain reaction. *Obstet. Gynecol.* 89, 81–84
4. Lerna, E., Matias-Guiu, X., Lee, S. J., and Prat, J. (1999) Squamous cell carcinoma of the vulva: study of ploidy, HPV, p53, and pRb. *Int. J. Gynecol. Pathol.* 18, 191–197
5. Pinto, A. P., Schlecht, N. F., Pintos, J., Kaiano, J., Franco, E. L., Crum, C. P., and Villa, L. L. (2004) Prognostic significance of lymph node variables and human papillomavirus DNA in invasive vulvar carcinoma. *Gynecol. Oncol.* 92, 856–865
6. Skapa, P., Zamecniik, J., Hansikova, E., Salakova, M., Smahelova, J., Jandova, K., Robova, H., Rob, L., and Tachezy, R. (2007) Human papillomavirus (HPV) profiles of vulvar lesions: possible implications for the classification of vulvar squamous cell carcinoma precursors and for the efficacy of prophylactic HPV vaccination. *Am. J. Surg. Pathol.* 31, 1834–1843
7. Knopp, S., Nesland, J. M., Tropé, C., and Holm, R. (2006) p14ARF, a prognostic predictor in HPV-negative vulvar carcinoma. *Am. J. Clin. Pathol.* 125, 266–276
8. Madsen, B. S., Jensen, H. L., van den Brule, A. J., Wohlfahrt, J., and Frisch, M. (2008) Risk factors for invasive squamous cell carcinoma of the vulva and vagina-population-based case-control study in Denmark. *Int. J. Cancer* 122, 2827–2834
9. Sutton, B. C., Allen, R. A., Moore, W. E., and Dunn, S. T. (2008) Distribution of human papillomavirus genotypes in invasive squamous carcinoma of the vulva. *Mod. Pathol.* 21, 345–354
10. Lindell, G., Nasman, A., Jonsson, C., Ehrsson, R. J., Jacobsson, H., Danielsson, K. G., Dalianis, T., Kallström, B. N., and Larsson, B. (2010) Presence of human papillomavirus (HPV) in vulvar squamous cell carcinoma (VSCC) and sentinel node. *Gynecol. Oncol.* 117, 312–316
11. van de Nieuwenhof, H. P., van Kempen, L. C., de Hullu, J. A., Bekkers, R. L., Buiten, J., Melchers, W. J., and Massuger, L. F. (2009) The etiologic role of HPV in vulvar squamous cell carcinoma fine tuned. *Cancer Epidemiol. Biomarkers Prev.* 18, 2591–2597
12. van der Avoort, I. A., Shirango, H., Hoevenaars, B. M., Greffe, J. M., de Hullu, J. A., de Wilde, P. C., Buiten, J., Melchers, W. J., and Massuger, L. F. (2006) Vulvar squamous cell carcinoma is a multifactorial disease following two separate and independent pathways. *Int. J. Gynecol. Pathol.* 25, 22–29
Individual VSCC Tumor Pathway Data for Group Comparison

50. Jiang, L. J., Zhang, N. N., Ding, F., Li, X. Y., Chen, L., Zhang, H. X., Zhang, W., Chen, S. J., Wang, Z. G., Li, J. M., Chen, Z., and Zhu, J. (2011) RA-inducible gene-I induction augments STAT1 activation to inhibit leukemia cell proliferation. Proc. Natl. Acad. Sci. U.S.A. 108, 1897–1902

51. Tsai, F. M., Shyu, R. Y., and Jiang, S. Y. (2007) RIG1 suppresses Ras activation and induces cellular apoptosis at the GoGi apparatus. Cell Signal. 19, 989–999

52. Kim, H. S., and Lee, M. S. (2007) STAT1 as a key modulator of cell death. Cell Signal. 19, 454–465

53. Xu, X., Xu, X. Y., Plate, J., and Chong, A. S. (1998) IFN-gamma induces cell growth inhibition by Fas-mediated apoptosis: requirement of STAT1 protein for up-regulation of Fas and FasL expression. Cancer Res. 58, 2832–2837

54. Watts, C., Matthews, S. P., Mazzeo, D., Manoury, B., and Moss, C. X. (2007) Asparaginyl endopeptidase: case history of a class II MHC compartment protease. Immunol. Rev. 207, 218–228

55. Manoury, B., Mazzeo, D., Li, D. N., Billson, J., Loak, K., Benaroch, P., and Watts, C. (2003) Asparaginyl endopeptidase can initiate the removal of the MHC class II invariant chain chaperone. Immunity 18, 489–498

56. Liu, C., Sun, C., Huang, H., Janda, K., and Edgington, T. (2003) Overexpression of legumain in tumors is significant for invasion/metastasis and a candidate enzymatic target for prodrug therapy. Cancer Res. 63, 2957–2964

57. Gawenda, J., Traub, F., Luck, H. J., Kreipe, H., and von Wasielewski, R. (2007) Legumain expression as a prognostic factor in breast cancer patients. Breast Cancer Res. Treat. 102, 1–6

58. Wu, C. C., Hsu, C. W., Chen, C. D., Yu, C. J., Chang, K. P., Tai, D. I., Liu, H. P., Su, W. H., Chang, Y. S., and Yu, J. S. (2010) Candidate serological biomarkers for cancer identified from the secretomes of 23 cancer cell lines and the human protein atlas. Mol. Cell. Proteomics 9, 1100–1117

59. Miller, G., Matthews, S. P., Reinheckel, T., Fleming, S., and Watts, C. (2011) Asparaginyl endopeptidase is required for normal kidney physiology and homeostasis. FASEB J. 25, 1608–1617

60. Straight, S. W., Herman, B., and McCance, D. J. (1995) The E5 oncoprotein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain. Int. J. Cancer 60, 425–429

61. Schmeink, C. E., Melchers, W. J., Siebers, A. G., Quint, W. G., Massuger, L. F., and Bekkers, R. L. (2011) Human papillomavirus persistence in young unscreened women, a prospective cohort study. PLoS One 6, e27937

62. McCormack, S. J., Brazinski, S. E., Moore, J. L. Jr., Werness, B. A., and Goldstein, D. J. (1997) Activation of the focal adhesion kinase signal transduction pathway in cervical carcinoma cell lines and human genital epithelial cells immortalized with human papillomavirus type 18. Oncogene 15, 265–274

63. Du, M., Fan, X., Hong, E., and Chen, J. J. (2002) Interaction of oncogenic papillomavirus E6 proteins with fibulin-1. Biochem. Biophys. Res. Commun. 296, 962–969

64. Hantschmann, P., Jeschke, U., and Friese, K. (2005) TGF-alpha, c-erbB-2 expression and neoangiogenesis in vulvar squamous cell carcinoma. Anticancer Res. 25, 1731–1737

65. Gordinier, M. E., Steinhoff, M. M., Hogan, J. W., Peipert, J. F., Gajewski, W. H., Falkenberry, S. S., and Granai, C. O. (1997) S-Phase fraction, p53, and HER-2/neu status as predictors of nodal metastasis in early vulvar cancer. Gynecol. Oncol. 67, 200–202

66. Johnson, G. A., Mannel, R., Khalifa, M., Walker, J. L., Wren, M., Min, K. W., and Benbrook, D. M. (1997) Epidermal growth factor receptor in vulvar malignancies and its relationship to metastasis and patient survival. Gynecol. Oncol. 65, 425–429

67. Oonk, M. H., de Bock, G. H., van der Veen, D. J., van der Zee, A. G., and van der Zee, A. G. (2007) EGFR expression is associated with groin node metastases in vulvar cancer, but does not improve their prediction. Gynecol. Oncol. 104, 109–113

68. Hermeking, H., Lengauer, C., Poylik, K., He, T. C., Zhang, L., Thiagalingam, S., Kinzler, K. W., and Vogelstein, B. (1997) 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression. Mol. Cell 1, 3–11

69. Wang, Z., Nesland, J. M., Suo, Z., Trope, C. G., and Holm, R. (2011) The prognostic value of 14–3–3 isoforms in vulvar squamous cell carcinoma cases: 14–3–3beta and epsilon are independent prognostic factors for these tumors. PLoS One 6, e24843

70. Ashrati, G. H., Haghshenas, M., Marchetti, B., and Campbell, M. S. (2006) E5 protein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain. Int. J. Cancer 119, 2105–2112

71. Tomas, M. I., Kucic, N., Mahmutefeidic, H., Blagovevic, G., and Lucin, P. (2010) Murine cytomegalovirus perturbs endosomal trafficking of major histocompatibility complex class I molecules in the early phase of infection. J. Virol. 84, 11101–11112

In order to cite this article properly, please include all of the following information: Sandberg, A.S., Lindell, G., Källström, B. N., Branca, R. M., Danielsson, K. G., Dahlberg, M., Larson, B., Forshed, J., and Lehtio, J. (2012) Tumor Proteomics by Multivariate Analysis on Individual Pathway Data for Characterization of Vulvar Cancer Phenotypes. Mol. Cell. Proteomics 11(7):M112.016998. DOI: 10.1074/mcp.M112.016998.