

The Primary Effect on the Proteome of *ARID1A*-mutated Ovarian Clear Cell Carcinoma is Downregulation of the Mevalonate Pathway at the Post-transcriptional Level[§]

Aaron R. Goldman[‡], Benjamin G. Bitler[§], Zachary Schug[‡], Jose R. Conejo-Garcia[¶],
 Rugang Zhang[§], and David W. Speicher^{‡||**}

Inactivating mutations in *ARID1A*, which encodes a subunit of the SWI/SNF chromatin-remodeling complex, are found in over half of ovarian clear cell carcinoma cases and more broadly across most types of cancers. To identify *ARID1A*-dependent changes in intracellular signaling pathways, we performed proteome analyses of isogenic ovarian clear cell carcinoma cell lines with or without *ARID1A* expression. Knockout of *ARID1A* in an ovarian clear cell carcinoma cell line with wild-type *ARID1A*, OVCA429, primarily resulted in downregulation of the mevalonate pathway, an important metabolic pathway involved in isoprenoid synthesis, cholesterol synthesis, and other downstream pathways. In a complementary experiment, expression of wild-type *ARID1A* in an ovarian clear cell carcinoma cell line containing mutated *ARID1A*, OVISE, affected the mevalonate pathway in a reciprocal manner. A striking aspect of these analyses was that, although only 5% of the detected proteome showed significant abundance changes, most proteins in the mevalonate pathway were coordinately affected by *ARID1A* status. There were generally corresponding changes when comparing the proteomics data to our previously published microarray data for ectopic expression of *ARID1A* in the OVISE cell line. However, *ARID1A*-dependent changes were not detected for genes within the mevalonate pathway. This discrepancy suggests that the

mevalonate pathway is not regulated directly by *ARID1A*-mediated transcription and may be regulated post-transcriptionally. We conclude that *ARID1A* status indirectly influences the mevalonate pathway and probably influences other processes including glycogen metabolism and 14-3-3-mediated signaling. Further, our findings demonstrate that changes in mRNA levels are sometimes poor indicators of signaling pathways affected by gene manipulations in cancer cells. *Molecular & Cellular Proteomics* 15: 10.1074/mcp.M116.062539, 3348–3360, 2016.

Epithelial ovarian cancer is the fifth highest cause of cancer mortality among women in the United States with an estimated death toll of 14,240 in 2016 (1). Ovarian clear cell carcinomas (OCCC)¹ accounts for an average of 5% of all epithelial ovarian cancers with occurrence rates greater than 20% in certain Asian populations (2). OCCC tumors are refractory to standard treatment regimens and carry a poor prognosis at late-stage detection (3).

OCCC cells have a distinct phenotype compared with other ovarian cancer histological subtypes that includes cytosolic glycogen stores that give them their characteristic “clear” appearance and a unique gene signature (4, 5). Loss-of-function mutations in *ARID1A* are found in the majority of OCCCs and one-third of ovarian endometrioid carcinomas (6, 7). *ARID1A* mutations are also prevalent in cancers of other tissues, including subtypes of breast and gastric cancers (8–10). *ARID1A* encodes the ARID1A/BAF250a subunit of the switch/sucrose nonfermentable (SWI/SNF) chromatin remodeling complex, an epigenetic regulator that modulates gene expression and DNA repair through nucleosome repositioning (11, 12). *ARID1A* may confer sequence specificity to the SWI/SNF complex through its inherent DNA-binding activity (13–

From the [‡]Molecular and Cellular Oncogenesis Program, The Wistar Institute, 3601 Spruce St., Philadelphia, Pennsylvania 19104; [§]Gene Expression and Regulation Program, The Wistar Institute, 3601 Spruce St., Philadelphia, Pennsylvania 19104; [¶]Tumor Microenvironment and Metastasis Program, The Wistar Institute, 3601 Spruce St., Philadelphia, Pennsylvania 19104; ^{||}The Center for Systems and Computational Biology, The Wistar Institute, 3601 Spruce St., Philadelphia, Pennsylvania 19104

^{**} Author's Choice—Final version free via Creative Commons CC-BY license.

Received July 18, 2016, and in revised form, September 14, 2016
 Published, MCP Papers in Press, September 21, 2016, DOI 10.1074/mcp.M116.062539

Author contributions: A.R.G., B.G.B., R.Z., and D.W.S. designed research; A.R.G. and B.G.B. performed research; B.G.B. contributed new reagents or analytic tools; A.R.G., Z.S., J.R.C., R.Z., and D.W.S. analyzed data; A.R.G. and D.W.S. wrote the paper.

¹ The abbreviations used are: OCCC, Ovarian clear cell carcinoma; BAF, BRG1-associated factor; CRISPR, Clustered regularly interspaced short palindromic repeats; FDR, False discovery rate; LFQ, Label-free quantitation; PBAF, Polybromo BRG1-associated factor; PTM, Post-translational modification; SREBP, Sterol regulatory element-binding protein; SWI/SNF, Switch/sucrose non-fermentable.

16) or via interactions with transcriptional regulators such as p53 (17). Mutation of *ARID1A* likely drives cancer progression by increasing dependence on alternative SWI/SNF complexes (18) with partially overlapping transcriptional profiles (19).

Efforts to uncover the global effects of *ARID1A* mutation in OCCC have thus focused primarily on changes in gene transcription. Mutation of *ARID1A* in OCCC cells results in loss of homeostasis between SWI/SNF and polycomb repressive complex 2, which were found to antagonistically regulate gene expression (20). Restoring wild-type *ARID1A* or inhibiting EZH2, the catalytic subunit of polycomb repressive complex 2, inhibits Phosphoinositide 3-kinase/AKT signaling through up-regulation of *PIK3IP1*. Profiling of *ARID1A*-deficient tumors in a mouse ovarian cancer model showed enrichment for genes associated with mesenchymal-epithelial transition (21). However, steady-state protein levels are often better indicators of the functional state of the cell (22). A preliminary targeted proteome approach using reverse phase protein arrays profiled 31 OCCC tumor samples with known *ARID1A* mutational status using 116 antibodies and observed differential expression of pAKT-Thr³⁰⁸ (23).

To better understand the global impact of *ARID1A* loss-of-function mutations on intracellular signaling networks, we assessed changes in the proteome because of *ARID1A* knockout in a *ARID1A* wild-type OCCC cell line, OVCA429, in an unbiased, in-depth manner using high resolution LC-MS/MS. Although *ARID1A* status had a minimal impact on the proteome overall, extensive effects on specific metabolic signaling pathways were observed. Notably, enzymes that function in the mevalonate pathway, which is involved in critical processes such as cholesterol biosynthesis and protein prenylation, showed decreased protein levels when *ARID1A* expression was abrogated. We validated these findings in an *ARID1A*-mutated OCCC cell line, OVISE, and observed an increase in abundance of mevalonate pathway enzymes when wild-type *ARID1A* was ectopically expressed. These results suggest a potential mechanism by which *ARID1A* mutation may contribute to OCCC progression and provides potential therapeutic targets for this difficult to treat form of ovarian cancer.

EXPERIMENTAL PROCEDURES

Cell Lines and 2D Culturing Conditions—OVCA429 parental cell line and OVISE cell line expressing tetracycline-inducible wild-type *ARID1A* (17) originated from the laboratory of I. M. Shih. Cell lines were tested for mycoplasma (University of Pennsylvania School of Medicine, Cell Culture Services). All cell lines were cultured on polystyrene in a 2D format in the presence of 5% CO₂ at 37 °C. OVCA429 cells were maintained in RPMI 1640 (Corning, Corning, NY, cat. no. 10-092-CM) supplemented with 10% heat-inactivated FCS (Sigma-Aldrich, St. Louis, MO, cat. no. F4135). OVISE cells expressing inducible wild-type *ARID1A* were maintained in RPMI 1640 supplemented with 10% Tet System Approved FCS (Clontech, Mountain View, CA, cat. no. 631107). All media contained 1% Penicillin-Streptomycin (Corning, cat. no. 30-002-CI).

Construction and Use of CRISPR Plasmids—The Clustered regularly interspaced short palindromic repeats (CRISPR) plasmids were a kind gift from Dr. Cigall Kadoch. *ARID1A*-specific guide RNA (gRNA; 5'-CGGGTTCGCCAGGCTGCTGGcgg-3') was inserted into pSpCas9 (BB)-2A-Puro (PX459) using *AgeI/EcoRI*. Fugene6 (Promega, Madison, WI) transfection agent was utilized to transfect plasmid into cells and cells were selected with puromycin (1 μg/ml). Isogenic clonal cell lines were established based on *ARID1A* protein expression.

Immunoblotting—Primary antibodies used for immunoblots were *ARID1A* (Cell Signaling, Beverly, MA, cat. no. 12354) at 1:1000 and *GAPDH* (Sigma-Aldrich, cat. no. G8795) at 1:100,000. Secondary antibodies were Peroxidase-conjugated Anti-Rabbit IgG (Sigma-Aldrich, cat. no. A0545) at 1:50,000 and Peroxidase-conjugated Anti-Mouse IgG (Sigma-Aldrich, cat. no. A9044) at 1:50,000.

Sample Preparation for Proteome Analysis—OVCA429 cells were plated on tissue culture-treated Petri dishes (Corning, cat. no. 430599). OVISE cells harboring inducible wild-type *ARID1A* were similarly plated and treated with either DMSO or 1 μg/ml doxycycline in DMSO to induce *ARID1A* expression (96 h treatment duration, refreshed after 72 h). Cells were harvested at 60–80% confluency. Cell lysates were prepared using SDS lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% SDS, 1 mM EDTA) with freshly added 1 mM DTT and protease inhibitors (150 μM PMSF, 1 μg/ml Pepstatin A, and 1 μg/ml Leupeptin). Lysates were sonicated briefly to shear genomic DNA. Protein content of clarified lysates was quantified by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Boston, MA, cat. no. 23227). For proteome analyses, 12.5 to 25 μg lysate was separated on 10% Bis-Tris NuPAGE minigels (Thermo Fisher Scientific, cat. no. NP0301) with MES running buffer (Thermo Fisher Scientific, cat. no. NP0002) until the dye front migrated 0.5 cm. In-gel trypsin digestion was performed as described previously (24).

Proteome Analysis—LC-MS/MS of tryptic peptides was performed using a nanoACQUITY UPLC (Waters, Milford, MA) in-line with a Q Exactive Plus or Q Exactive HF mass spectrometer (Thermo Fisher Scientific). Typically, 0.25–1 μg of each tryptic digest, estimated assuming a 50% recovery of the amount of protein loaded into the gel lane, was loaded onto a 180 μm x 20 mm nanoACQUITY UPLC Symmetry C18 trap column with 5 μm particle size and 100 angstrom pore size (Waters, cat. no. 186006527) with 0.1% formic acid in Milli-Q water (solvent A). Analytical separation was performed on a 1.7 μm x 250 mm nanoACQUITY UPLC Peptide BEH C18 column with 1.7 μm particle size and 130 angstrom pore size (Waters, cat. no. 186003546) using a 245 min gradient with 0.1% formic acid in acetonitrile (solvent B) as follows: 5–30% B over 225 min, 30–80% B over 5 min, and constant 80% B for 15 min. A blank was run between each sample to minimize carryover by injecting water and using a 30 min gradient with the same solvents. Peptides were analyzed using the same parameters for the Q Exactive Plus and HF instruments unless otherwise indicated. Full MS spectra were recorded at a resolution of 70,000 for the Plus and 60,000 for the HF with a scan range of 400–2000 *m/z* in profile mode. Full MS automatic gain control target and maximum injection time were set to 3e6 and 50 ms, respectively. MS2 spectra were recorded at 17,500 resolution for the Plus and 15,000 resolution for the HF. MS2 automatic gain control target and maximum injection time were set to 5e4 and 50 ms, respectively. Data-dependent analysis was performed on the 20 most abundant ions using an isolation width of 1.5 *m/z* and an underfill of 1%, corresponding to a minimum threshold of 1e4. Peptide match was set to preferred, and unassigned and singly charged ions were rejected. Dynamic exclusion was set to 30 s for the Plus and 45 s for the HF.

Experimental Design and Statistical Rationale—All experimental and control samples were analyzed using biological triplicates to

allow for robust statistics when performing label-free quantitative comparisons. This is more advantageous than technical replicates as it accounts for variability between cell culture plates and at all stages of sample preparation and processing. The OVCA429 and OVISe proteome comparisons each generated 6 RAW files, 3 corresponding to experimental conditions and three corresponding to control, that contain all acquired Full MS and MS2 spectra. Base peak chromatograms were inspected visually in Xcalibur Qual Browser version 3.0.63 (Thermo Fisher Scientific). RAW files were processed by MaxQuant version 1.5.1.2 using default parameters unless otherwise specified (<http://www.maxquant.org>) (25). All RAW files for a given parental cell line were analyzed together in a single MaxQuant run. Database searches were performed using the Andromeda search engine included with the MaxQuant release (26) with the UniProt human sequence database (July 28, 2014; 145,433 sequences; 53,453,851 residues) and an in-house contaminants database of common laboratory contaminants, including keratins, bovine proteins detected in FCS, trypsin, and mycoplasma proteins to detect potential mycoplasma contamination of cell cultures (Sequences obtained from UniProt, July 28, 2014; 3671 sequences; 1,338,375 residues). Precursor mass tolerance was set to 4.5 ppm in the main search, and fragment mass tolerance was set to 20 ppm. Digestion enzyme specificity was set to Trypsin/P with a maximum of 2 missed cleavages. A minimum peptide length of 7 residues was required for identification. Up to 5 modifications per peptide were allowed; acetylation (protein N-terminal) and oxidation (Met) were set as variable modifications, and carbamidomethyl (Cys) was set as a fixed modification. No Andromeda score threshold was set for unmodified peptides. A minimum Andromeda score of 40 was required for modified peptides. Peptide and protein false discovery rates (FDR) were both set to 1% based off a target-decoy reverse database. Proteins that shared all identified peptides were combined into a single protein group. If all identified peptides from one protein were a subset of identified peptides from another protein, these proteins were also combined into that group. Peptides that matched multiple protein groups (“razor” peptides) were assigned to the protein group with the most unique peptides. “Match between runs” based on accurate *m/z* and retention time was enabled with a 0.7 min match time window and 20 min alignment time window. Label-free quantitation (LFQ) was performed using the MaxLFQ algorithm built into MaxQuant (27). Briefly, peaks were detected in Full MS, and a three-dimensional peak was constructed as a function of peak centroid *m/z* (7.5 ppm threshold) and peak area over time. Following de-isotoping, peptide intensities were determined by extracted ion chromatograms based on the peak area at the retention time with the maximum peak height. Peptide intensities were normalized to minimize overall proteome difference based on the assumption that most peptides do not change in intensity between samples. Protein LFQ intensities were calculated from the median of pairwise intensity ratios of peptides identified in two or more samples and adjusted to the cumulative intensity across samples. Quantification was performed using razor and unique peptides, including those modified by acetylation (protein N-terminal) and oxidation (Met). A minimum peptide ratio of 1 was required for protein intensity normalization, and “Fast LFQ” was enabled.

Data processing and cluster analysis was performed using Perseus version 1.5.0.31 (<http://www.perseus-framework.org>) (28), and statistical analysis was performed using Excel 2013 and 2016 (Microsoft, Redmond, WA). Contaminants and protein groups identified by a single peptide were filtered from the data set. FDR was calculated as the percentage of reverse database matches out of total forward and reverse matches. Protein group LFQ intensities were \log_2 transformed to reduce the effect of outliers. For cluster analysis and statistical comparisons between proteomes, protein groups missing LFQ values

were assigned values using imputation. Missing values were assumed to be biased toward low abundance proteins that were below the MS detection limit, referred to as “missing not at random” (29), an assumption that is frequently made in proteomics studies (28). The missing values were replaced with random values taken from a median downshifted Gaussian distribution to simulate low abundance LFQ values (demonstrated in supplemental Fig. S4). Imputation was performed separately for each sample from a distribution with a width of 0.3 and downshift of 1.8. Hierarchical clustering was performed on Z-score normalized, \log_2 LFQ intensities using Euclidean distance and average linkage with k-means preprocessing (300 clusters). Log ratios were calculated as the difference in \log_2 LFQ intensity averages between experimental and control groups. Two-tailed, unpaired, homoscedastic Student’s *t* test calculations were used in statistical tests as histograms of LFQ intensities showed that all data sets approximated normal distributions. $p < 0.05$ was considered statistically significant. Base 10-fold-change values for ratios < 1 are represented as negative reciprocals of the ratios.

Data Availability—The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the MassIVE partner repository (<http://massive.ucsd.edu/ProteoSAFe/static/massive.jsp>) with the data set identifier PXD004570.

Comparison of Proteomics and Microarray Data Sets—Microarray data for OVISe cells with and without *ARID1A* induction were obtained from the Gene Expression Omnibus database, accession number GSE54979 (20). Protein groups from the proteomics analysis were mapped to Illumina microarray probes based on Gene Name and Ensembl Gene Accession Number. When duplicate probes were present for a single gene, the probe with the most significant *p* value was selected.

Canonical Pathway Analysis—Pathway analysis was performed using Qiagen’s (Valencia, CA) Ingenuity Pathway Analysis (IPA) using default parameters unless otherwise specified (<http://www.ingenuity.com/pa/>). Tab-delimited files containing matrices of identified proteins and associated log ratios and *t* test *p* values were uploaded into the IPA service as “Ingenuity File Format A or B.” Proteins were mapped from UniProt gene names to “IPA IDs” based on “Gene symbol - human.” Unmapped proteins were reconciled based on alternative gene names if possible. Core Analyses were performed using a log ratio threshold of 1 and *p* value threshold of 0.05. Proteins that increased or decreased in abundance were considered. The entire imported data set for a given comparison was used as a reference set in statistical calculations.

RESULTS

Generation of Isogenic *ARID1A* Knockout and *ARID1A* Wild-type OCCC Cell Lines for Proteome Analysis—To assess the impact of *ARID1A* frameshift or nonsense mutations typically found in OCCC (6, 7), we performed proteome comparisons between cells that express wild-type *ARID1A* and cells that have lost functional *ARID1A* expression through mutation. However, distinct *ARID1A* wild-type and *ARID1A*-mutated OCCC cell lines have genetic and epigenetic differences that may obfuscate the contribution of *ARID1A* mutational status to proteome differences (20). To address this concern, we utilized CRISPR gene editing to knockout *ARID1A* expression in the OCCC cell line OVCA429 (30), which expresses wild-type *ARID1A* protein at a level representative of the majority of commonly used *ARID1A* wild-type OCCC cell lines (17). *ARID1A* knockout resulted in the loss of detectable

ARID1A by immunoblot (supplemental Fig. S1). There were no significant phenotypic effects on cell morphology or cell proliferation (data not shown).

Proteome Analysis of *ARID1A* Knockout in OVCA429—We performed proteome analysis to elucidate the effects of ARID1A loss in OVCA429 using total cell lysates in biological triplicates from ARID1A knockout and control cells. LC-MS/MS was performed on a Thermo Q Exactive Plus mass spectrometer with an unfractionated, 4-hour run per sample. Data from biological triplicates of knockout and control cells were analyzed together in MaxQuant using a protein and peptide FDR of 1% and “matching between runs” based on accurate mass and retention time. A total of 54,037 tryptic peptides were identified when acetylation (protein N-terminal) and oxidation (Met) were counted separately from unmodified versions of the 50,365 sequence unique peptides that were identified across all runs (supplemental Table S1). These peptides were assigned to 5,681 distinct proteins groups. Single peptide hits were filtered from the data set to yield a list of 4,973 high confidence protein groups identified at 0.52% FDR (supplemental Table S2). These protein groups were the focus of all analyses. An average of 10 peptides were identified per protein group, yielding an average sequence coverage of 29%. The depth of proteome analysis was estimated to be 50% based on the expected approximate 10,000 proteins expressed by a cancer cell line at a given time (31).

Label-free Comparisons Between *ARID1A* Knockout and Control Proteomes—We used label-free quantitation (LFQ) to compare relative abundances of protein groups across the ARID1A knockout and control proteomes. The MaxLFQ algorithm built into the MaxQuant utilizes maximal ratios of MS signal intensities for peptides identified in multiple samples coupled with delayed normalization to allow for accurate comparisons across proteomes (27). The overall effect of loss of ARID1A protein expression was assessed using a global analysis of the biological triplicate knockout and control proteomes. Unsupervised hierarchical clustering based on LFQ intensities resulted in independently clustered knockout and control proteomes, indicating that ARID1A status is the main determinant of differences between the proteomes (Fig. 1A). We also compared different passages for each cell line based on average LFQ intensities of protein groups from biological triplicate samples. As has been observed in prior studies, low abundance proteins had a wider spread of signal intensities due to less accurate quantitation near the noise threshold (32). However, there was excellent overall reproducibility between the different culture passage numbers (R^2 of 0.9803 and 0.9873 for control and knockout cells, respectively), which demonstrates that independent sample preparation had minimal effects on the proteomes (supplemental Fig. S2).

***ARID1A* Knockout Has a Low Overall Impact on the Proteome**—We examined the specific proteins affected by ARID1A knockout through quantitative proteome comparisons. Protein groups that significantly changed in abundance

as a function of ARID1A status were selected based on magnitude of changes between mean LFQ intensities of the biological replicates as well as the statistical significance of the changes to increase precision because fold-change alone is sensitive to outliers (27). A minimum fold-change of 2 in either direction was required based on the rationalization that smaller changes are less likely to exert a biological effect. A Student's t test p value of less than 0.05 was used as a statistical cutoff. When comparing the ARID1A knockout and control proteomes, 430 and 2606 proteins satisfied the fold-change and statistical criteria, respectively (Table S3). Combining these relatively relaxed stringency criteria selected only 264 proteins, corresponding to 5% of the identified proteins, as significantly changed in level in the knockout relative to the control (Fig. 1B, supplemental Table S3). Of these, 95 proteins increased in abundance and 169 proteins decreased in abundance. ARID1A protein was undetectable in the ARID1A wild-type control proteome, likely due to its low abundance. Other components of the BRG1-associated factor (BAF) SWI/SNF complex (33), including SMARCC1/2, SMARCD1/2, DPF2, SMARCE1, SMARCA4, and SMARCB1, were detected but did not exhibit ARID1A-dependent changes. ARID2, a subunit exclusive to the polybromo BRG1-associated factor (PBAF) SWI/SNF complex, significantly increased in level in the knockout, suggesting possible compensation for ARID1A loss via an alternative SWI/SNF complex with an overlapping transcriptional profile (19).

***ARID1A* Knockout Primarily Affects Enzymes Involved in the Mevalonate Pathway**—To uncover signaling pathways that are perturbed when ARID1A expression is lost, we performed canonical pathway analysis (IPA) on the set of 264 protein groups that significantly changed between ARID1A knockout and control, of which 257 were curated in the IPA knowledgebase. Interestingly, a number of metabolic pathways were significantly overrepresented in the ARID1A-dependent proteins compared with all identified proteins (Fig. 2). Four of the 11 highest scoring canonical pathways were related to the mevalonate pathway, which synthesizes a variety of isoprenoids derived from acetyl-CoA that are required for diverse cellular processes, including cholesterol, prenyl groups (geranylgeranyl pyrophosphate and farnesyl pyrophosphate), dolichol, ubiquinone, and heme A (34, 35). Associated pathways included: “Superpathway of Geranylgeranyl-diphosphate Biosynthesis I (via Mevalonate),” “Mevalonate Pathway I,” “Superpathway of Cholesterol Biosynthesis,” and “Trans, trans-farnesyl Diphosphate Biosynthesis.” Coverage of these pathways was extensive with 27.3% to 100% of identified pathway components significantly changing in the knockout proteome (Fig. 2). In total, 6 enzymes in the mevalonate pathway (ACAT2, HMGCS1, MVK, MVD, IDI1, and FDPS) decreased in abundance in the ARID1A knockout (Table I). Proteins relevant to the mevalonate biosynthesis pathway (listed in Table I) were identified in all knockout and control proteomes or not identified in any sample; therefore, imputation

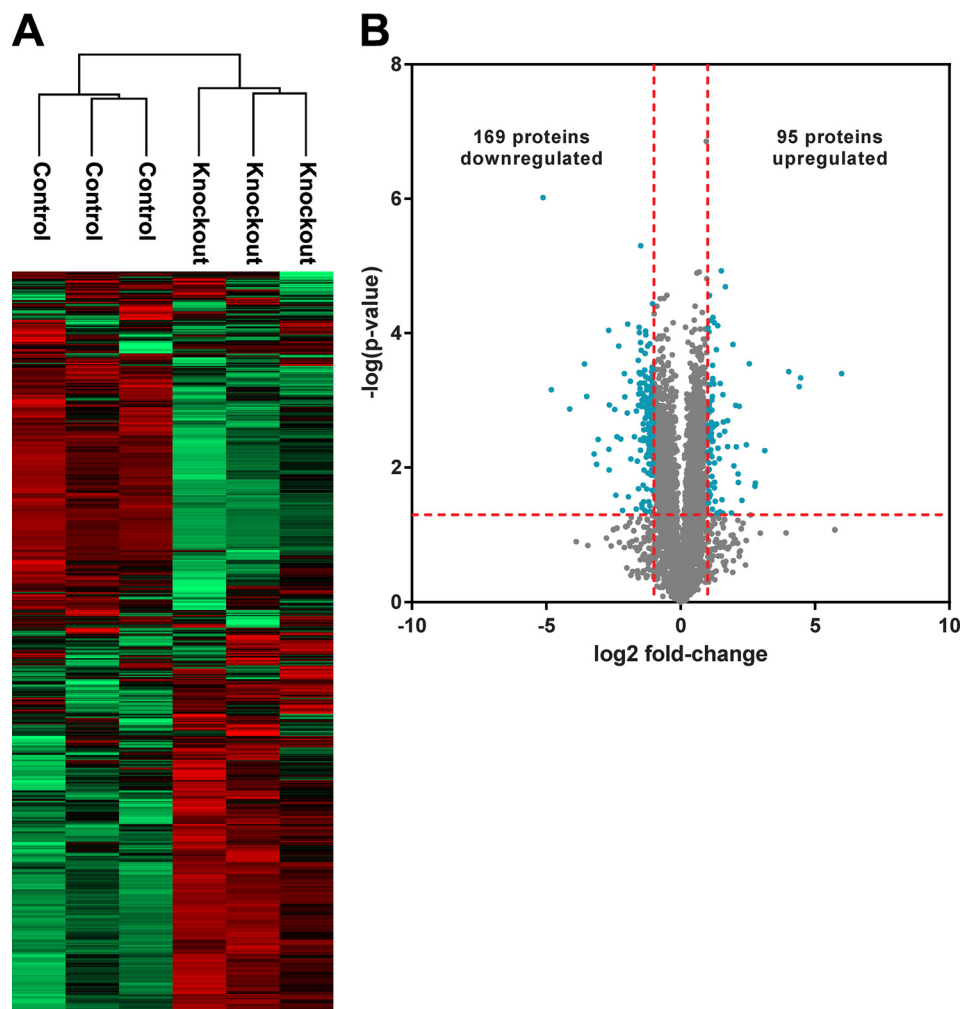


FIG. 1. Characterization of the OVCA429 proteome with or without *ARID1A* CRISPR knockout. *A*, Two-way, unsupervised hierarchical clustering of OVCA429 proteomes with and without *ARID1A* knockout based on LFQ intensities. *B*, Volcano plot comparing *ARID1A* knockout and control proteomes. Log ratios of LFQ intensities in *ARID1A* knockout versus control were plotted against negative log p values from the Student's t test based on biological triplicates. Vertical lines: fold-changes of ± 2 . Horizontal line: Student's t test p of 0.05. Blue points: proteins that meet both criteria for significant change between *ARID1A* knockout and control (i.e. fold-change in abundance > 2 and $p < 0.05$). Gray points: proteins that do not meet both of these criteria.

of missing values did not affect the calculated fold-changes and p values (supplemental Tables S2 and S3). Pathways associated with glycogen and glucose metabolism were also among the top scoring pathways, including “Glycogen Degradation II,” “Glycogen Degradation III,” “Glucose and Glucose-1-phosphate Degradation,” and “GDP-glucose Biosynthesis.” Within these pathways there was a significant decrease in abundance of phosphoglucomutase (PGM1, PGM2, and PGM3) and phosphorylase (PYGL) which should reduce glycogen degradation (Fig. 2). As noted above, the clear cell morphology is because of excessive glycogen accumulation in the cytosol. Another high scoring pathway was “14-3-3-mediated signaling.” 14-3-3 proteins interact with target proteins to regulate a diverse array of signaling pathways (36). Of the 7 human isoforms (β , γ , ϵ , η , σ , τ , and ζ), 14-3-3 σ (SFN) and 14-3-3 τ (YWHAQ) were significantly

decreased in the *ARID1A* knockout relative to control (supplemental Table S3). Loss of 14-3-3 σ in particular frequently occurs in cancer and is associated with metabolic changes that promote cancer progression (36, 37). Similar results were obtained from analysis of a second *ARID1A* CRISPR knockout clone (supplemental Fig. S1, data not shown).

*A Complementary System for Assessing the Proteome Impact of *ARID1A* Perturbation*—We performed a complementary proteomics analysis to validate the effects of *ARID1A* knockout observed in OVCA429 in an alternative biological context. Wild-type *ARID1A* was reintroduced into the *ARID1A*-mutated OCCC cell line OVISe (38), which completely lacks expression of *ARID1A* protein (17), using a tetracycline-inducible construct (supplemental Fig. S3). This experimental approach was recently used to study the transcriptional effects of ectopic expression of wild-type *ARID1A* in OVISe cells (20).

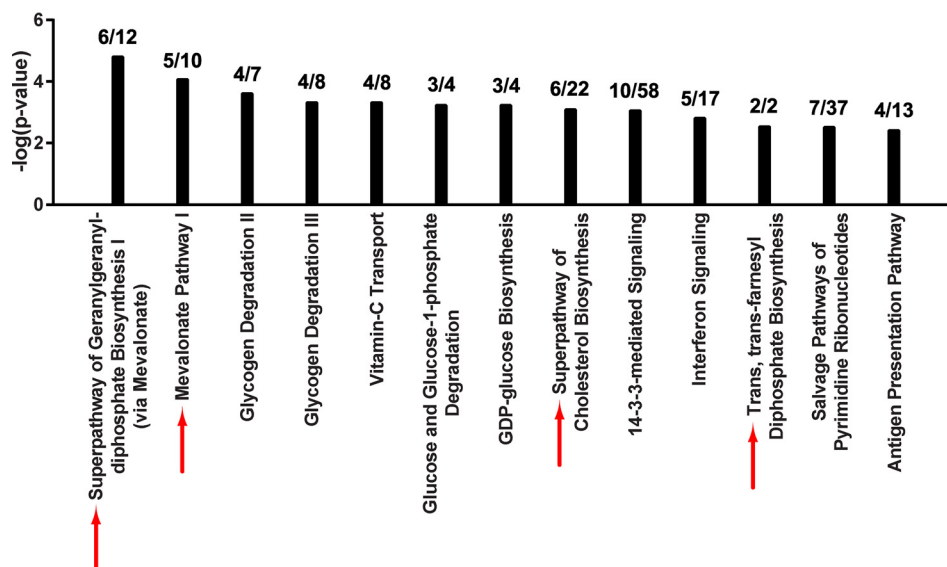


FIG. 2. **Knockout of *ARID1A* downregulates the mevalonate pathway.** Top scoring canonical pathways associated with *ARID1A* knockout in OVCA429 are shown (IPA, $p < 0.005$ by Fisher's exact test right-tailed). *Black bars*: negative log p values for each canonical pathway. The ratio of annotated proteins that significantly changed in level in *ARID1A* knockout to total identified proteins for a given canonical pathway is shown above each bar. *Red arrows*: canonical pathways that include the mevalonate pathway and are also significantly enriched among proteins that changed upon *ARID1A* induction. Refer to Fig. 4.

Proteome Analysis of *ARID1A* Induction in OVISE—We analyzed cell lysates from biological triplicates of *ARID1A*-induced and control (mutated *ARID1A*) cells by LC-MS/MS on a Thermo Q Exactive HF mass spectrometer using single unfractionated, 4-hour runs. All samples were analyzed together in MaxQuant using a 1% FDR for peptides and proteins, as described above. A total of 66,829 tryptic peptides were identified when acetylation (protein N-terminal) and oxidation (Met) were counted separately from unmodified versions of the 64,245 sequence unique peptides that were identified across all runs (supplemental Table S4). These peptides were assigned to a total of 6,717 protein groups. We further analyzed the set of 5785 protein groups identified by more than a single peptide, which corresponded to an FDR of 0.28% (supplemental Table S5). There was an average of 11 peptides identified per protein group, yielding an average sequence coverage of 30%.

Label-free Comparisons Between *ARID1A*-induced and Control Proteomes—We performed label-free comparisons between *ARID1A*-induced and control proteomes. Unsupervised hierarchical clustering resulted in independent clustering based on wild-type *ARID1A* expression, indicating that *ARID1A* status had a global effect on the proteomes (Fig. 3A). The percentage of protein groups that showed *ARID1A* dependence compared with all identified protein groups was similar in magnitude to the changes observed for the *ARID1A* knockout. 387 protein groups significantly changed between conditions, corresponding to 7% of the identified proteome (Fig. 3B, supplemental Table S6). The cells with induced wild-type *ARID1A* showed 191 protein groups at higher levels and 196 protein groups at lower levels relative to the control

cells. Given that in this system *ARID1A* is overexpressed relative to a null background, it is unsurprising that the *ARID1A* protein had the greatest magnitude fold-change among the significantly changed proteins with a 223-fold increase in protein level. In contrast to the *ARID1A* knockout experiment, the core SWI/SNF complex subunits SMARCB1, SMARCC1, and DPF2 (33) significantly increased in abundance after *ARID1A* induction. *ARID1B*, a component of BAF SWI/SNF complexes that is mutually exclusive with *ARID1A*, and PBRM1, a subunit of the PBAF SWI/SNF complex, both significantly decreased in abundance in response to increased *ARID1A* levels.

***ARID1A* Induction Up-regulates the Mevalonate Pathway**—Pathway analysis was performed on the 378 of the 387 *ARID1A*-dependent proteins that could be mapped to the IPA database. Seven of the top 10 scoring pathways were associated with the mevalonate pathway: “Superpathway of Geranylgeranyl-diphosphate Biosynthesis I (via Mevalonate),” “Mevalonate Pathway I,” “Superpathway of Cholesterol Biosynthesis,” “Cholesterol Biosynthesis I,” “Cholesterol Biosynthesis II (via 24,25-dihydrolanosterol),” “Cholesterol Biosynthesis III (via Desmosterol),” and “Epoxy-squalene Biosynthesis” (Fig. 4). There was extensive coverage of pathway components with 38.5% to 100% of identified pathway proteins significantly changing in abundance, similar to the *ARID1A* knockout, although the change in protein abundance was in the opposite direction. Specifically, the 10 enzymes that changed among these pathways (ACAT2, HMGCS1, MVK, IDI1, FDPS, FDFT1, SQLE1, LSS1, TM7SF2, and MSMO1) increased in abundance when *ARID1A* was induced (Table I). All mevalonate pathway-related proteins except

Effect of *ARID1A* Loss in Ovarian Clear Cell Carcinoma

TABLE I
Mevalonate pathway enzymes in proteomics and microarray datasets

Changes in abundance of enzymes that participate in mevalonate pathway and related downstream pathways. Inclusion in table based on IPA curation. Refer to supplemental Tables S3 and S6 for protein quantitation. For proteome analyses, significant changes were defined as fold-change > 2 and Student's *t*-test $p < 0.05$. For enzymes in the proteome analysis that significantly changed in one condition, a slightly relaxed fold-change was allowed in the other condition (i.e. fold-change of 1.8 with Student's *t*-test $p < 0.05$). For the microarray analysis, significant changes were defined as $p < 0.05$. Red and blue background: significantly higher or lower protein/gene level after *ARID1A* knockout/induction compared to control. Grey background: no significant change in level after *ARID1A* knockout/induction compared to control. N.D. indicates not detected in a given analysis. Significant p values ($p < 0.05$) are bolded.

Gene name	Protein name	OVCA429 <i>ARID1A</i> knockout vs. control (proteome)		OVISE <i>ARID1A</i> induction vs. control (proteome)		OVISE <i>ARID1A</i> induction vs. control (microarray) ^a		# of PTMs ^b		
		Fold-change	P-value	Fold-change	P-value	Fold-change	P-value	Ub	Ac	Phos
ACAT2	Acetyl-CoA acetyltransferase, cytosolic	-2.68	<0.001	2.50	<0.001	-1.18	0.0413	1	3	2
HMGCS1	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	-6.41	<0.001	4.20	<0.001	1.07	0.3854	13	4	5
HMGCR	3-hydroxy-3-methylglutaryl-coenzyme A reductase	-1.25	0.0521	2.35	0.0580	-1.06	0.4778	21	0	1
MVK	Mevalonate kinase	-4.24	<0.001	3.34	<0.001	1.13	0.1116	0	0	0
PMVK	Phosphomevalonate kinase	-1.72	0.0293	1.20	0.0600	-1.02	0.7635	4	0	1
MVD	Diphosphomevalonate decarboxylase	-4.28	<0.001	1.94	<0.001	1.05	0.4947	2	0	7
IDI1	Isopentenyl-diphosphate Delta-isomerase 1	-2.44	<0.001	2.79	<0.001	-1.01	0.8763	6	1	1
FDPS	Farnesyl pyrophosphate synthase	-2.24	0.0010	2.15	<0.001	-1.20	0.0241	3	2	0
GGPS1	Geranylgeranyl pyrophosphate synthase	N.D.	N.D.	1.73	<0.001	1.12	0.1267	3	1	5
FDFT1	Squalene synthase	-1.31	0.0132	5.21	<0.001	-1.19	0.0497	5	0	0
SQLE	Squalene monooxygenase	-1.27	0.0719	2.19	<0.001	-1.10	0.2452	2	0	0
LSS	Lanosterol synthase	-1.01	0.8819	2.59	<0.001	1.08	0.3160	7	0	2
CYP51A1	Lanosterol 14-alpha demethylase	-1.05	0.6547	1.32	0.0018	-1.23	0.0185	14	0	1
DHCR24	Delta(24)-sterol reductase	1.06	0.5957	1.83	<0.001	1.11	0.1519	7	0	2
TM7SF2	Delta(14)-sterol reductase	N.D.	N.D.	2.92	0.0028	1.08	0.2978	0	0	0
MSMO1	Methylsterol monooxygenase 1	-1.10	0.1488	3.19	<0.001	-1.10	0.2421	4	0	0
NSDHL	Sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating	1.18	0.1521	1.29	0.0015	1.12	0.1296	3	0	0
HSD17B7	3-keto-steroid reductase	-1.00	0.9977	1.84	0.0033	-1.16	0.0484	6	1	0
EBP	3-beta-hydroxysteroid-Delta(8),Delta(7)-isomerase	1.28	0.0415	N.D.	N.D.	-1.13	0.1108	1	0	0
SC5D	Lathosterol oxidase	N.D.	N.D.	N.D.	N.D.	1.10	0.2194	3	0	1
DHCR7	7-dehydrocholesterol reductase	1.13	0.1615	1.92	<0.001	1.02	0.7912	5	0	2

^a Microarray data were obtained from the Gene Expression Omnibus database, accession number GSE5497 (20).

^b Ubiquitination (Ub), acetylation (Ac), and phosphorylation (Phos) post-translational modifications (PTMs) from Sharpe and Brown, 2013 (52).

HMGCR were identified in all OVISE proteomes or not detected in any sample. HMGCR was detected in a single control proteome (Table S5), therefore two imputed values were used in fold-change and p value calculations (supplemental Table S6).

Comparison of OVISE Cell Proteomics and Microarray Data—To further explore the basis for the above changes in enzyme levels, we compared the proteomics data to microarray data from OVISE cells ectopically expressing wild-type

ARID1A using the same lentivirus transduced cell line (20). Most mevalonate pathway enzymes that significantly increased at the protein level upon *ARID1A* induction did not show significant changes in gene expression (Table I). For the few genes that significantly changed in level, the magnitude of change was very low. A better concordance between *ARID1A*-dependent changes in protein and gene levels was observed for proteins that exhibited the greatest changes upon *ARID1A* induction and their associated protein-coding

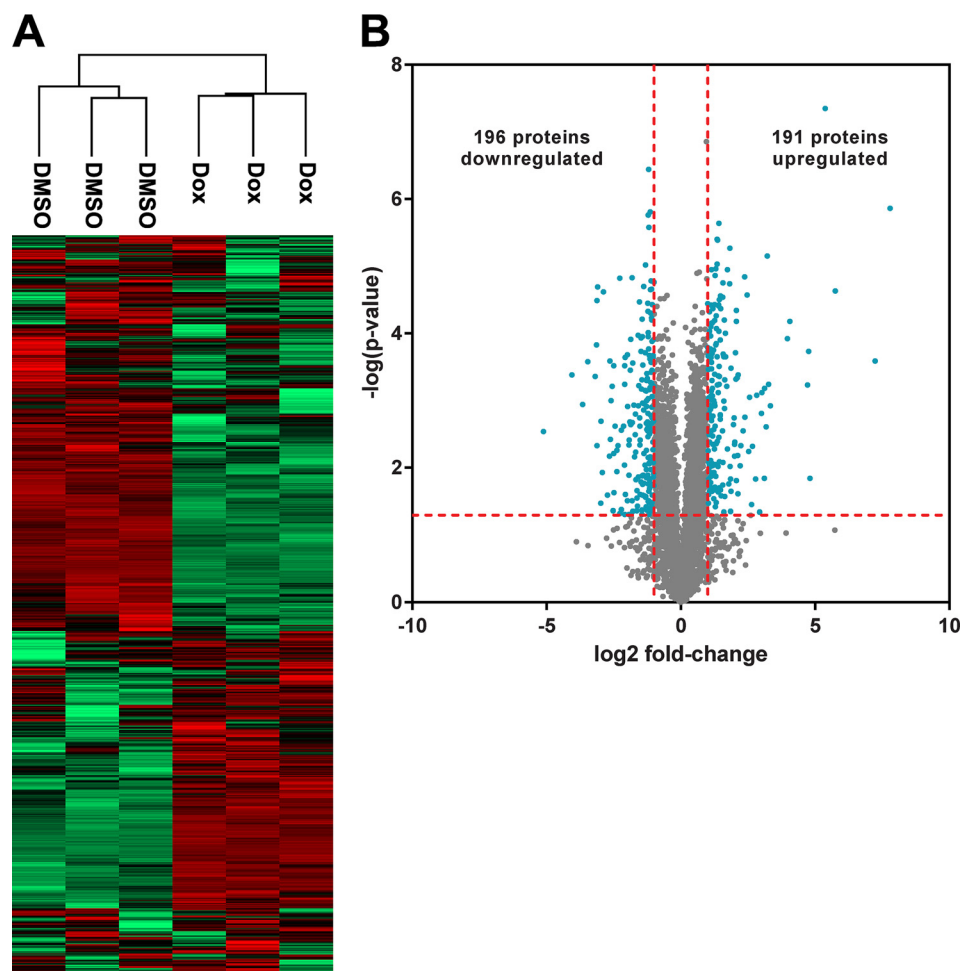


FIG. 3. Characterization of the OVISE proteome with or without induction of wild-type *ARID1A* expression. *A*, Two-way, unsupervised hierarchical clustering of OVISE proteomes based on LFQ intensities. *B*, Volcano plot comparing proteomes from *ARID1A*-induced cells and control cells. Log ratios of LFQ intensities in *ARID1A* induction versus control were plotted against negative log *p* values from the Student's *t* test based on biological triplicates. Vertical lines: fold-changes of ± 2 . Horizontal line: Student's *t* test *p* of 0.05. Blue points: proteins that meet both criteria for significant change between *ARID1A* induction and control (i.e. fold-change in abundance > 2 and $p < 0.05$). Gray points: proteins that do not meet both of these criteria.

genes (Table II). Thirteen of the top 25 entries showed consistent protein and gene regulation; however, gene expression changes were much smaller than protein level changes. In some cases, discrepancy in magnitude of change could be because of limitations of using imputed values for proteins that were not detected in either condition. When the effects of imputation were evaluated, 679 protein groups or 12% of all identified protein groups, were not quantified in at least 1 OVISE proteome (supplemental Table S5) and therefore were assigned imputed values (supplemental Table S6, supplemental Fig. S4). Eight proteins listed in Table II required two or more imputed values for quantitation because of lack of detection in either the *ARID1A*-induced or control proteomes, though none of these proteins required imputed values for both conditions (Table S7). For example, ARHGDI1 had an apparent 10-fold increase in protein abundance compared with a 1.5-fold increase in gene expression following *ARID1A*

induction; however, ARHGDI1 protein was not detected in control and, therefore, the calculated fold-change is dependent entirely upon imputed values. Without imputation, fold-change calculations for proteins such as ARHGDI1 would result in a division by zero. Approximately 4% of all identified protein groups (204 protein groups) fall into this category and would be classified as being observed in either the *ARID1A*-induced or control proteomes if imputation had not been used (supplemental Table S5). For proteins present at quantifiable levels in both *ARID1A*-induced and control proteomes, calculations performed with and without imputation resulted in comparable fold-changes and *p* values (supplemental Table S7). For example, EFNB1 was not detected in two control proteomes. With imputation, the fold-change between *ARID1A* induction versus control was 28.1 with a *p* value of 0.014. With imputation, these values were 20.0 and 0.034, respectively, which still met our criteria for significantly changed proteins.

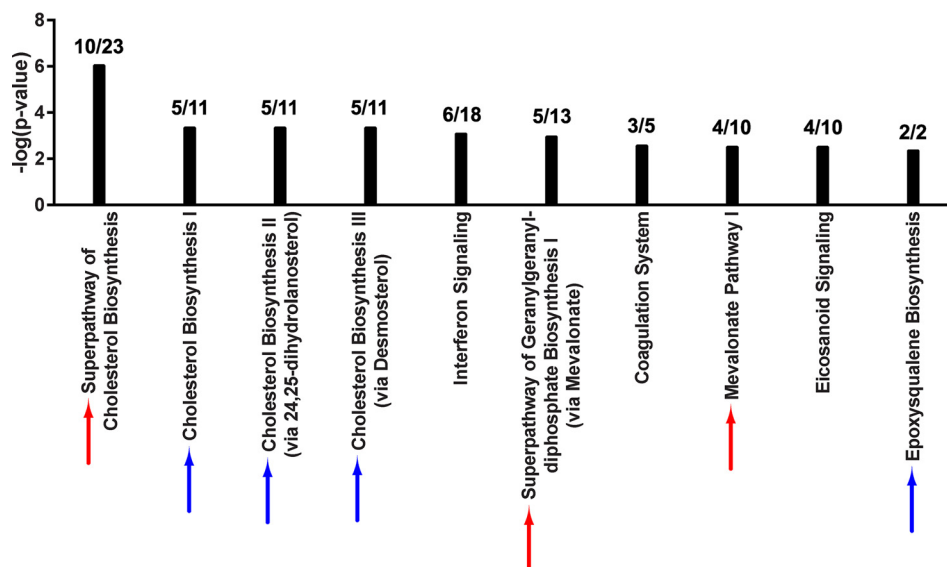


FIG. 4. **Induction of ARID1A up-regulates the mevalonate pathway.** Top scoring canonical pathways associated with ARID1A restoration in OVISe are shown (IPA, $p < 0.005$ by Fisher's exact test right-tailed). *Black bars*: negative log p values for each canonical pathway. The ratio of annotated proteins that significantly changed in level following ARID1A induction to total identified proteins for a given canonical pathway is shown above each bar. *Red arrows*: canonical pathways that involve the mevalonate pathway and are also significantly enriched among proteins that changed due to ARID1A knockout. Refer to Fig. 2. *Blue arrows*: additional canonical pathways involving the mevalonate pathway.

We also searched the proteomics data for known *bona fide* ARID1A targets. SMAD3 and PIK3IP1 (17, 20) were not detected in the proteome data set, while CDKN1A (17) increased moderately, but significantly, in abundance upon ARID1A induction in OVISe (supplemental Table S6).

DISCUSSION

Here, we show that 5–7% of the proteomes in OCCC cell lines significantly change following perturbation of ARID1A expression (Fig. 1B, Fig. 3B). Despite the relatively moderate global effect on the proteome, the mevalonate pathway and related downstream pathways were associated with both ARID1A knockout in OVCA429 and ectopic expression of wild-type ARID1A in OVISe (Fig. 2, Fig. 4). Examination of specific pathway components revealed that most enzymes in the mevalonate pathway showed abundance changes that correlated with ARID1A status (Table I). That is, most enzymes decreased in abundance in the ARID1A knockout and increased in abundance upon restoration of wild-type ARID1A. Enzymes showing these reciprocal abundance changes include FDPS, which catalyzes production of farnesyl pyrophosphate, a metabolite that is upstream of the branch point that specifies prenylation, cholesterol biosynthesis, and other pathway outputs (Fig. 5). Overexpression of ARID1A is also associated with increased abundance of enzymes in the cholesterol biosynthesis branch downstream of the mevalonate pathway (Table I, Fig. 5). These findings suggest that the mevalonate pathway and related pathways are strongly affected by ARID1A.

Dysregulation of the mevalonate pathway has previously been implicated in cancers and other diseases, mainly due to increased cholesterol biosynthesis and aberrant protein prenylation that occurs when the pathway is spuriously up-regulated (39–41). In breast cancer, mutant p53 up-regulates mevalonate pathway components at least in part through the sterol regulatory element-binding proteins (SREBP) family of transcription factors, and increased mevalonate pathway flux is necessary and sufficient for the mutant p53 phenotype observed in culture (42). Since ARID1A is a tumor suppressor (17), it was surprising that the mevalonate pathway was downregulated in OCCC when ARID1A was knocked out (Table I). Since most oppositely changed enzymes are located in the mevalonate trunk of the pathway before it branches into alternative products (Fig. 5), further study is required to determine the biological importance and relative contribution of the alternative pathway outputs to the OCCC disease phenotype. However, a recently published large Danish demographic study of the linkage between statin use and risk of ovarian cancer showed that OCCC was the only subtype with increased cancer incidence with use of statins and that risk of OCCC increased with length of use (43). This is consistent with downregulation of the mevalonate pathway having an important role in development of OCCC because statins inhibit HMGCR, which catalyzes the rate-limiting step of the pathway (34). This is also consistent with our finding that ARID1A inactivation downregulates the mevalonate pathway.

Given that ARID1A regulates gene transcription through the SWI/SNF complex (17), direct targets of ARID1A should be

TABLE II
Effects of wild-type ARID1A induction on protein and gene expression levels

The 25 proteins showing the largest magnitude, significant changes in protein abundance after ARID1A induction in OVISe cell line. Proteins unable to be mapped to at least one Illumina microarray probe were pre-filtered from the list. Refer to supplemental Table S6 for protein quantitation. For the proteome analysis, significant changes were defined as fold-change > 2 and Student's t-test $p < 0.05$. For the microarray analysis, significant changes were defined as $p < 0.05$. Red and blue background: significantly higher or lower protein/gene level after ARID1A induction. Grey background: no significant change in level after ARID1A induction. N.D. indicates not detected in a given analysis. Significant p values ($p < 0.05$) are bolded.

Gene name	Protein name	OVISe ARID1A induction vs. control (proteome)		OVISe ARID1A induction vs. control (microarray) ^a		# of PTMs ^b		
		Fold-change	P-value	Fold-change	P-value	Ub	Ac	Phos
ARID1A	AT-rich interactive domain-containing protein 1A	223.48	<0.001	2.81	<0.001	4	8	79
ABAT	4-aminobutyrate aminotransferase, mitochondrial	151.37	<0.001	1.54	<0.001	0	2	1
MAP1B	Microtubule-associated protein 1B	53.86	<0.001	1.31	0.0039	3	18	243
MYL4	Myosin light chain 4	41.70	<0.001	3.76	<0.001	0	1	4
PTGS2	Prostaglandin G/H synthase 2	-34.53	0.0029	-2.72	<0.001	0	0	2
EFNB1	Ephrin-B1	28.13	0.0144	3.13	<0.001	2	0	15
ASS1	Argininosuccinate synthase	27.16	<0.001	2.20	<0.001	2	3	20
DPF2	Zinc finger protein ubi-d4	16.70	<0.001	1.10	0.1844	4	5	21
ZNF384	Zinc finger protein 384	-16.59	<0.001	-1.12	0.1411	0	1	15
LMCD1	LIM and cysteine-rich domains protein 1	15.66	<0.001	1.10	0.2722	0	1	11
RAB2B	Ras-related protein Rab-2B	-12.60	0.0011	1.05	0.5055	1	0	11
TPX2	Targeting protein for Xklp2	-11.02	<0.001	-1.19	0.0338	7	11	50
ARHGDI1	Rho GDP-dissociation inhibitor 2	10.16	0.0012	1.50	<0.001	14	15	14
SLC4A11	Sodium bicarbonate transporter-like protein 11	9.62	<0.001	2.39	<0.001	6	0	7
L1CAM	Neural cell adhesion molecule L1	9.31	<0.001	1.58	<0.001	0	1	23
GCNT3	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase 3	-9.12	<0.001	-1.16	0.0588	0	0	2
SAMD9	Sterile alpha motif domain-containing protein 9	9.07	0.0025	1.30	0.0075	3	1	23
CWC27	Peptidyl-prolyl cis-trans isomerase CWC27 homolog	-8.79	<0.001	-1.08	0.2959	2	2	15
ZHX2	Zinc fingers and homeoboxes protein 2	-8.71	0.0047	1.05	0.4835	0	0	26
OBSL1	Obscurin-like protein 1	8.70	<0.001	-1.21	0.0273	0	0	13
PLAU	Urokinase-type plasminogen activator	-8.67	<0.001	-1.02	0.8216	1	0	4
AIM1L	Absent in melanoma 1-like protein	8.64	0.0145	1.05	0.5806	1	0	1
ZNF280C	Zinc finger protein 280C	-8.57	<0.001	-1.14	0.0734	3	1	35
CYP2S1	Cytochrome P450 2S1	7.99	0.0016	1.90	<0.001	0	1	9
ZBTB38	Zinc finger and BTB domain-containing protein 38	-7.87	0.0341	1.05	0.5409	4	3	18

^a Microarray data were obtained from the Gene Expression Omnibus database, accession number GSE5497 (20).

^b Ubiquitination (Ub), acetylation (Ac), and phosphorylation (Phos) post-translational modifications (PTMs) curated by PhosphoSitePlus (<http://www.phosphosite.org>) (54). Only sites identified in human were considered.

detectable by monitoring gene expression. However, mevalonate pathway up-regulation after reintroduction of wild-type ARID1A was observed at the protein level but not at the mRNA level (Table I). Although this discrepancy could be because of the fact that separate batches of cells, prepared at different times under somewhat different conditions, were used for the proteome and transcriptome analyses, decoupled responses for groups of genes and proteins are

commonly observed when comparing parallel global proteomics and transcriptomics data (44, 45). The better concordance observed between protein and gene expression changes for the 25 proteins with the largest fold-changes (Table II) suggests, but does not prove, that the mevalonate pathway is particularly sensitive to regulation at the post-transcriptional level. A number of studies support this hypothesized post-transcriptional regulation of the mevalonate

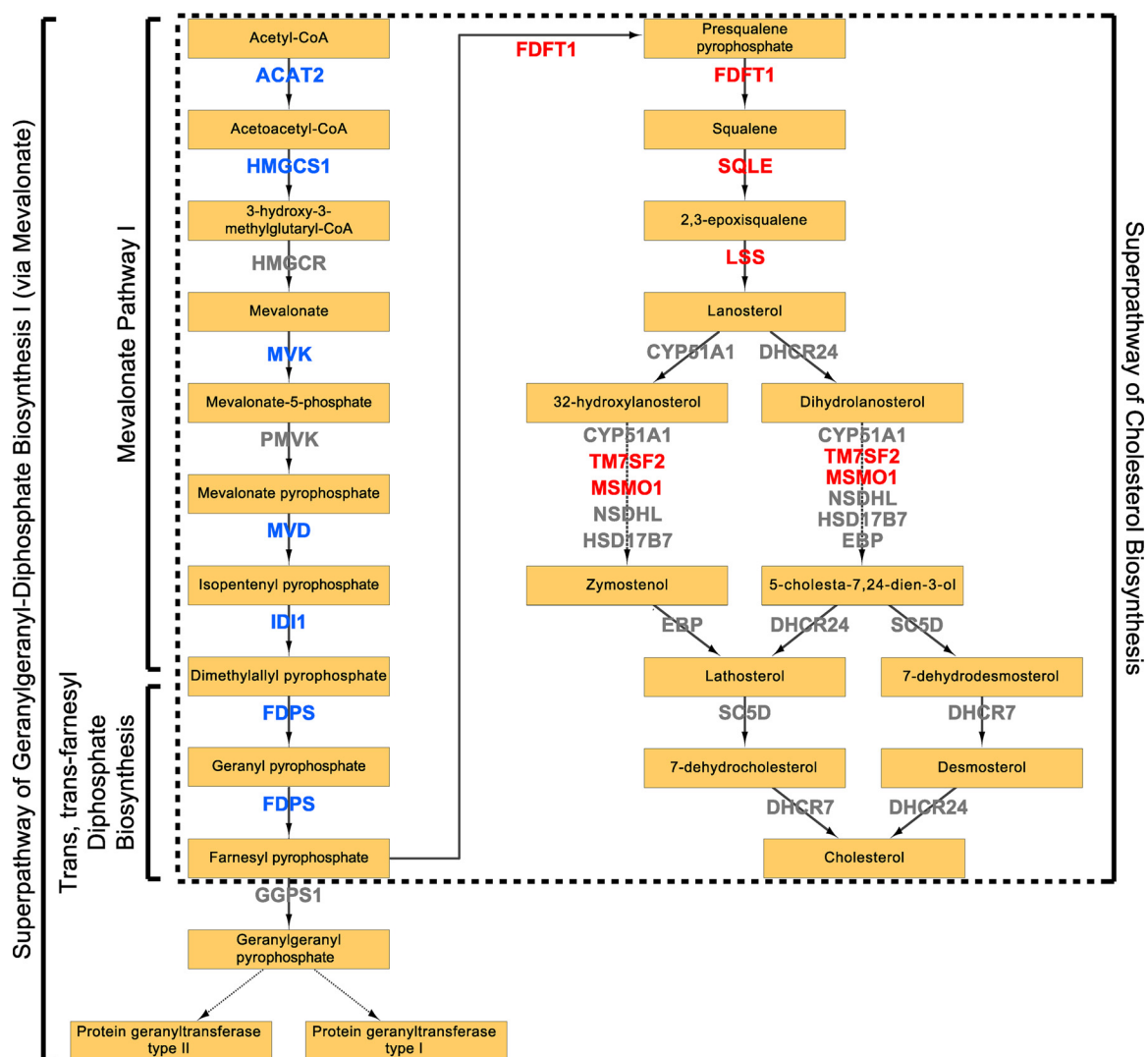


FIG. 5. **Regulation of mevalonate pathway by *ARID1A*.** Schematic diagram of mevalonate pathway based on IPA curation. Canonical pathways overrepresented upon *ARID1A* knockout in OVCA429 and induction in OVISE are labeled. *Blue*: enzymes significantly decreased in abundance in knockout and significantly increased in abundance in induction. For enzymes that significantly changed in one condition, we allowed for a slightly relaxed fold change in the other condition (*i.e.* 1.8-fold change with Student's *t* test $p < 0.05$). *Red*: enzymes significantly increased in abundance in induction and not significantly changed in knockout. Refer to Table I.

pathway. Several of these proteins (HMGCR, SQLE, and HMGCS1) have been reported to undergo ubiquitin-regulated proteasomal degradation (46–51). Further, almost all of the other proteins in this pathway have also been reported to be ubiquitinated, and many have been reported to be acetylated or phosphorylated (Table I) (52). Of the proteins that changed in level with *ARID1A* status, only MVK had no reported modifications. Therefore, the discordance between protein and mRNA changes for the mevalonate pathway when wild-type *ARID1A* is restored may be because of inhibition of protein degradation through the ubiquitin-proteasome system (53).

Further examination of the 25 proteins with the largest fold-changes illustrates the likely complex and indirect regulation of the proteome by *ARID1A* (Table II). Similar to mevalonate pathway proteins, DPF2, LMCD1, AIM1L, and OBSL1

showed increased protein levels that did not positively correlate with changes in gene expression. Eleven proteins showed concordant increases in protein and mRNA levels with much larger fold-changes at the protein level (*ARID1A*, ABAT, MAP1B, MYL4, EFHB1, ASS1, ARHGDI1B, SLC4A11, L1CAM, SAMD9, and CYP2S1), consistent with a combinatorial effect of inhibition of protein degradation and activation of gene transcription. However, these mechanisms do not fully explain the relationship between protein and mRNA levels observed among the top 25 proteins. PTGS2 and TPX2 showed concordant decreases in protein and mRNA levels, but the protein changes were an order-of-magnitude greater. Further, eight proteins (ZNF384, RAB2B, GCNT3, CWC27, ZHX2, PLAU, ZNF280C, and ZBTB38) did not display significant changes in mRNA levels but had large decreases at the protein level. Six

of these proteins have reported ubiquitination sites, and the other two proteins are known to be either acetylated or phosphorylated (Table II), suggesting that the basis for the decreased protein levels could be due to increased protein degradation through ubiquitination or changes in other post-translational modifications. These observations suggest that ARID1A may inhibit degradation of some proteins and promote degradation of other proteins possibly by differentially affecting specific ubiquitin ligases. Further studies are required to identify specific targets of ARID1A involved in post-transcriptional regulation, including potential changes in protein degradation as postulated here.

This study is, to the best of our knowledge, the first in-depth proteome analysis addressing the role of ARID1A in OCCC. Protein level changes rather than transcriptional changes best reflect the state of the cell when wild-type ARID1A is knocked out or restored. Identification of the mevalonate pathway as the primary target of ARID1A loss of function suggests areas for further investigation and potential therapeutic targeting for ARID1A-mutated cancers.

Acknowledgments—The authors gratefully acknowledge the assistance of the Wistar Institute Proteomics and Metabolomics Core Facility, as well as the Wistar Institute Bioinformatics Core Facility.

* This study was supported by National Institutes of Health Grant CA131582 (DWS) and an Ovarian Cancer Research Fund Program Project (DWS, JCG, RZ). ARG was supported by NCI training grant CA009171 to the Wistar Institute. BGB was supported by National Institutes of Health/NCI grant K99CA194318. Support for the Proteomics and Metabolomics Core Facility and the Bioinformatics Core Facility was provided by Cancer Center Support Grant CA010815 to the Wistar Institute.

☐ This article contains [supplemental material](#).

** To whom correspondence should be addressed: The Wistar Institute, 3601 Spruce St., Room 272a, Philadelphia, PA 19104. Tel.: 215-898-3972; E-mail: speicher@wistar.org.

REFERENCES

- Siegel, R. L., Miller, K. D., and Jemal, A. (2016) Cancer statistics, 2016. *CA Cancer J. Clin.* **66**, 7–30
- Itamochi, H., Kigawa, J., and Terakawa, N. (2008) Mechanisms of chemoresistance and poor prognosis in ovarian clear cell carcinoma. *Cancer Sci.* **99**, 653–658
- Pectasides, D., Pectasides, E., Psyrris, A., and Economopoulos, T. (2006) Treatment issues in clear cell carcinoma of the ovary: a different entity? *Oncologist* **11**, 1089–1094
- Kurman, R. J., and Shih Ie, M. (2008) Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. *Int. J. Gynecol. Pathol.* **27**, 151–160
- Yamaguchi, K., Mandai, M., Oura, T., Matsumura, N., Hamanishi, J., Baba, T., Matsui, S., Murphy, S. K., and Konishi, I. (2010) Identification of an ovarian clear cell carcinoma gene signature that reflects inherent disease biology and the carcinogenic processes. *Oncogene* **29**, 1741–1752
- Wiegand, K. C., Shah, S. P., Al-Agha, O. M., Zhao, Y., Tse, K., Zeng, T., Senz, J., McConechy, M. K., Anglesio, M. S., Kalloger, S. E., Yang, W., Heravi-Moussavi, A., Giuliany, R., Chow, C., Fee, J., Zayed, A., Prentice, L., Melnyk, N., Turashvili, G., Delaney, A. D., Madore, J., Yip, S., McPherson, A. W., Ha, G., Bell, L., Fereday, S., Tam, A., Galletta, L., Tonin, P. N., Provencher, D., Miller, D., Jones, S. J., Moore, R. A., Morin, G. B., Oloumi, A., Boyd, N., Aparicio, S. A., Shih Ie, M., Mes-Masson, A. M., Bowtell, D. D., Hirst, M., Gilks, B., Marra, M. A., and Huntsman, D. G. (2010) ARID1A mutations in endometriosis-associated ovarian carcinomas. *N. Engl. J. Med.* **363**, 1532–1543
- Jones, S., Wang, T. L., Shih Ie, M., Mao, T. L., Nakayama, K., Roden, R., Glas, R., Slamon, D., Diaz, L. A., Jr, Vogelstein, B., Kinzler, K. W., Velculescu, V. E., and Papadopoulos, N. (2010) Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science* **330**, 228–231
- Lawrence, M. S., Stojanov, P., Mermel, C. H., Robinson, J. T., Garraway, L. A., Golub, T. R., Meyerson, M., Gabriel, S. B., Lander, E. S., and Getz, G. (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501
- Wang, K., Kan, J., Yuen, S. T., Shi, S. T., Chu, K. M., Law, S., Chan, T. L., Kan, Z., Chan, A. S., Tsui, W. Y., Lee, S. P., Ho, S. L., Chan, A. K., Cheng, G. H., Roberts, P. C., Rejto, P. A., Gibson, N. W., Pocalyko, D. J., Mao, M., Xu, J., and Leung, S. Y. (2011) Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat. Genet.* **43**, 1219–1223
- Mamo, A., Cavallone, L., Tuzmen, S., Chabot, C., Ferrario, C., Hassan, S., Edgren, H., Kallioniemi, O., Aleynikova, O., Przybytkowski, E., Malcolm, K., Mousses, S., Tonin, P. N., and Basik, M. (2012) An integrated genomic approach identifies ARID1A as a candidate tumor-suppressor gene in breast cancer. *Oncogene* **31**, 2090–2100
- Helming, K. C., Wang, X., and Roberts, C. W. (2014) Vulnerabilities of mutant SWI/SNF complexes in cancer. *Cancer Cell.* **26**, 309–317
- Kadoch, C., and Crabtree, G. R. (2015) Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447
- Chandler, R. L., Brennan, J., Schisler, J. C., Serber, D., Patterson, C., and Magnuson, T. (2013) ARID1a-DNA interactions are required for promoter occupancy by SWI/SNF. *Mol. Cell. Biol.* **33**, 265–280
- Patsialou, A., Wilsker, D., and Moran, E. (2005) DNA-binding properties of ARID family proteins. *Nucleic Acids Res.* **33**, 66–80
- Dallas, P. B., Pacchione, S., Wilsker, D., Bowrin, V., Kobayashi, R., and Moran, E. (2000) The human SWI-SNF complex protein p270 is an ARID family member with non-sequence-specific DNA binding activity. *Mol. Cell. Biol.* **20**, 3137–3146
- Nie, Z., Xue, Y., Yang, D., Zhou, S., Deroo, B. J., Archer, T. K., and Wang, W. (2000) A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex. *Mol. Cell. Biol.* **20**, 8879–8888
- Guan, B., Wang, T. L., and Shih Ie, M. (2011) ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers. *Cancer Res.* **71**, 6718–6727
- Helming, K. C., Wang, X., Wilson, B. G., Vazquez, F., Haswell, J. R., Manchester, H. E., Kim, Y., Kryukov, G. V., Ghandi, M., Aguirre, A. J., Jagani, Z., Wang, Z., Garraway, L. A., Hahn, W. C., and Roberts, C. W. (2014) ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat. Med.* **20**, 251–254
- Raab, J. R., Resnick, S., and Magnuson, T. (2015) Genome-Wide Transcriptional Regulation Mediated by Biochemically Distinct SWI/SNF Complexes. *PLoS Genet.* **11**, e1005748
- Bitler, B. G., Aird, K. M., Garipov, A., Li, H., Amatangelo, M., Kossenkov, A. V., Schultz, D. C., Liu, Q., Shih Ie, M., Conejo-Garcia, J. R., Speicher, D. W., and Zhang, R. (2015) Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. *Nat. Med.* **21**, 231–238
- Zhai, Y., Kuick, R., Tipton, C., Wu, R., Sessine, M., Wang, Z., Baker, S. J., Fearon, E. R., and Cho, K. R. (2016) Arid1a inactivation in an Apc- and Pten-defective mouse ovarian cancer model enhances epithelial differentiation and prolongs survival. *J. Pathol.* **238**, 21–30
- Vogel, C., and Marcotte, E. M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232
- Wiegand, K. C., Hennessy, B. T., Leung, S., Wang, Y., Ju, Z., McGahren, M., Kalloger, S. E., Finlayson, S., Stemke-Hale, K., Lu, Y., Zhang, F., Anglesio, M. S., Gilks, B., Mills, G. B., Huntsman, D. G., and Carey, M. S. (2014) A functional proteogenomic analysis of endometrioid and clear cell carcinomas using reverse phase protein array and mutation analysis: protein expression is histotype-specific and loss of ARID1A/BAF250a is associated with AKT phosphorylation. *BMC Cancer* **14**, 120
- Beer, L. A., Tang, H. Y., Sriswasdi, S., Barnhart, K. T., and Speicher, D. W. (2011) Systematic discovery of ectopic pregnancy serum biomarkers

- using 3-D protein profiling coupled with label-free quantitation. *J. Proteome Res.* **10**, 1126–1138
25. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372
 26. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805
 27. Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526
 28. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (pro)teomics data. *Nat. Methods* **13**, 731–740
 29. Lazar, C., Gatto, L., Ferro, M., Bruley, C., and Burger, T. (2016) Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. *J. Proteome Res.* **15**, 1116–1125
 30. Berchuck, A., Rodriguez, G., Olt, G., Whitaker, R., Boente, M. P., Arrick, B. A., Clarke-Pearson, D. L., and Bast, R. C., Jr. (1992) Regulation of growth of normal ovarian epithelial cells and ovarian cancer cell lines by transforming growth factor-beta. *Am. J. Obstet. Gynecol.* **166**, 676–684
 31. Nagaraj, N., Wisniewski, J. R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Paabo, S., and Mann, M. (2011) Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* **7**, 548
 32. Geiger, T., Wehner, A., Schaab, C., Cox, J., and Mann, M. (2012) Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol. Cell. Proteomics* **11**, M111 014050
 33. Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M., Workman, J. L., and Crabtree, G. R. (1996) Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15**, 5370–5382
 34. Goldstein, J. L., and Brown, M. S. (1990) Regulation of the mevalonate pathway. *Nature* **343**, 425–430
 35. Buhaescu, I., and Izzedine, H. (2007) Mevalonate pathway: a review of clinical and therapeutical implications. *Clin. Biochem.* **40**, 575–584
 36. Morrison, D. K. (2009) The 14–3–3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. *Trends Cell Biol.* **19**, 16–23
 37. Phan, L., Chou, P. C., Velazquez-Torres, G., Samudio, I., Parreno, K., Huang, Y., Tseng, C., Vu, T., Gully, C., Su, C. H., Wang, E., Chen, J., Choi, H. H., Fuentes-Mattei, E., Shin, J. H., Shiang, C., Grabiner, B., Blonska, M., Skerl, S., Shao, Y., Cody, D., Delacerda, J., Kingsley, C., Webb, D., Carlock, C., Zhou, Z., Hsieh, Y. C., Lee, J., Elliott, A., Ramirez, M., Bankson, J., Hazle, J., Wang, Y., Li, L., Weng, S., Rizk, N., Wen, Y. Y., Lin, X., Wang, H., Wang, H., Zhang, A., Xia, X., Wu, Y., Habra, M., Yang, W., Pusztai, L., Yeung, S. C., and Lee, M. H. (2015) The cell cycle regulator 14–3–3sigma opposes and reverses cancer metabolic reprogramming. *Nat. Commun.* **6**, 7530
 38. Gorai, I., Nakazawa, T., Miyagi, E., Hirahara, F., Nagashima, Y., and Minaguchi, H. (1995) Establishment and characterization of two human ovarian clear cell adenocarcinoma lines from metastatic lesions with different properties. *Gynecol. Oncol.* **57**, 33–46
 39. Berndt, N., Hamilton, A. D., and Sebt, S. M. (2011) Targeting protein prenylation for cancer therapy. *Nat. Rev. Cancer* **11**, 775–791
 40. Clendening, J. W., Pandya, A., Boutros, P. C., El Ghamrasni, S., Khosravi, F., Trentin, G. A., Martirosyan, A., Hakem, A., Hakem, R., Jurisica, I., and Penn, L. Z. (2010) Dysregulation of the mevalonate pathway promotes transformation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15051–15056
 41. Thurnher, M., Nussbaumer, O., and Gruenbacher, G. (2012) Novel aspects of mevalonate pathway inhibitors as antitumor agents. *Clin. Cancer Res.* **18**, 3524–3531
 42. Freed-Pastor, W. A., Mizuno, H., Zhao, X., Langerod, A., Moon, S. H., Rodriguez-Barrueco, R., Barsotti, A., Chicas, A., Li, W., Polotskaia, A., Bissell, M. J., Osborne, T. F., Tian, B., Lowe, S. W., Silva, J. M., Borresen-Dale, A. L., Levine, A. J., Bargonetti, J., and Prives, C. (2012) Mutant p53 disrupts mammary tissue architecture via the mevalonate pathway. *Cell* **148**, 244–258
 43. Baandrup, L., Dehlendorff, C., Friis, S., Olsen, J. H., and Kjaer, S. K. (2015) Statin use and risk for ovarian cancer: a Danish nationwide case-control study. *Br. J. Cancer* **112**, 157–161
 44. Fathi, A., Pakzad, M., Taei, A., Brink, T. C., Pirhaji, L., Ruiz, G., Sharif Tabatabaee, M., Gourabi, H., Adjaye, J., Baharvand, H., and Salekdeh, G. H. (2009) Comparative proteome and transcriptome analyses of embryonic stem cells during embryoid body-based differentiation. *Proteomics* **9**, 4859–4870
 45. Sap, K. A., Bezstarosti, K., Dekkers, D. H., van den Hout, M., van Ijcken, W., Rijkers, E., and Demmers, J. A. (2015) Global quantitative proteomics reveals novel factors in the ecdysone signaling pathway in *Drosophila melanogaster*. *Proteomics* **15**, 725–738
 46. Burg, J. S., and Espenshade, P. J. (2011) Regulation of HMG-CoA reductase in mammals and yeast. *Prog. Lipid Res.* **50**, 403–410
 47. Jo, Y., and Debose-Boyd, R. A. (2010) Control of cholesterol synthesis through regulated ER-associated degradation of HMG CoA reductase. *Crit. Rev. Biochem. Mol. Biol.* **45**, 185–198
 48. Hampton, R. Y., and Bhakta, H. (1997) Ubiquitin-mediated regulation of 3-hydroxy-3-methylglutaryl-CoA reductase. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12944–12948
 49. Ravid, T., Doolman, R., Avner, R., Harats, D., and Roitelman, J. (2000) The ubiquitin-proteasome pathway mediates the regulated degradation of mammalian 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *J. Biol. Chem.* **275**, 35840–35847
 50. Gill, S., Stevenson, J., Kristiana, I., and Brown, A. J. (2011) Cholesterol-dependent degradation of squalene monoxygenase, a control point in cholesterol synthesis beyond HMG-CoA reductase. *Cell Metab.* **13**, 260–273
 51. Sapir, A., Tsur, A., Koorman, T., Ching, K., Mishra, P., Bardenheier, A., Podolsky, L., Bening-Abu-Shach, U., Boxem, M., Chou, T. F., Broday, L., and Sternberg, P. W. (2014) Controlled sumoylation of the mevalonate pathway enzyme HMGs-1 regulates metabolism during aging. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3880–E3889
 52. Sharpe, L. J., and Brown, A. J. (2013) Controlling cholesterol synthesis beyond 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR). *J. Biol. Chem.* **288**, 18707–18715
 53. Ciechanover, A. (2005) Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nat. Rev. Mol. Cell Biol.* **6**, 79–87
 54. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520