Landscape of somatic single-nucleotide and copy-number mutations in uterine serous carcinoma

Siming Zhao[a], Murim Choi[a], John D. Overton[a], Stefania Bellone[b], Dana M. Roque[b], Emiliano Cocco[b], Federica Guzzo[b], Diana P. English[b], Joyce Varughese[b], Sara Gasparrini[b], Ileana Bortolomai[b], Natalia Buza[c], Pei Hui[c], Maysa Abu-Khalaf[d], Antonella Ravaggi[e], Eliana Bignotti[e], Elisabetta Bandiera[e], Chiara Romani[e], Paola Todeschini[e], Renata Tassi[e], Laura Zanotti[e], Luisa Carrara[e], Sergio Pecorelli[e], Dan-Arin Silasi[b], Elena Ratner[b], Masoud Azodi[b], Peter E. Schwartz[b], Thomas J. Rutherford[b], Amy L. Stiegler[f], Shrikant Mane[a], Titus J. Boggon[f], Joseph Schlessinger[f], Richard P. Lifton[a,1], and Alessandro D. Santin[b]

[a]Department of Genetics, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT 06510; [b]Departments of Obstetrics, Gynecology, and Reproductive Sciences, [c]Pathology, [d]Internal Medicine and Oncology, and [f]Pharmacology, Yale University School of Medicine, New Haven, CT 06510; and [e]Department of Obstetrics, and Gynecology, "Angelo Nocivelli" Institute of Molecular Medicine, University of Brescia, 25123 Brescia, Italy

Contributed by Richard P. Lifton, December 27, 2012 (sent for review December 12, 2012)

Uterine serous carcinoma (USC) is a biologically aggressive subtype of endometrial cancer. We analyzed the mutational landscape of USC by whole-exome sequencing of 57 cancers, most of which were matched to normal DNA from the same patients. The distribution of the number of protein-altering somatic mutations revealed that 52 USC tumors had fewer than 100 (median 36), whereas 5 had more than 3,000 somatic mutations. The mutations in these latter tumors showed hallmarks of defects in DNA mismatch repair. Among the remainder, we found a significantly increased burden of mutation in 14 genes. In addition to well-known cancer genes (i.e., *TP53*, *PIK3CA*, *PPP2R1A*, *KRAS*, *FBXW7*), there were frequent mutations in *CHD4/Mi2b*, a member of the NuRD–chromatin-remodeling complex, and *TAF1*, an element of the core TFIID transcriptional machinery. Additionally, somatic copy-number variation was found to play an important role in USC, with 13 copy-number gains and 12 copy-number losses that occurred more often than expected by chance. In addition to loss of *TP53*, we found frequent deletion of a small segment of chromosome 19 containing *MBD3*, also a member of the NuRD–chromatin-modification complex, and frequent amplification of chromosome segments containing *PIK3CA*, *ERBB2* (an upstream activator of PIK3CA), and CCNE1 (a target of FBXW7-mediated ubiquitination). These findings identify frequent mutation of DNA damage, chromatin remodeling, cell cycle, and cell proliferation pathways in USC and suggest potential targets for treatment of this lethal variant of endometrial cancer.

endometrial carcinoma | uterine serous papillary cancer | cancer genomics

Endometrial cancer is the most prevalent gynecologic tumor in women, with an annual incidence of 47,130 new cases and 8,010 deaths in 2012 in the United States (1). On the basis of clinical and histopathological features, endometrial cancer is classified into type I and type II disease groups (2). Type I tumors, which constitute the majority of cases, are generally diagnosed at an early stage, are low grade and endometrioid in histology, are associated with a history of hyperestrogenism, and typically have a good prognosis. In contrast, type II cancers are poorly differentiated, often with serous papillary [uterine serous carcinoma (USC)] or clear cell histology. Although these tumors account for a minority of endometrial cancers, the majority of relapses and deaths occur in this group of patients (2).

Among type II cancers, USC represents the most biologically aggressive subtype (3, 4). Classically, the neoplastic epithelium is characterized by serous differentiation with psammoma bodies and a predominantly papillary architecture (3). Pleomorphic cytology with nuclear atypia, prominent nucleoli, a vescicular chromatin pattern, and high mitotic activity are seen. Clinically, USC has a propensity for early intra-abdominal and lymphatic spread (3) and is more commonly diagnosed in women of African ancestry (3–5). The overall 5-y survival of USC is only 30 ± 9% for all stages, and the recurrence rate after surgery is extremely high (50–80%) (3–5). This poor prognosis motivates determination of the molecular basis of this tumor's aggressive behavior in hope of developing new effective treatment modalities.

To address these issues, we have sequenced the exomes of a large cohort of USC and have identified genes with increased numbers of somatic single-nucleotide and copy-number variants. The results identify cancer genes and define the pathways involved in USC.

Results

Exome Sequencing of USC. Fifty-seven patients with uterine serous carcinoma were studied. Their clinical features are presented in *SI Appendix*, Table S1. Upon surgical removal of tumors, primary cell lines were prepared (15 tumors) or tumors were frozen (42 tumors). Exome sequencing was performed on all tumors; for 34 of these, DNA samples from normal tissue were available and sequenced. Exome sequencing was performed using the NimbleGen/Roche capture reagent followed by 74 base paired-end DNA sequencing on the Illumina HiSeq platform (6). By design, tumor samples were sequenced to greater depth of coverage to permit detection of somatic mutations in tumors despite admixture of normal and tumor cells in these samples. For tumors and normal DNA, each targeted base was sequenced by a mean of 187 and 100 independent reads, respectively (*SI Appendix*, Table S2). Of all targeted bases in tumors, 94.5% were read by 20 or more independent reads; mean per-base per read error rates were 0.42% for normal DNA and 0.48% for tumor DNA. Segments of loss of heterozygosity (LOH) were called from the difference in B-allele frequency between tumor-normal pairs (*SI Appendix*, Fig. S1), allowing estimates of tumor purity, which were above 60% for frozen tumors and higher for primary cell lines. Somatic mutations were identified by finding variant reads in tumors that were significantly more frequent than expected by chance (*Materials and Methods*). At the coverage levels studied, there was no significant relationship between tumor purity and the number of somatic variants detected, consistent with sufficient depth of coverage having been achieved to identify the vast majority of somatic mutations. Variants in genes implicated in the pathogenesis of

Author contributions: J.S., R.P.L., and A.D.S. designed research; J.D.O., S.B., D.M.R., E.C., F.G., S.G., I.B., N.B., P.H., A.R., E. Bignotti, E. Bandiera, C.R., P.T., R.T., L.Z., L.C., A.L.S., S.M., and T.J.B. performed research; D.P.E., J.V., M.A.-K., S.P., D.-A.S., E.R., M.A., P.E.S., and T.J.R. contributed new reagents/analytic tools; S.Z., M.C., and R.P.L. analyzed data; and S.Z., A.L.S., T.J.B., J.S., R.P.L., and A.D.S. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: richard.lifton@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1222577110/-/DCSupplemental.

USC were verified by direct Sanger sequencing and were found to be expressed in all available USC cell lines.

Tumors with Hypermutator Phenotype. The number of protein-altering somatic mutations per tumor markedly deviated from a normal distribution (Fig. 1*A*). In the discovery set of 34 USC with matched normal DNA, 30 tumors had fewer than 100 protein-altering somatic mutations (median 36), whereas 4 had more than 3,000 somatic mutations each. Only one of these tumors was from a cell line (with limited propagation), and none came from patients who had received chemotherapy before sample acquisition. These tumors with high mutation burden were also notable for having no LOH segments or copy-number variants (CNVs), a feature found in only five other tumors. These features suggest a hypermutator phenotype due to deficiency of mismatch repair (MMR) or polymerase ε (*POLE*) genes (7, 8). Consistent with this, these hypermutated tumors showed a paucity of T:A > A:T or C:G > A:T transversions (*SI Appendix*, Fig. S2) (9). Examination of the *POLE* and MMR genes showed no germ-line mutations; however, somatic mutations in these genes were highly prevalent in these tumors (mean of 4 per tumor, including 4 premature termination mutations for MMR genes and a mean of 4.5 per tumor for *POLE*) and more frequent than expected by chance ($P = 2.23 \times 10^{-3}$) (*SI Appendix*, Table S3). Among the cancers without matched normal DNA, one showed a similarly high prevalence of rare protein-altering variants (>3,000) and a skewed distribution of rare protein-altering transversions. Thus, 9% of USC in this cohort have a hypermutator phenotype. Because of the skewing effect of the large number of mutations in these tumors, they were not included in subsequent analyses of mutation burden.

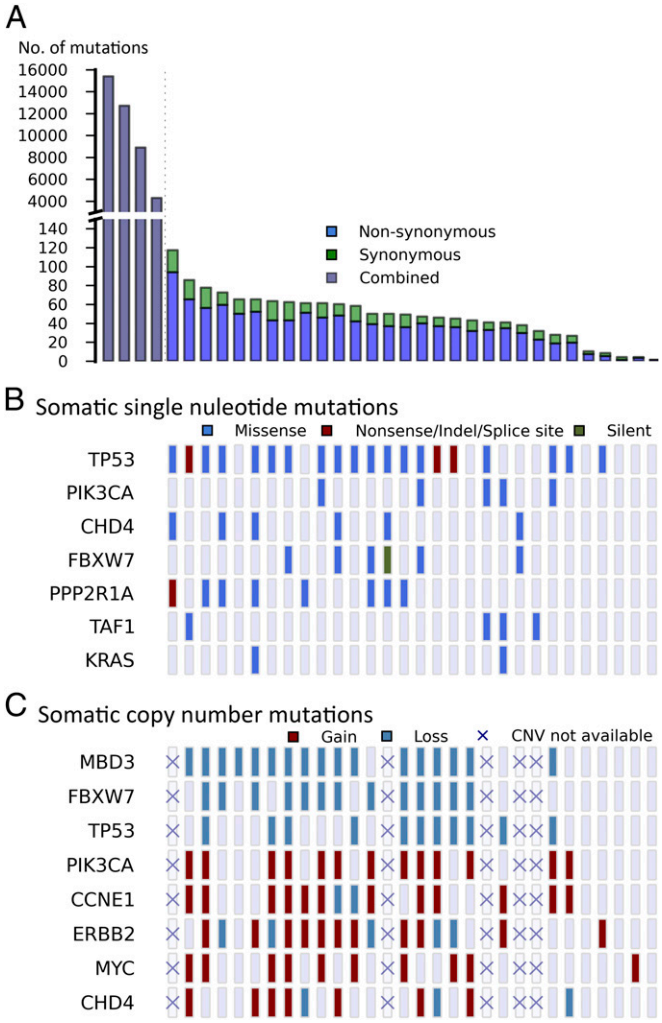

Fig. 1. Somatic variation pattern underlying USC. (*A*) Distribution of the number of protein-altering somatic mutations in 34 normal-tumor USC pairs. Subplot left, mutation spectrum in four hypermutator phenotype samples; subplot right, mutation spectrum in 30 moderately mutated samples. (*B* and *C*) Thirty tumors with moderate somatic burden are arranged by the total number of somatic point mutations from left to right. The four hypermutator phenotype tumors are excluded in this analysis. (*B*) Significantly mutated genes are listed vertically by the order of damaging or conserved *P* values shown in Table 1. (*C*) Genes with significant CNVs and genes of interest are listed. Copy neutral status is shown as lavender rectangles. Five samples without CNV information are marked by crosses.

Analysis of Single-Nucleotide Variants. Among somatic mutations in the 30 remaining matched tumors, we identified recurrences of somatic mutations at the same positions. Accounting for the rate of protein-altering somatic mutations in these tumors (1.1×10^{-6}) and the size of the exome, the likelihood of seeing the mutation twice by chance at any position among these tumors is $<10^{-3}$. We identified six genes with recurrent somatic mutations (Table 1). These included well-established activating mutations in *PIK3CA* (10), the catalytic subunit of phosphoinositide-3 kinase (five tumors); the well-established G12V mutation in *KRAS* (three tumors) (11); and a mutation at R465 in *FBXW7* in four tumors (12). FBXW7 is the targeting component of a SCF-type 3 ubiquitin ligase, and R465 occurs in the WD40 domain involved in substrate recognition; mutation at this site prevents targeting of cyclin E for ubiquitination and degradation (*SI Appendix*, Fig. S3) (12, 13). Recurrent mutations also occurred at two sites in *PPP2R1A*, the constant regulatory subunit of serine-threonine phosphatase 2a. The P179R and S256F mutations occurred four and two times, respectively, and have been previously reported (14). These mutations occur at the interface where PPP2R1A interacts with regulatory B subunits that target the phosphatase to specific substrates; inhibition of this interaction by SV40 small t antigen plays a role in viral transformation (*SI Appendix*, Fig. S4). Additional somatic mutations were found on the surface of PPP2R1A that interacts with the B or C (catalytic) subunit (Table 1). *TP53*, the well-characterized tumor suppressor gene, had five different positions mutated two or more times, and there were 19 additional single somatic mutations in this gene. Eighty-two percent of these mutations were in segments of somatic LOH (Table 1 and Fig. 1*B*).

In addition to these previously described recurrent mutations, a recurrent mutation was found in CHD4/Mi2b (chromodomain-helicase–DNA-binding protein 4), an ATP-dependent chromatin-remodeling protein that is a major subunit of the Mi2b/nucleosome remodeling and deacetylase (NuRD) complex. Mutations in *CHD4/Mi2b* have not been previously associated with cancer. In addition, there were 10 other somatic or rare mutations in *CHD4* among matched and unmatched tumors (see further discussion below).

We next sought genes with overall increased somatic mutation burden in the 30 matched tumor-normal pairs. In this analysis, we determined the probability of seeing $\geq n$ mutations in each gene, taking into account the overall rate of protein-altering somatic mutations in the matched tumor normal set (1.1×10^{-6}) and the length of the protein-coding region in each gene. We also adjusted for the level of expression of each gene from expression data in normal human endometrium (15) because we found a higher somatic mutation rate among genes with lower expression, consistent with effects of transcription-coupled DNA repair reducing the mutation rate among expressed genes (16). We considered *P* values $< 2.4 \times 10^{-6}$ to represent a significant increase in mutation burden compared with that expected under the null hypothesis, accounting for the testing of ~21,000 genes. We added to this set variants in the 22 unmatched tumors that occurred in genes that had at least one somatic mutation in the matched set and that had

Table 1. Genes with significant mutation burden in 52 USC

Gene	Recurrent (no.)	Coding length (bp)	No. nonsynonymous mutations	Nonsynonymous *P* value	No. damaging mutations	No. Cons MS	Dam+ Cons *P* value	No. silent mutations
TP53	R43 (2), C44 (2), R116 (4), G134 (3), R141 (2)	1,222	30	1.34*E*-75	3	14	1.01*E*-45	0
PIK3CA	E542 (3), N1044 (2)	3,287	13	6.39*E*-23	0	13	4.26*E*-27	0
CHD4	F1112 (2)	5,895	11	3.82*E*-17	1	10	1.05*E*-20	0
FBXW7	R465 (4)	2,168	9	1.05*E*-17	0	9	1.08*E*-20	1
PPP2R1A	P179 (4), S256 (2)	1,830	8	1.28*E*-16	0	8	2.60*E*-19	0
TAF1	—	5,834	7	2.71*E*-10	0	7	1.19*E*-12	0
KRAS	G12 (2)	586	3	2.61*E*-08	0	3	1.66*E*-09	0
PTEN	—	1,248	3	2.28*E*-07	1	2	1.46*E*-08	0
HCFC1R1	—	433	2	8.00*E*-07	0	2	1.01*E*-07	0
CDKN1A	—	503	2	1.25*E*-06	0	2	1.58*E*-07	0
CTDSPL	—	863	2	7.86*E*-06	0	2	9.99*E*-07	0
YIPF3	—	1,089	2	1.25*E*-05	0	2	1.59*E*-06	0
SPOP	—	1,161	2	1.51*E*-05	0	2	1.92*E*-06	0
FAM132A	—	941	2	1.51*E*-05	0	2	1.93*E*-06	0

Recurrent (no.), positions with recurrent mutations (no. of instances); Cons MS, missense mutations at conserved positions; Dam+Cons, damaging + conserved missense mutations.

never been seen in >7,000 exomes in the Yale University and National Heart, Lung, and Blood Institute exome databases. Because we found no novel variants in any of these genes in the 30 germ-line samples of tumor-normal pairs, we infer that virtually all of these represent somatic mutations.

In the resulting set, the six genes with recurrent mutations were among the most frequently mutated genes. Included in this set was *CHD4*, which had six somatic mutations and five more novel variants found in the 22 unmatched tumors (Table 1 and Fig. 1*B*). Many of these 11 *CHD4* mutations (Fig. 2), which all appear to be heterozygous, impair at least some normal CHD4 functions. CHD4 is a SWI2/SNF2 ATPase and part of the larger helicase superfamily 2 whose members share a similar catalytic core containing two RecA-like helicase domains. Conserved catalytic "signature" motifs have been well described and contain many residues required for catalysis of ATP hydrolysis and helicase activity (20). Three *CHD4* mutations (R957Q, R1127G, and R1162W) alter residues in these signature motifs (motif B, V, and VI, respectively) that are conserved from yeast to humans, and whose mutation has been shown to impair normal function. Similarly, there is a mutation in the second plant homeodomain (PHD) finger that normally binds methylated histone H3K9. This C464Y mutation disrupts one of the key cysteines that coordinate Zn^{2+} binding. In addition, there are four mutations (Q1106R, I1109T, and two instances of F1112L)

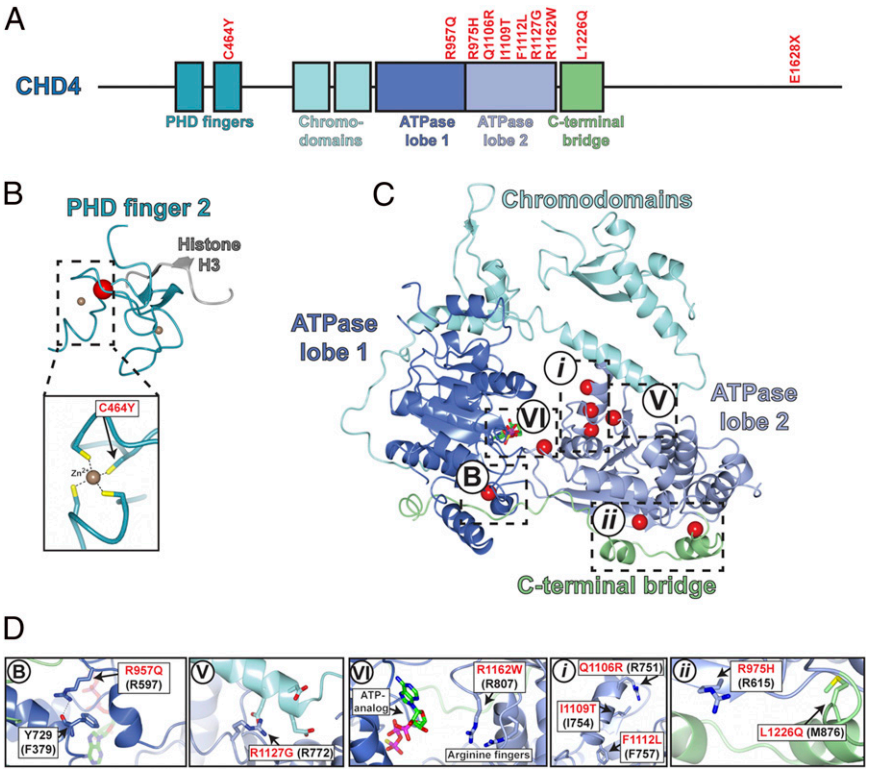

Fig. 2. Mapping of USC mutations onto the crystal structure of CHD4. (*A*) Schematic representation of somatic mutations found in CHD4. The horizontal bar represents full-length CHD4 protein with functional domains shown as boxes. Somatic mutations found in USC are marked in red text. All mutations are missense mutations except E1628X, which is a nonsense mutation. (*B*) C464 locates in the second PHD finger, which binds directly to histone H3 methylated at K9. An NMR structure of the second PHD domain of CHD4 (blue ribbon plot) has been determined [Protein Data Bank (PDB) ID: 2L75] (17). Red dot represents C464Y mutation. (*Lower*) A close-up view of the zinc–C464 interaction. (*C*) Somatic mutations in catalytic core of CHD4 mapped to the crystal structure of a related protein, human CHD1 (sequence identity of the ATPase lobes is 42%, homology is 57% over 572 residues; PDB ID: 3MWY) (18). Blue, ATPase lobe 1; light purple, ATPase lobe 2; cyan, chromodomains; green, C-terminal bridge. Three mutations (red dots) in CHD4 fall within known conserved motifs (motifs B, V, VI) (19); three mutations were found in unknown helicase motif *i*, and two mutations were found in unknown motif *ii*. (*D*) A close-up view of mutations in five motifs in *C*. Somatic mutations are labeled in red text; amino acid positions in parentheses represent homologous positions in CHD1.

clustered in a short α-helix in ATPase lobe 2; we speculate that alteration of this helix might alter interaction with another protein in the complex. Additional mutations include two in a C-terminal bridge that links ATPase lobes 1 and 2 at positions that appear to stabilize this segment. Finally, there is a premature termination near the normal C terminus (Fig. 2). The high prevalence of *CHD4* mutations, the clear implication of disrupted normal function by many *CHD4* mutations, and the common genomic amplifications of the *CHD4* gene (see CNV results below) implicate *CHD4* mutations in USC.

Another gene of interest was *TAF1*, an X-linked gene, which had four different somatic mutations and three additional variants in unmatched tumors (Table 1). TAF1 is the largest component and core scaffold of the TFIID basal transcription factor complex and has DNA-binding activity, histone acetyltransferase (HAT) activity, two kinase domains, and ubiquitin-activating/conjugating activity (21). It is known to be required for progression through the G1 phase of the cell cycle, promoting cyclin D expression (22). Most of the seven *TAF1* mutations lie in the HAT domain at positions that are extraordinarily well conserved—all are conserved in vertebrates and nearly all are conserved in yeasts (Fig. 3). Although the function(s) of these mutations in USC are uncertain, overexpression of *TAF1* has been previously reported in human lung and breast carcinoma and found to be associated with poor tumor differentiation and high mitotic activity (25).

Additional genes that meet thresholds for significantly increased burden in the entire set include *PTEN*, *CDKN1A*, and *SPOP*, as well as *HCFC1R1*, *CTDSPL*, *YIPF3*, and *FAM132A*, genes not previously implicated in cancer. For each of these genes, mutations are predominantly at highly conserved positions, there are few if any silent mutations in the same gene, and quantitative PCR demonstrated expression of each of these genes in all available USC cell lines.

Analysis of CNVs. We next assessed somatic CNVs. For the 25 tumors in which read coverage distribution showed distinct modes (*SI Appendix*, Fig. S5), CNVs were identified by comparing coverage depth of individual capture intervals from tumor and normal samples (*Materials and Methods*); CNVs were supported by significant deviation of the B allele frequency from the genome-wide average. The significance of CNVs affecting specific chromosome segments was assessed by Monte Carlo simulation, randomly distributing CNVs of the empirically observed sizes and numbers in each tumor in 10^8 permutations to assess the distribution expected by chance alone. A significance threshold was established that provided a false discovery rate <0.25. Within each significant copy-number gain or loss, all CNVs that contained the most frequently altered segment were removed, and the remaining CNVs were reassessed to see if independent signals could be detected. We identified 13 chromosome segments with more frequent gains of copy number and 12 with more frequent deletions than expected by chance (Fig. 4). Among these, we found focal amplification of the segment of chromosome 17 that contains *ERBB2* in 11 of the 25 tumors (44%) (*SI Appendix*, Fig. S6), large duplications that include the *PIK3CA* locus in 60%, and a small duplication of chromosome 19 containing *CCNE1* in 48% (*SI Appendix*, Table S4). There was also amplification of a large segment of chromosome 8 containing *MYC* in 11 (44%) tumors and amplification of a segment of chromosome 12 that included *CHD4* in 7 (28%) tumors (Fig. 1*C*). Among deletions, *TP53* was deleted in 44% of tumors. The most frequent somatic deletions were small (0.5 Mb) deletions on chromosomes 19 and 22, which occurred in 68% and 72% of tumors, respectively (*SI Appendix*, Table S5 and Fig. S7). Most interestingly, the chromosome 19 interval contains *MBD3*, which is a component of the same SWI/SNF complex as *CHD4* (19). The chromosome 22 interval includes a number of interesting genes, including three in the MAP kinase pathway, *HDAC10*, and *PPP6R2*.

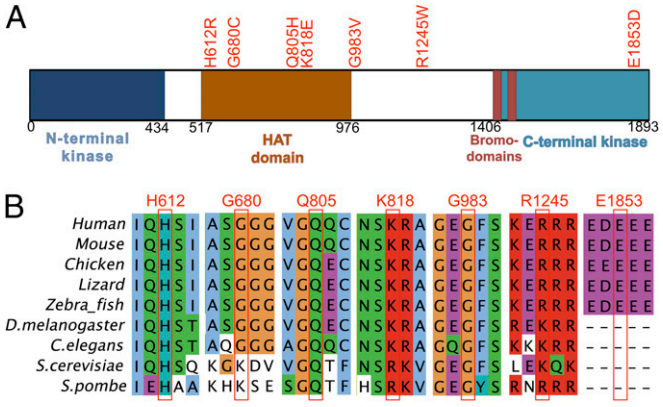

Fig. 3. Schematic representation of TAF1 functional domains and mutation conservation analysis. (*A*) Functional domains in TAF1 are represented by colored boxes with domain names noted below (23). HAT domain, histone acetyltransferase domain. Mutations found in USC are labeled at the top by red text. (*B*) Multiple sequence alignment across vertebrate and invertebrate species around the seven mutations found in USC. Mutation positions in human TAF1 are labeled in red at the top. Sequence aligned by Clustal W 2.0 (24).

Discussion

We report exome sequencing of a USC cohort five times larger than those recently reported (26, 27). The results define the genetic hallmarks of uterine serous cancer. We have found significantly increased mutation burden in 14 genes, including the previously identified and well-recognized cancer genes *TP53*, *PIKC3A*, *PPP2R1A*, *KRAS*, and *PTEN*, *FBXW7*, and *CDKN1A*.

The high frequency of single-nucleotide variants (SNVs) in *CHD4*, which was mutated in 19% of tumors and was the third most frequently mutated gene, was noteworthy. These mutations were diverse and predominantly at highly conserved positions from yeast to humans, and several have been previously shown to cause loss of function (19). Nonetheless, because CHD4 has many functional domains, it is possible that not all CHD4 functions are lost. Indeed, there appears to be clustering of mutations in particular domains, and there are seven copy-number gains that include *CHD4*.

Similarly, mutations in *TAF1*, a component of the core RNA polymerase II machinery, are found in 13% of tumors, with mutations at positions conserved throughout yeasts. Because TAF1 has diverse biochemical functions, the observed clustering of mutations in the HAT domain does not require that these mutations are null for all TAF1 functions. One known function of TAF1 is promotion of cyclin D expression; overexpression of cyclin D is itself known to promote cell cycle progression and proliferation and is frequently amplified in cancers (28). Notably, seven tumors had amplification of the segment of chromosome 11 containing *CCND1* (Fig. 4). It will be of interest to determine the biochemical and phenotypic consequences of *TAF1* mutations.

Several other genes show marginal statistical significance; these include known cancer genes such as *PTEN* and *CDKN1A* as well as *SPOP*, which targets proteins for ubiquitination via its MATH domain. SPOP has recently been shown to have clustered mutations in its MATH domain in prostate cancer (29); the two USC mutations are at different sites in the MATH domain. Additional genes not previously implicated in cancer include *HCFC1R1*, *CTDSPL*, *YIPF3*, and *FAM132A*. Further efforts will be required to determine the impact of these latter genes on USC and other cancers.

Our results demonstrate that somatic CNVs play a major role in the pathogenesis of USC, one that is likely at least as important as somatic point mutations. Interestingly, the most frequent CNV was a small deletion found in 68% of tumors affecting a short

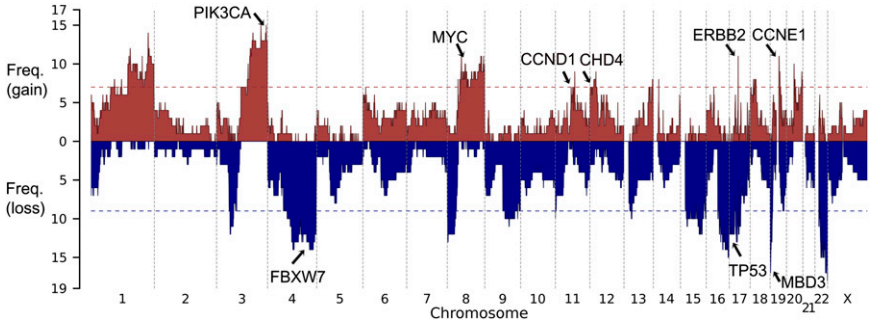

Fig. 4. Copy-number profile of 25 USC tumors. Frequency of copy-number gain (red) and copy-number loss (blue) are plotted along the genome. Horizontal dotted line, genome-wide significance level for CNV gain (red) and CNV loss (blue). Genes of interest in significant CNV peak regions are labeled.

segment of chromosome 19 that contains only 17 genes. Among these genes is *MBD3*, which is part of the same chromatin-remodeling complex—NuRD—as *CHD4*. Additionally, there were seven copy-number gains of the segment of chromosome 12 that includes *CHD4*, all of which were in samples with *MBD3* deletions. This complex deacetylates histones, repressing gene expression. Collectively, these findings add to the growing list of genes involved in chromatin remodeling that are mutated in cancer (30, 31).

CNV analysis of USC also identified frequent amplifications, including the well-known cancer genes *PIK3CA* (60%) and *ERBB2* (encoding HER2/neu; 44%) (32). ErbB2 overexpression has been previously reported to be associated with cancer cell proliferation, poor survival, and resistance to therapy in multiple human tumors including USC (5). Moreover, ErbB2 functions as an upstream regulator of the PIK3CA/AKT/mTOR-signaling pathway. These findings suggest common involvement of this pathway in USC and the possible utility of Food and Drug Administration-approved antibodies (i.e., trastuzumab, pertuzumab) or small molecule TK inhibitors used either alone or in combination with anti-mTOR, AKT, and/or PIK3CA active agents (5).

Another frequent somatic amplification (found in 44% of tumors) included a small segment of chromosome 19 that harbors *CCNE1*. *CCNE1* encodes cyclin E1 and is known to regulate the transition from the G1 phase to the S phase. High levels of *CCNE1* accelerate the transition through the G1 phase, and its accumulation is common in a number of cancers (33, 34). Most interestingly, CCNE1 degradation is mediated by binding to FBXW7 followed by ubiquitination via the SCF complex. Seventeen percent of USC harbor recurrent mutations in FBXW7 that abrogate CCNE1 binding (*SI Appendix*, Fig. S3) (35). These observations suggest that inhibition of *CCNE1* activity may have efficacy in patients harboring mutation in this pathway.

Collectively, our results implicate frequent mutations in several pathways in USC, including specific genes in DNA damage, chromatin remodeling, cell cycle, and cell proliferation pathways (Fig. 5). Analysis of correlation and anticorrelation of all possible pairs of significant mutations did not provide evidence of strong associations (*SI Appendix*, Fig. S8). Nonetheless, a large fraction of tumors shared mutations affecting genes in different pathways such as *TP53*, *PIK3CA*, *MBD3*, and *FBXW7*.

We also found that 9% of USC have a very high number of somatic mutations with many somatic mutations in mismatch repair and *POLE* genes. This distribution is distinct from the remainder (median 36 protein-altering mutations, all <100). Mutator phenotypes in the absence of germ-line mutations in mismatch repair genes have been seen in several other cancers (7, 8). These USC tumors are notable for being relatively frequent and for having a uniformly very high number of mutations, more than those seen in 90% of colon cancers with the mutator phenotype (36). Despite the remarkable somatic mutation burden, these tumors had no identified CNVs.

The establishment of 15 USC cell lines with different mutation profiles (*SI Appendix*, Fig. S9) provides the opportunity for in vitro assessment of whether a mutation profile is predictive of

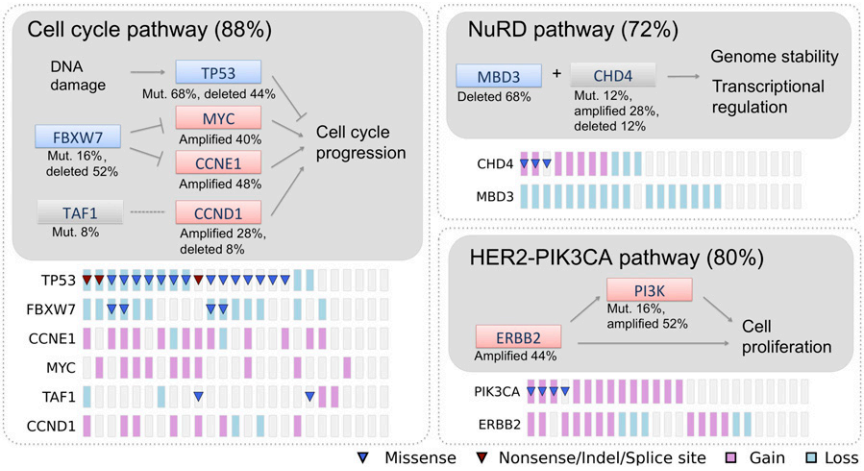

Fig. 5. Major altered pathways in USC. The altered percentages shown for genes and pathways come from the 25 matched tumors with CNV information. Genes are colored based on their activity in the pathway diagram. Pink, predicted activated; blue, predicted inactivated; gray, uncertain at this stage; lines with blunt end, inhibiting effect; lines with pointed end, promoting effect; dotted line, uncertain. Mutation and CNV status for each gene across the 25 samples are shown at the bottom following the pathway diagram.

drug response. For example, the finding that MMR-defective colorectal cancers may respond favorably to poly (ADP ribose) polymerase inhibitors (37) raises the question of whether the same may apply to USC with the hypermutator phenotype.

USC and high-grade serous ovarian carcinoma (HG-SOC) are histologically similar gynecological tumors characterized by a highly aggressive biologic behavior. Exome sequencing of HG-SOC has been recently reported by The Cancer Genome Atlas Research Network (38). TP53 was mutated in 95% of these cancers, with no other gene with somatic SNVs in more than 6% and only four (*BRCA1*, *BRCA2*, *CSMD3*, and *FAT3*) that were mutated in more than 3%. USC shows a lower frequency of *TP53* mutation (59%), 5 genes mutated in 13–23% of tumors, 10 more genes mutated in 3–10%, and no *BRCA1* or *BRCA2* mutations (*SI Appendix*, Table S6). These findings indicate substantial differences in the genetics of USC and HG-SOC.

While this article was in preparation, Khun et al. (26) and Le Gallo et al. (27) reported results from exome sequencing of a small number of USC (10 and 13 USC in discovery sets, respectively). These studies also found mutations in *TP53*, *PIK3CA*, *PPP2R1A*, *FBXW7*, *CHD4*, and *SPOP* genes, but were underpowered to detect other genes and CNVs that we report.

Our results define the genetic landscape of USC and identify specific pathways that are frequently mutated in these tumors. These findings will guide further research and targeted therapies against this highly aggressive variant of endometrial cancer.

Materials and Methods

Patients and Specimens. The study protocol was approved by the Yale University Human Investigation Committee. DNA and RNA were purified from tumors and normal tissues. Libraries were prepared as described (6). Additional details are provided in *SI Materials and Methods*.

Exome Sequencing and Analysis. Genomic DNA was captured on the NimbleGen 2.1M human exome array and subjected to 74 base paired-end reads on the Illumina HiSeq instrument as described (6). Sequence reads were mapped to the reference genome (hg18) using the ELAND program (6). For matched normal-tumor pairs, somatic mutations were called by comparing reference and nonreference reads from the matched pair as described (39). For unmatched tumors, SAMtools was used to call variant bases appended with quality scores. Additional details are provided in *SI Materials and Methods*.

Somatic CNV. CNVs were identified by comparing coverage depth of individual capture intervals (0.5-Mb bins) from tumor and normal samples after normalization for mean coverage depth of each exome. A permutation-based strategy was used to assess the significance of recurrent CNVs with a false-discovery-rate cutoff of 0.25 after Benjamini–Hochberg correction. Independent significant CNV peaks were called based on a GISTIC-like peel algorithm (40).

Quantitative RT-PCR. Quantitative RT-PCR was performed for all genes showing significant mutation burden using RNA from 15 USC cell lines. Additional details are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. This work was supported in part by National Institutes of Health (NIH) Grants R01 CA122728-01A4 and R01 CA154460-01A1 (to A.D.S.) and by Gilead Sciences Inc. This investigation was also supported by NIH Research Grant CA-16359 from the National Cancer Institute.

1. Siegel R, Naishadham D, Jemal A (2012) Cancer statistics, 2012. *CA Cancer J Clin* 62(1):10–29.
2. Bokhman JV (1983) Two pathogenetic types of endometrial carcinoma. *Gynecol Oncol* 15(1):10–17.
3. Hendrickson M, Ross J, Eifel P, Martinez A, Kempson R (1982) Uterine papillary serous carcinoma: A highly malignant form of endometrial adenocarcinoma. *Am J Surg Pathol* 6(2):93–108.
4. Lee KR, et al. (2003) Tumors of the ovary and peritoneum. *World Health Organization Tumours of the Breast and Female Genital System*, eds Tavassoéli F, Devilee P (IARC Press, Lyon, France), pp 113–145.
5. El-Sahwi KS, Schwartz PE, Santin AD (2012) Development of targeted therapy in uterine serous carcinoma, a biologically aggressive variant of endometrial cancer. *Expert Rev Anticancer Ther* 12(1):41–49.
6. Choi M, et al. (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA* 106(45):19096–19101.
7. Loeb LA (2011) Human cancers express mutator phenotypes: Origin, consequences and targeting. *Nat Rev Cancer* 11(6):450–457.
8. Yoshida R, et al. (2011) Concurrent genetic alterations in DNA polymerase proofreading and mismatch repair in human colorectal cancer. *Eur J Hum Genet* 19(3):320–325.
9. Greenman C, et al. (2007) Patterns of somatic mutation in human cancer genomes. *Nature* 446(7132):153–158.
10. Samuels Y, et al. (2005) Mutant PIK3CA promotes cell growth and invasion of human cancer cells. *Cancer Cell* 7(6):561–573.
11. Boguski MS, McCormick F (1993) Proteins regulating Ras and its relatives. *Nature* 366(6456):643–654.
12. Welcker M, Clurman BE (2008) FBW7 ubiquitin ligase: A tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat Rev Cancer* 8(2):83–93.
13. Hao B, Oehlmann S, Sowa ME, Harper JW, Pavletich NP (2007) Structure of a Fbw7-Skp1-cyclin E complex: Multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. *Mol Cell* 26(1):131–143.
14. Shih IeM, et al. (2011) Somatic mutations of PPP2R1A in ovarian and uterine carcinomas. *Am J Pathol* 178(4):1442–1447.
15. Talbi S, et al. (2006) Molecular phenotyping of human endometrium distinguishes menstrual cycle phases and underlying biological processes in normo-ovulatory women. *Endocrinology* 147(3):1097–1121.
16. Pleasance ED, et al. (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 463(7278):191–196.
17. Mansfield RE, et al. (2011) Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. *J Biol Chem* 286(13):11779–11791.
18. Hauk G, McKnight JN, Nodelman IM, Bowman GD (2010) The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol Cell* 39(5):711–723.
19. Flaus A, Martin DM, Barton GJ, Owen-Hughes T (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* 34(10):2887–2905.
20. Lai AY, Wade PA (2011) Cancer biology and NuRD: A multifaceted chromatin remodelling complex. *Nat Rev Cancer* 11(8):588–596.
21. Wassarman DA, Sauer F (2001) TAF II 250: A transcription toolbox interaction partners. *J Cell Sci* 114(Pt 16):2895–2920.
22. Hilton TL, Li Y, Dunphy EL, Wang EH (2005) TAF1 histone acetyltransferase activity in Sp1 activation of the cyclin D1 promoter. *Mol Cell Biol* 25(10):4321–4332.
23. Kloet SL, Whiting JL, Gafken P, Ranish J, Wang EH (2012) Phosphorylation-dependent regulation of cyclin D1 and cyclin A gene transcription by TFIID subunits TAF1 and TAF7. *Mol Cell Biol* 32(16):3358–3369.
24. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.
25. Wada C, Kasai K, Kameya T, Ohtani H (1992) A general transcription initiation factor, human transcription factor IID, overexpressed in human lung and breast carcinoma and rapidly induced with serum stimulation. *Cancer Res* 52(2):307–313.
26. Kuhn E, et al. (2012) Identification of molecular pathway aberrations in uterine serous carcinoma by genome-wide analyses. *J Natl Cancer Inst* 104(19):1503–1513.
27. Le Gallo M, et al. (2012) Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat Genet* 44(12):1310–1315.
28. Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL (2011) Cyclin D as a therapeutic target in cancer. *Nat Rev Cancer* 11(8):558–572.
29. Barbieri CE, et al. (2012) Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* 44(6):685–689.
30. Turcan S, et al. (2012) IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* 483(7390):479–483.
31. Wang K, et al. (2011) Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat Genet* 43(12):1219–1223.
32. Yarden Y, Sliwkowski MX (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2(2):127–137.
33. Geisen C, Moroy T (2002) The oncogenic activity of cyclin E is not confined to Cdk2 activation alone but relies on several other, distinct functions of the protein. *J Biol Chem* 277(42):39909–39918.
34. Donnellan R, Chetty R (1999) Cyclin E in human cancers. *FASEB J* 13(8):773–780.
35. Thompson BJ, et al. (2007) The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. *J Exp Med* 204(8):1825–1835.
36. Cancer Genome Atlas Network (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487(7407):330–337.
37. Takahashi M, Koi M, Balaguer F, Boland CR, Goel A (2011) MSH3 mediates sensitization of colorectal cancer cells to cisplatin, oxaliplatin, and a poly(ADP-ribose) polymerase inhibitor. *J Biol Chem* 286(14):12157–12165.
38. Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* 474(7353):609–615.
39. Choi M, et al. (2011) K+ channel mutations in adrenal aldosterone-producing adenomas and hereditary hypertension. *Science* 331(6018):768–772.
40. Beroukhim R, et al. (2007) Assessing the significance of chromosomal aberrations in cancer: Methodology and application to glioma. *Proc Natl Acad Sci USA* 104(50):20007–20012.