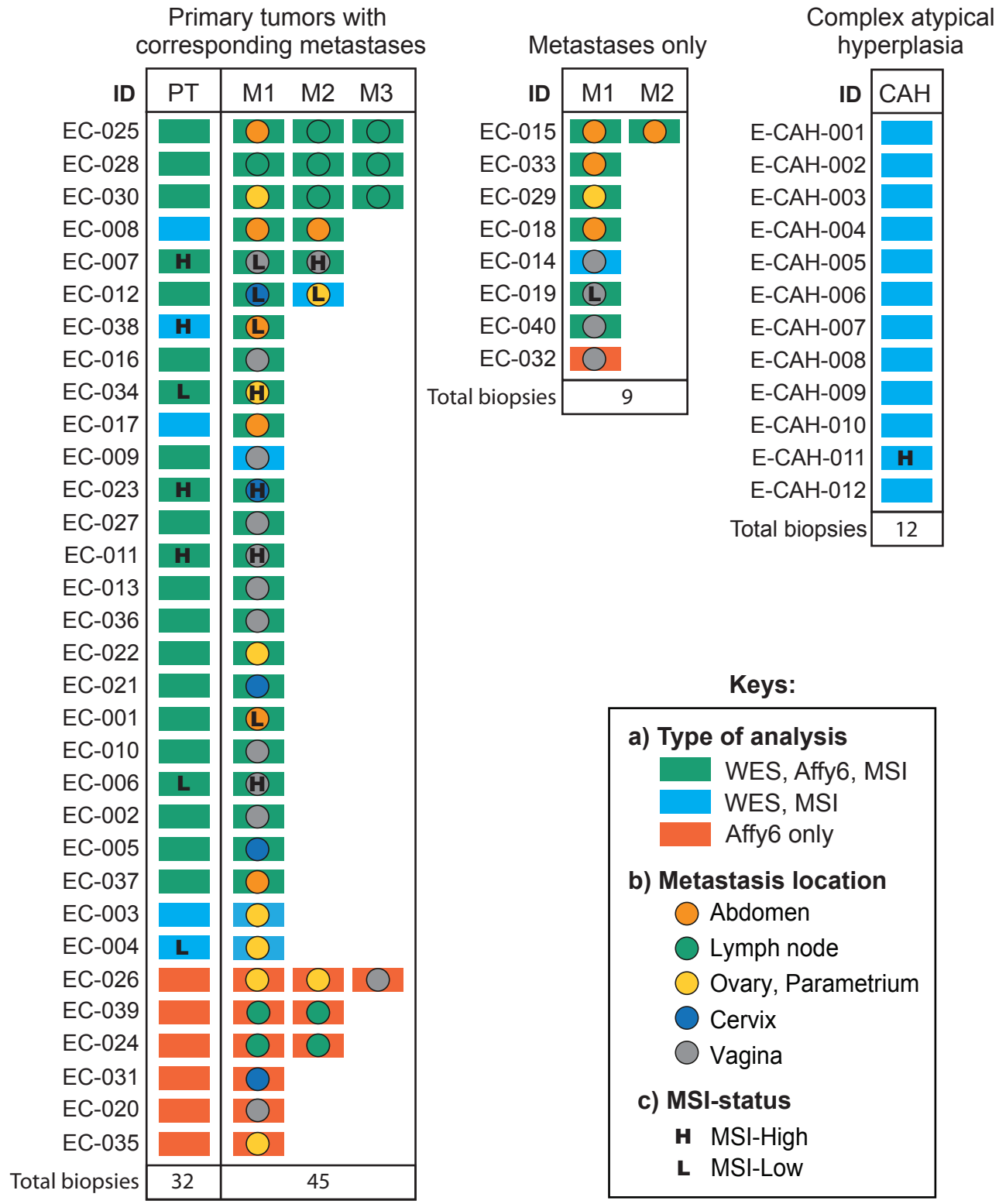
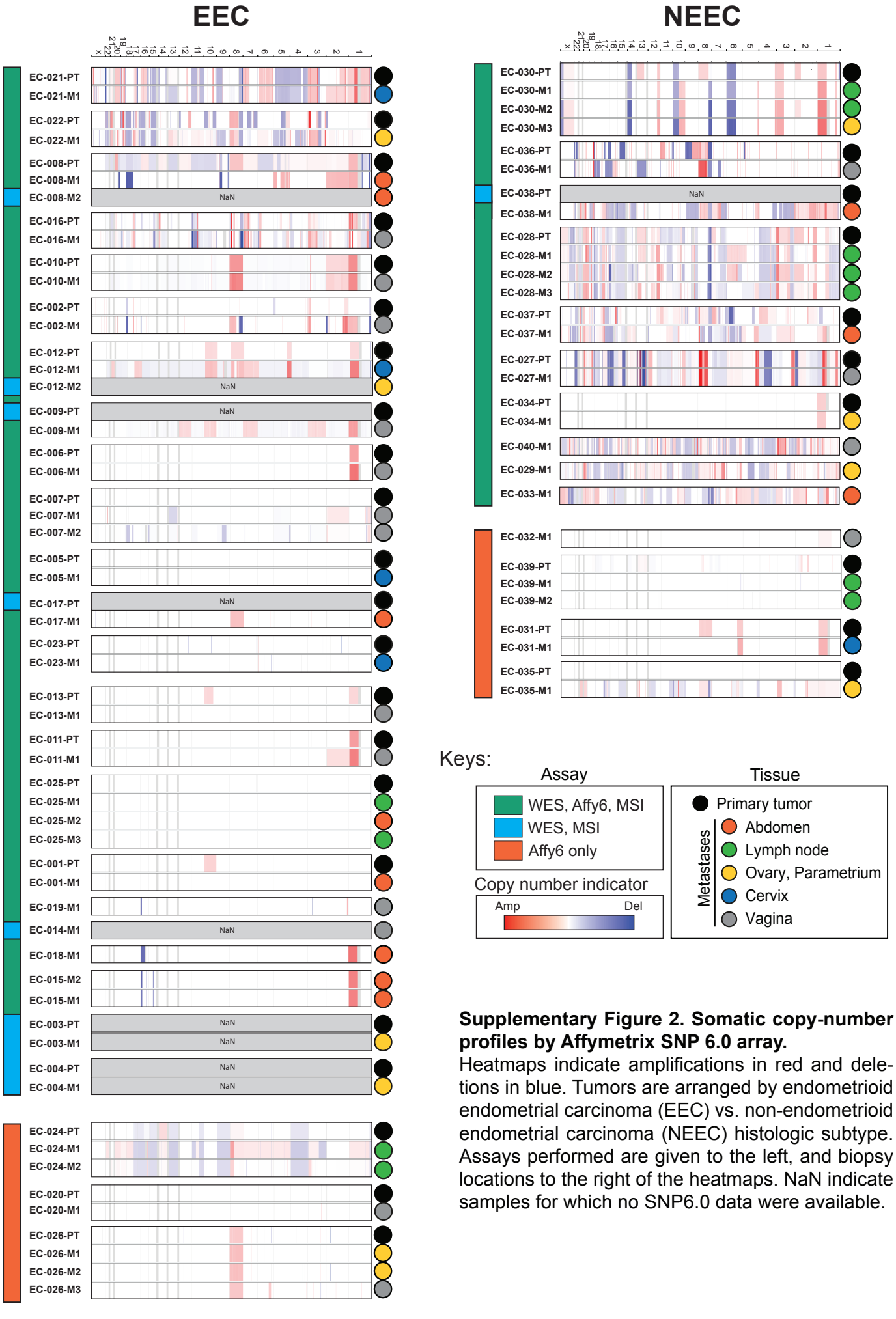
Supplementary Figure 1



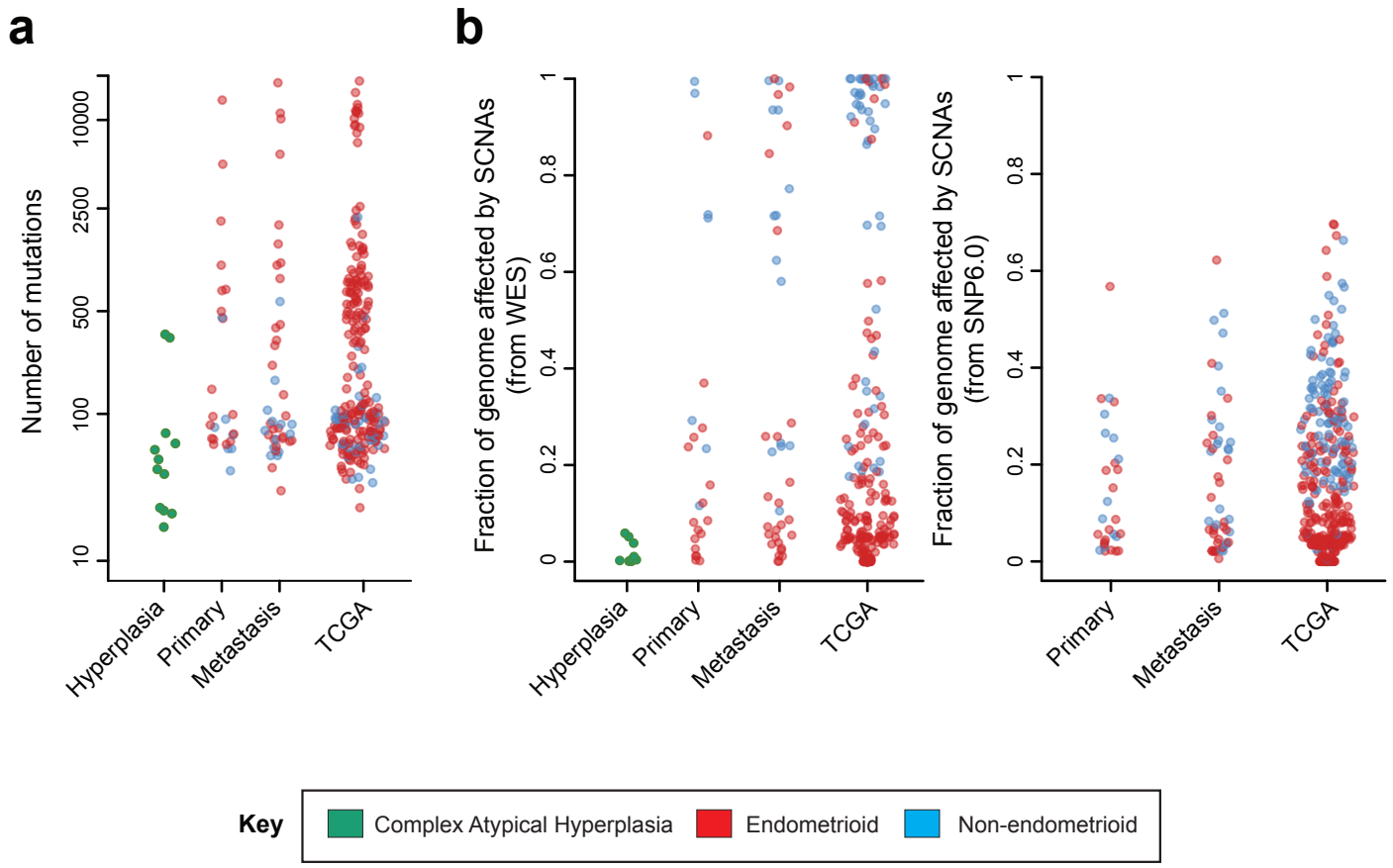
Supplementary Figure 1. Overview of sample set and data generated.
Samples from primary tumors (PT) with paired metastases from the same patient (left), metastases without matching primary tumors (middle), and complex atypical hyperplasias (CAH; right). Information about type of analysis (key a): Most biopsies underwent Whole-Exome Sequencing (WES). Copy-number profiles were determined with Affymetrix SNP 6.0 arrays, in large overlapping with WES samples. Location of abdominal metastasis (M) for the samples are indicated by dots with colours corresponding to key b. For some patients multiple anatomically distinct metastases were obtained. Microsatellite instability status by a 7-marker panel was performed on samples analyzed by WES, as indicated by key c (No indication; MS-stable). The total number of biosies is 98, taken from 52 patients.

Supplementary Figure 2



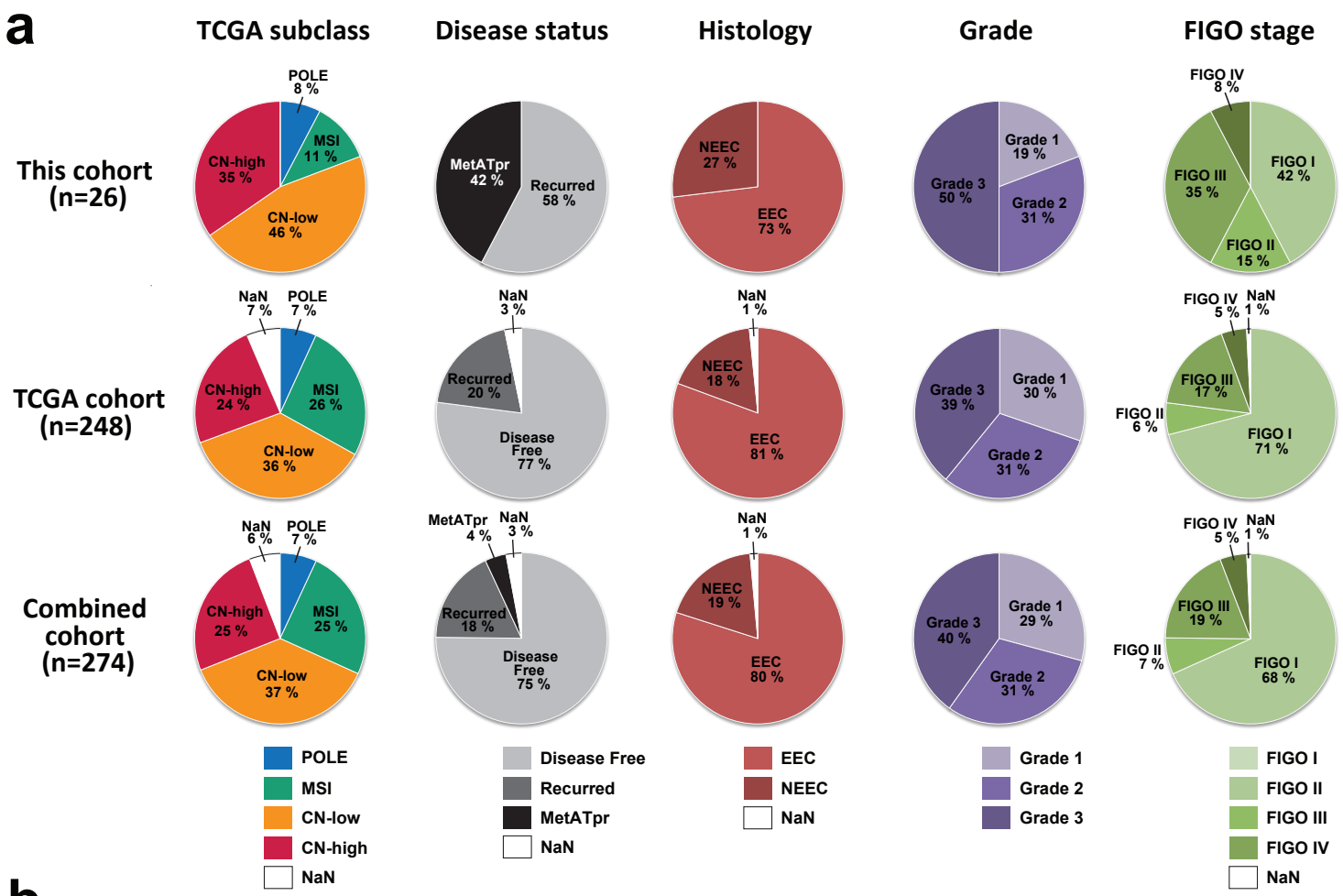
Supplementary Figure 2. Somatic copy-number profiles by Affymetrix SNP 6.0 array. Heatmaps indicate amplifications in red and deletions in blue. Tumors are arranged by endometrioid endometrial carcinoma (EEC) vs. non-endometrioid endometrial carcinoma (NEEC) histologic subtype. Assays performed are given to the left, and biopsy locations to the right of the heatmaps. NaN indicate samples for which no SNP6.0 data were available.

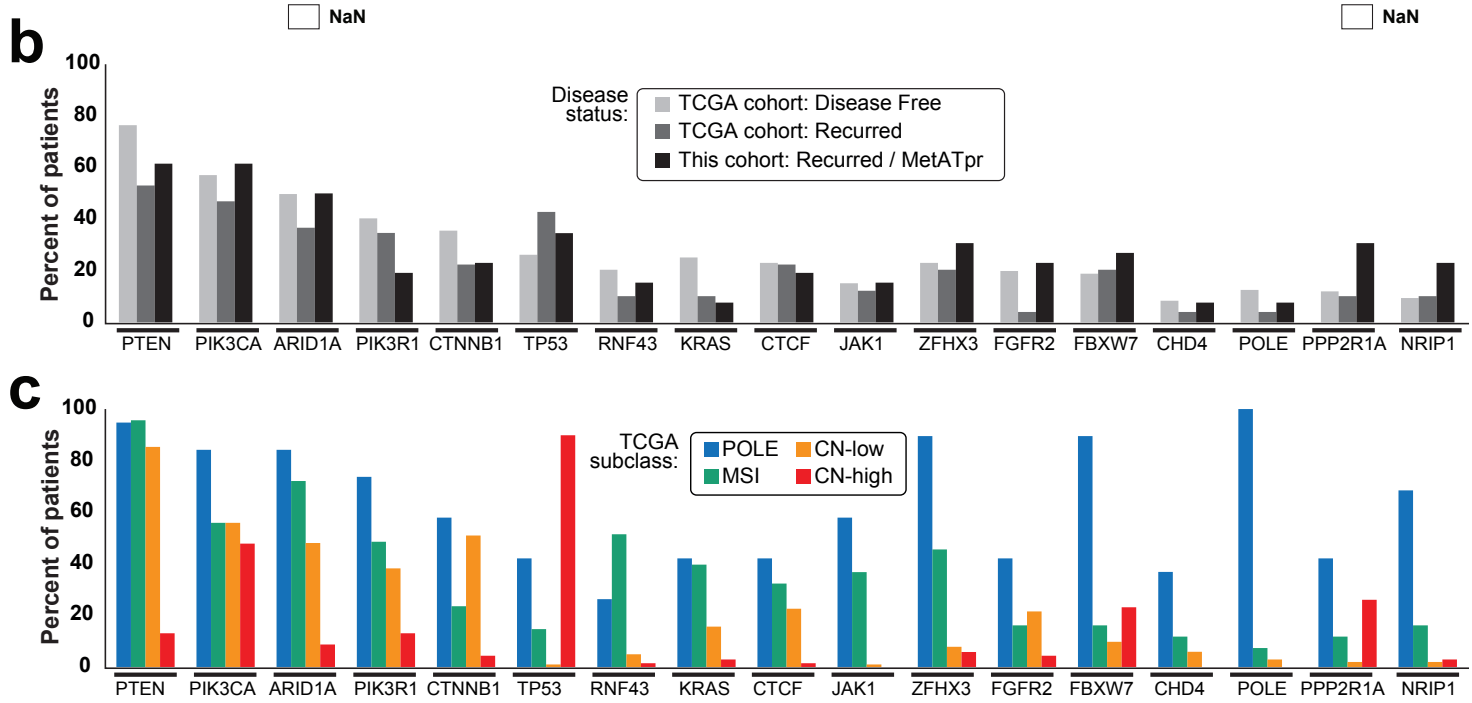
Supplementary Figure 3



Supplementary Figure 3. Somatic alteration rates by tissue type.
a, Number of exonic mutations detected in Endometrial complex atypical hyperplasias, primary tumors and metastases. Data from TCGA primary tumors are displayed on the right.
b, Fraction of the genome that is affected by SCNAs in endometrial complex atypical hyperplasias, primary tumors and metastases as determined by Whole Exome Sequencing (WES, left) or SNP6 array analysis (right). For WES-analyzed samples, allelic copy number profiles were computed and the fraction of the genome affected by SCNAs was defined as regions where there was not one copy of each allele. In this analysis, samples affected by whole genome doubling obtain higher values for the fraction of the genome affected by SCNAs than an analysis of total relative copy-number would yield. For samples analyzed with SNP 6.0 arrays, the fraction of the genome affected by SCNAs was defined as the fraction of the genome where the total relative copy level (normalized to a median of 2 per sample) was below 1.6 or above 2.5.

Supplementary Figure 4

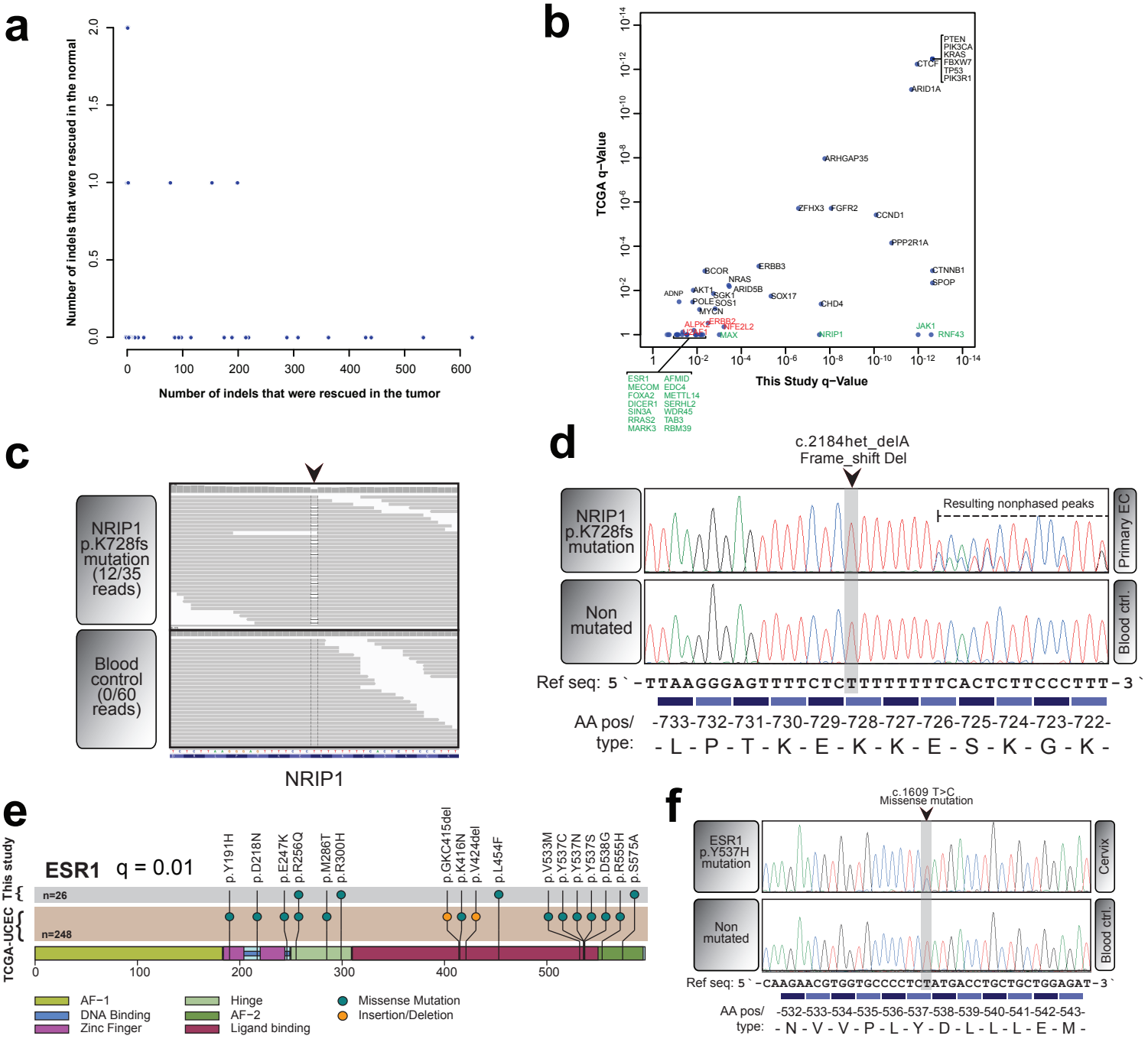




Supplementary Figure 4. Key features of this cohort compared to TCGA cohort

a, Composition of molecular subgroups, clinical variables and histology of this cohort compared to that of the Uterine Corpus Endometrioid Carcinoma (UCEC) of TCGA. This cohort included more of the copy-number (CN) high subgroup tumors, non-endometrioid tumors (NEEC), and more of cases with higher grade and FIGO stages. All cases in this cohort progressed to metastatic disease or were metastatic at the time of primary surgery (MetATpr). NaN indicates samples that were not assigned to a TCGA subclass or key variable. **b,** Percentage of patients harboring mutations in driver genes by disease status, comparing this cohort to that of TCGA. **c,** Percentage of patients with mutations in driver genes by TCGA molecular subgroups in the combined cohort (n=274).

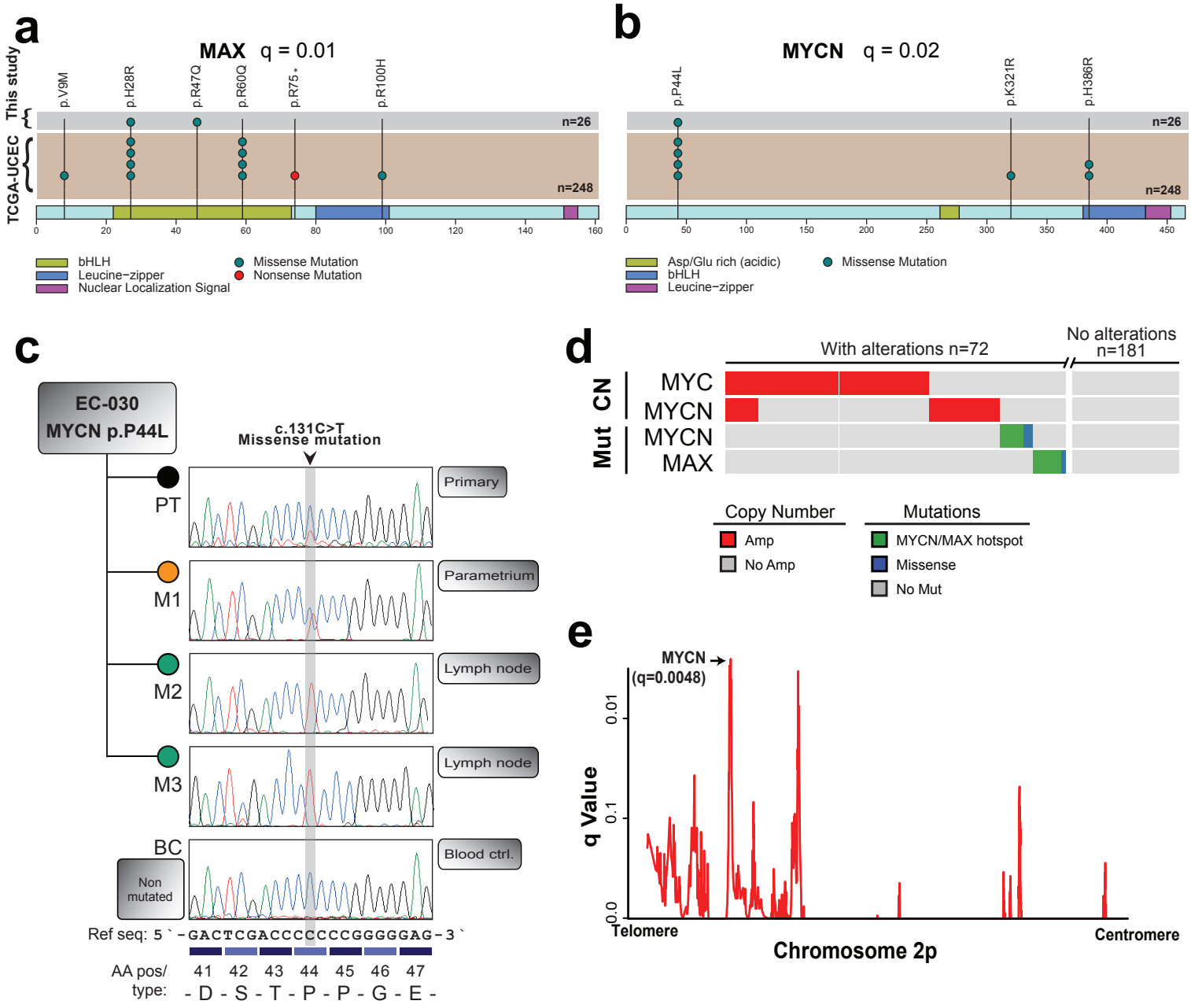
Supplementary Figure 5



Supplementary Figure 5. Indel rescue and estrogen pathway mutations.

a, Accuracy of modification to indeLocator to “rescue” indels. Insertions/Deletions meeting modified IndeLocator criteria were called in tumor-normal pairs (x-axis). To test the specificity of these calls, tumor/normal labels were swapped and the number of called indels was recorded for each normal-tumor pair (y-axis). A mean of 0.16 mutations were called in the swap, indicating a low false positive rate, whereas a mean of 62 mutations were rescued in the tumor-normal comparisons. **b,** Comparison of significance analysis results to TCGA. The 49 significantly mutated genes in our study (x-axis, purple dots) were plotted by their FDR q-value towards that obtained from the TCGA cohort (y-axis). Genes along the x-axis are significantly mutated in this study but not in that of the TCGA. Red genes indicate novel significant genes in our study that are also recurrently mutated in other cancer types. Green genes indicate novel significant genes in our study that are not recurrently mutated in other cancer types. **c,** The NRIP1 p.K728fs mutation shown based on WES data for a vaginal metastasis in case EC-011 visualized by Intergrated Genomics Viewer. **d**, Sanger sequencing chromotograms validating NRIP1 p.K728fs mutation in an external sample set of primary endometrial cancer. Two of 37 samples in the validation cohort carried the mutation. **e,** Stick plot of mutations observed in ESR1 by data source. “This study” refers to the 26 primary cancers included in our cohort. Mutations cluster in the ligand binding domain of the estrogen receptor. The Y537 mutations have previously been observed to confer tamoxifen resistance in breast cancers. However, we observed one patient in our cohort who demonstrated a p.Y537H mutation without any prior hormonal therapy as shown in **f**. Chromotogram of Sanger sequencing validate this ESR1 p.Y537H hotspot mutation in this individual (EC-012).

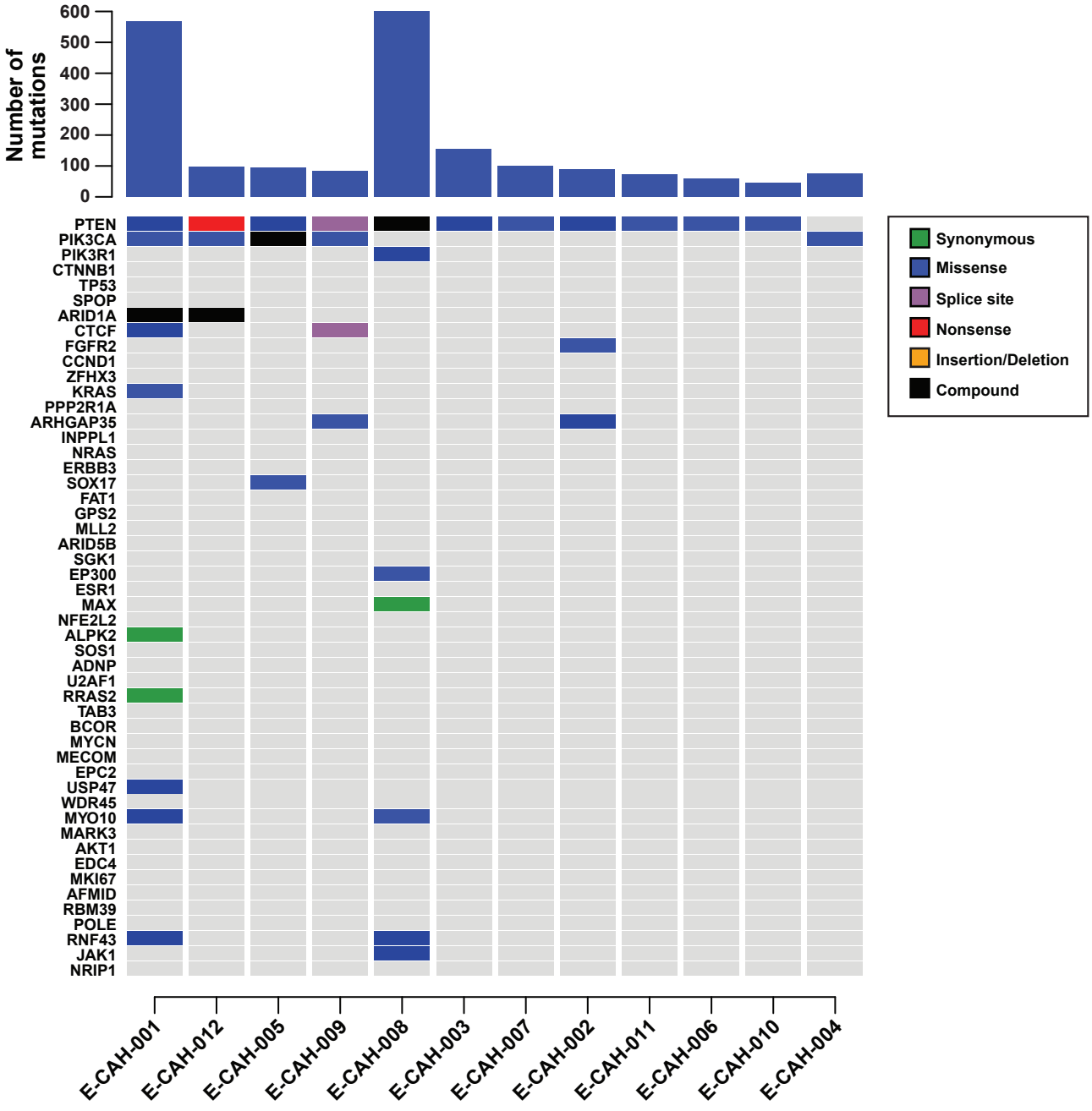
Supplementary Figure 6



Supplementary Figure 6. Frequency of co-alteration of MYC pathway copy number alterations and mutations.

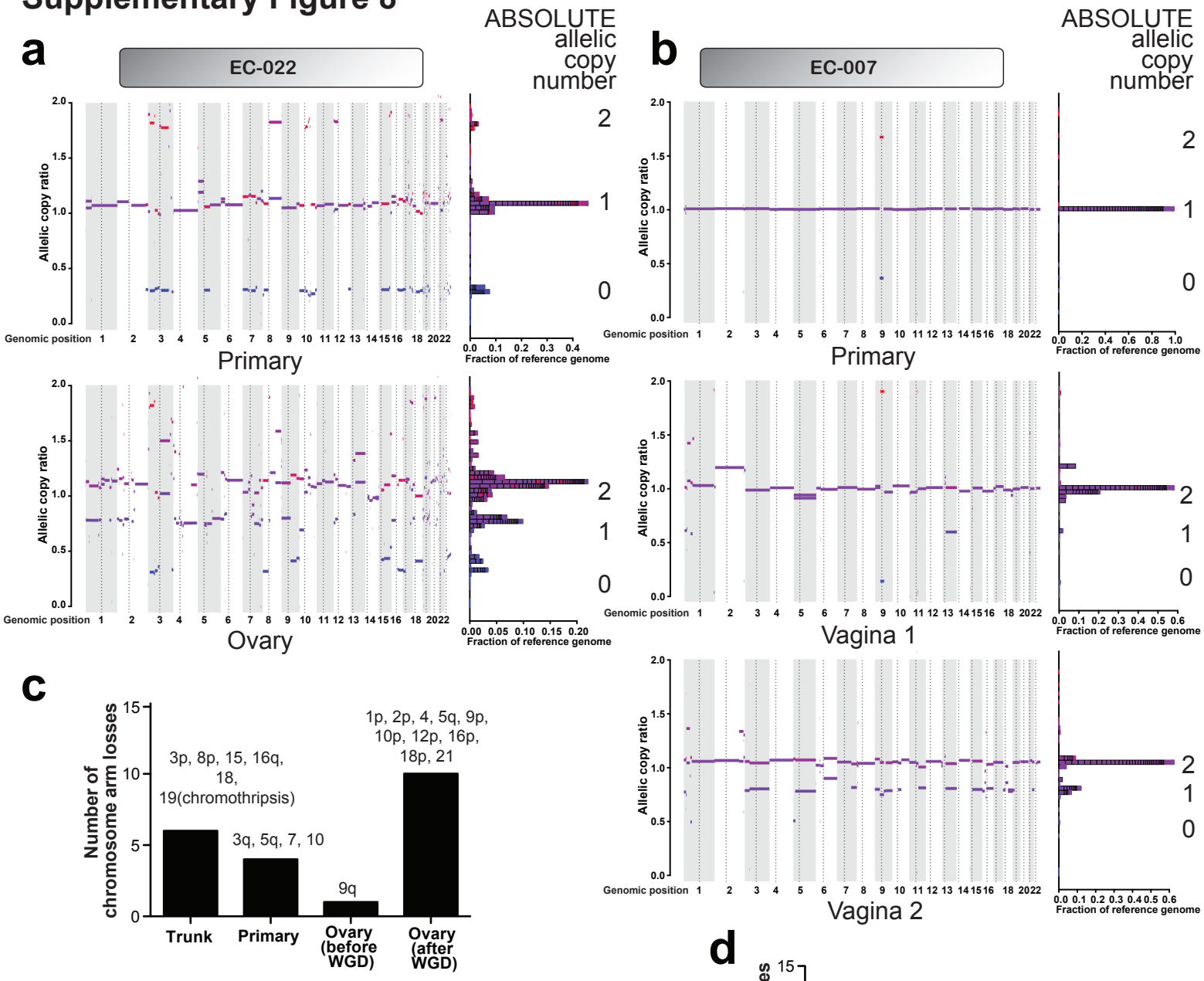
a, Stick plot of mutations in MAX by data source. We identified two hotspots in MAX, p.H28R (n=5) and p.R60Q (n=4). **b,** Stick plot of mutations in MYCN by data source. Two recurrent sites were identified, P44L (n=5) and H386R (n=2). **c,** Sanger sequencing validation of p.P44L mutation of MYCN in case EC-030, all tumors. **d,** All samples in our cohort and TCGA were analyzed to determine the patterns of coalteration between significantly altered MYC pathway members. Both MYCN and MAX mutations were mutually exclusive with other MYC pathway alterations. None of these anti-correlations reached statistical significance (p=0.29 and 0.36 for MYCN and MAX, respectively). **e,** MYCN is the only gene in a GISTIC peak in endometrial cancer. GISTIC q-values are plotted on the y-axis, with genomic position on the x-axis.

Supplementary Figure 7



Supplementary Figure 7. Mutations in endometrial complex atypical hyperplasias.
(top) Number of exomic mutations detected in endometrial hyperplasias (bottom). Genes that were significantly mutated in primary tumors are depicted in the row of the comutation plot. Only PTEN and PIK3CA were mutated in every hyperplasia.

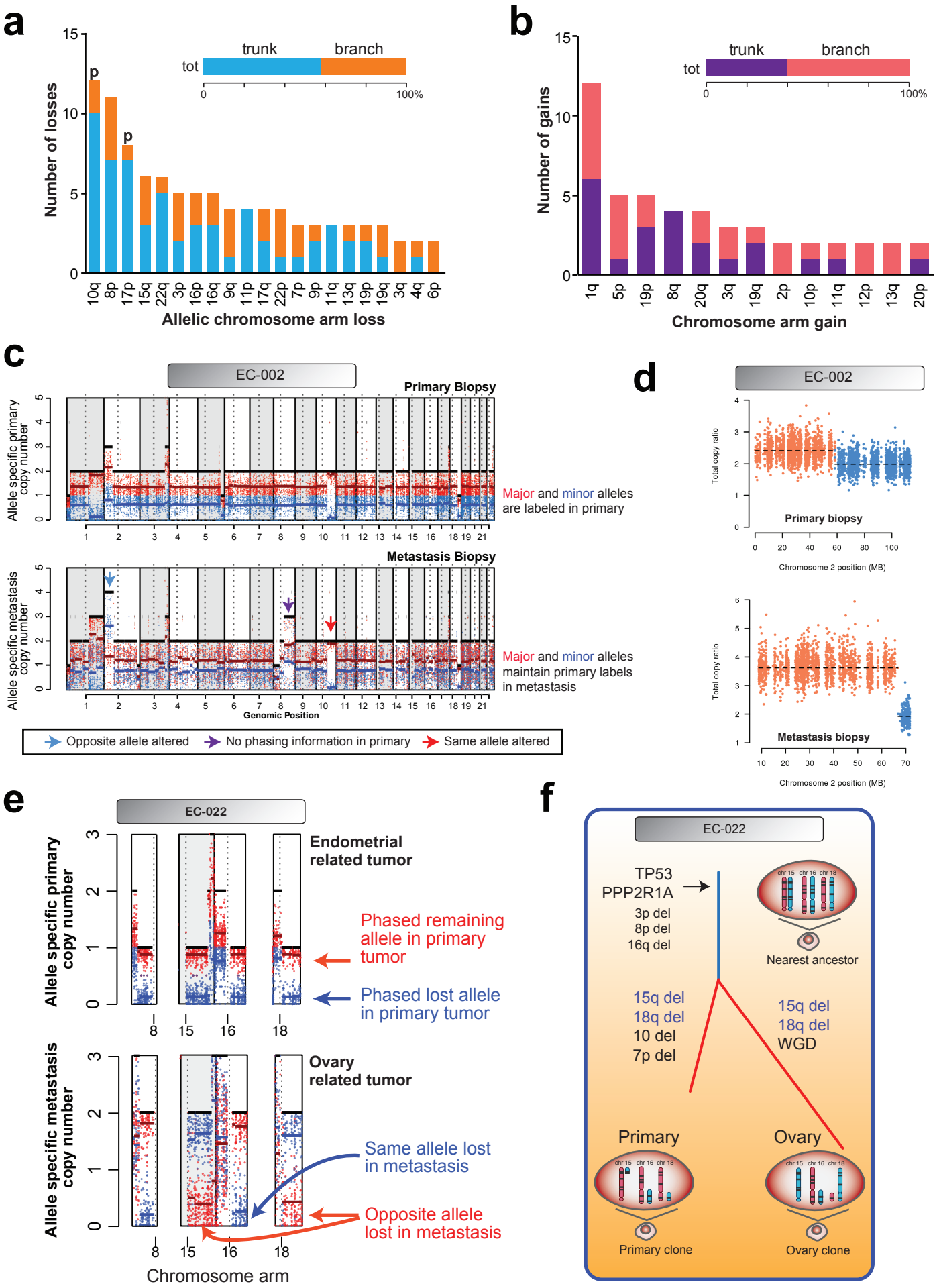
Supplementary Figure 8



Supplementary Figure 8. Whole genome doubling (WGD) is followed by elevated rates of localized structural events

a-b, Allelic copy number plots for phylogenies with non-truncal WGD for cases EC-022 and EC-007. While the primary biopsy for patient EC-007 harbored few SCNAs that would allow accurate determinate of biopsy purity, we were able to assign a purity of 67% from the variant allelic fractions of 1,027 mutations. At this purity, we were adequately powered to detect clonal SCNAs if they occured. We detected neither WGD nor arm-level copy number events. **c-d,** Number of arm-level events occuring in each biopsy in a) and b) respectively. All single-copy losses were deemed to have occured after WGD. Losses of two copies of the same allele could have occured prior to or after WGD; we called all of these prior to WGD to be conservative. (**Lower part, d)** We performed flow cytometry to confirm that the primary in EC-007 was diploid. The second vaginal biopsy was obtained 12 months after the resection of the first vaginal metastasis and may therefore represent either recurrence of the first metastasis or an independent metastasis. Due to this ambiguity, EC-007 was excluded from multi-metastasis analyses.

Supplementary Figure 9



Supplementary Figure 9. Phylogenetics of somatic copy number alterations.

a-b, Phylogenetic placement of chromosome arm losses **(a)** and gains **(b)**. Pi (p) indicates bias towards truncal alteration (binomial test $p<0.05$).

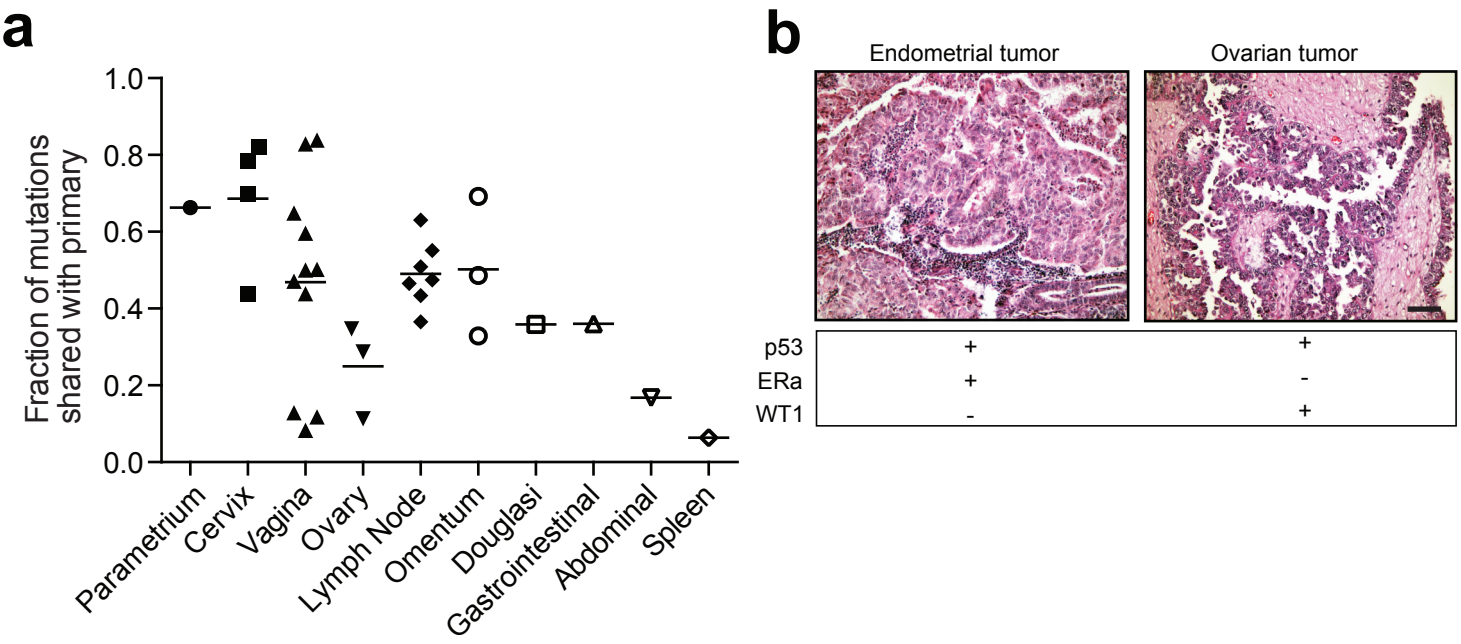
c, Haplotype tracking across tumor biopsies for EC-002 reveals convergent SCNAs. Copy-number adjusted allelic fractions from germline heterozygous sites are plotted for the PB (top) and MB (bottom). All variants at germline heterozygous sites are labeled in the PB by whether more reads of either allele are present (major allele, red). These labels were maintained in the MB (the sites labeled as red in the PB are red in the MB) and the same analysis was performed in the MB. Where a genomic segment is deleted or amplified, a haplotype is revealed. Often, the same haplotype is gained/lost in both biopsies (eg 10q in this example, red arrowhead). Occasionally, genomic regions undergo convergent copy number gains/losses in which the opposite allele is gained/lost in the MB compared to the PB (e.g. 2p in this example, blue arrowhead).

d, The convergent SCNAs identified by allelic analysis in EC-002 also displayed different breakpoints in total copy ratio data.

e, Allele-specific copy-number profiles at germline heterozygous sites in primary and metastasis of EC-022.

f, Schematic outlining process of convergent loss of chromosome arms 15q and 18q indicated by data in **e**. WGD = Whole Genome Doubling.

Supplementary Figure 10



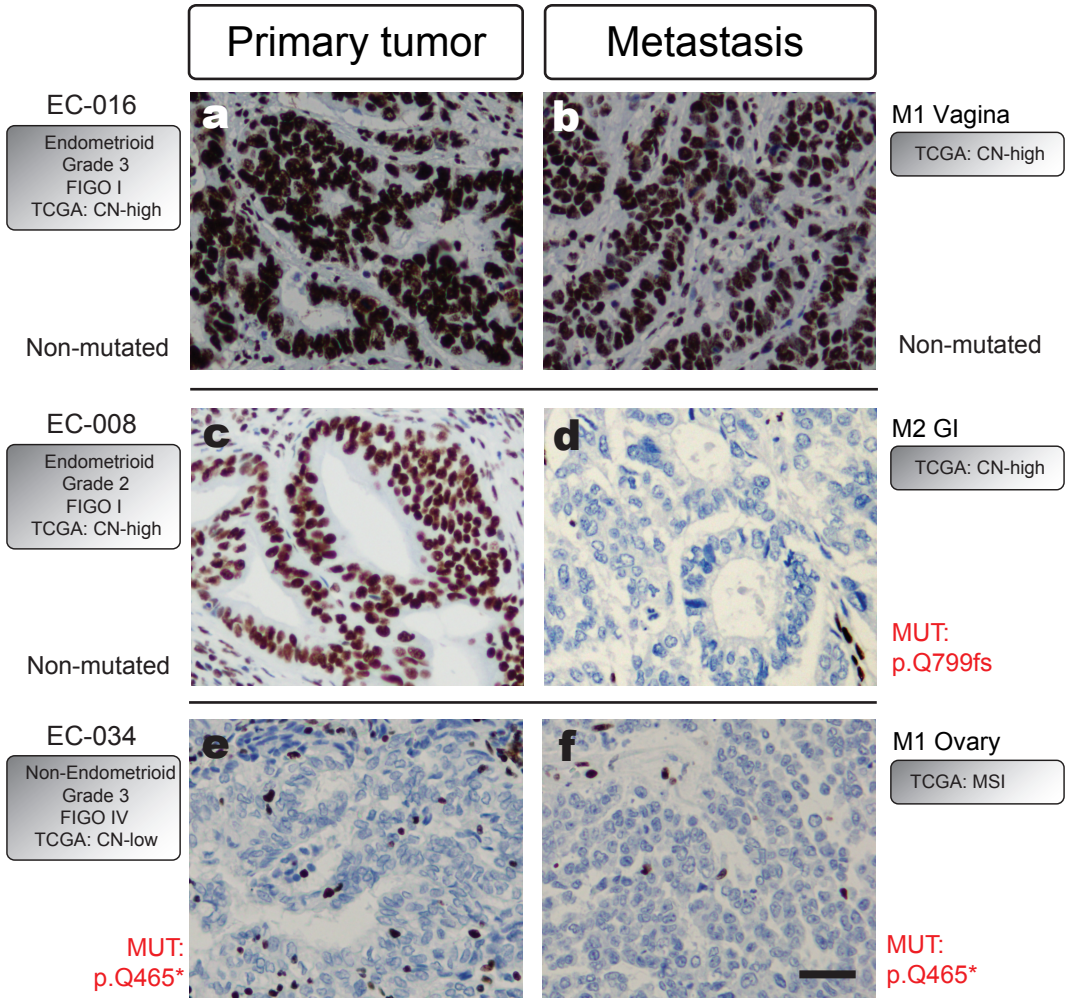
Supplementary Figure 10. Trends in phylogenetic similarity.

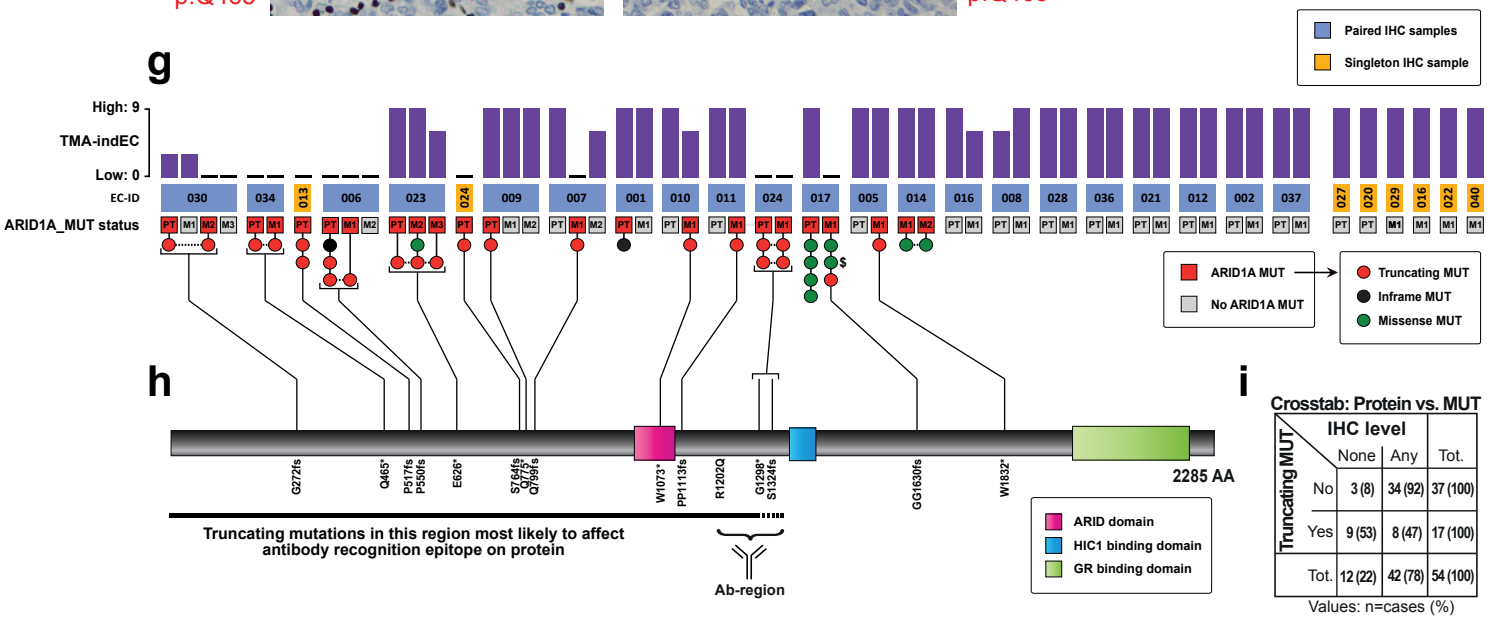
a, The fraction of mutations each metastasis shared with its paired primary appeared to decrease with increasing distance to the primary tumor. Although not statistically significant (p=0.27). For purposes of statistical analysis, sites were grouped into four categories:

close mets= Cervix, Parametrium
gynecologic mets= Vagina,Ovary
semi-gynecologic metastases= Douglasi, Lymph Node
abdominal metastases= Omentum, Gastrointesinal, Spleen, Abdominal

b, EC-022 was re-evaluated by the patologist during this study and considered to represent two independent synchronous cancers based on morphology and immunohistochemical markers, but the two biopsies shared 76 mutations, indicating a single cancer. (Left) Tissue (HE) section from the uterine cavity showed glandular morphology consistent with endometrioid adenocarcinoma. Immunohistochemistry (IHC) indicated the tumor to be positive for p53 and estrogen hormone receptor (ER alpha) and negative for WT1 (Wilms` Tumor 1). (Right) Tissue (HE) section from the right ovary showed a papillary morphology consistent with a high grade serous adenocarcinoma. Immunohistochemistry indicated that this tumor was positive for p53 and WT1 and negative for ER alpha. Bar = 50 mm.

Supplementary Figure 11





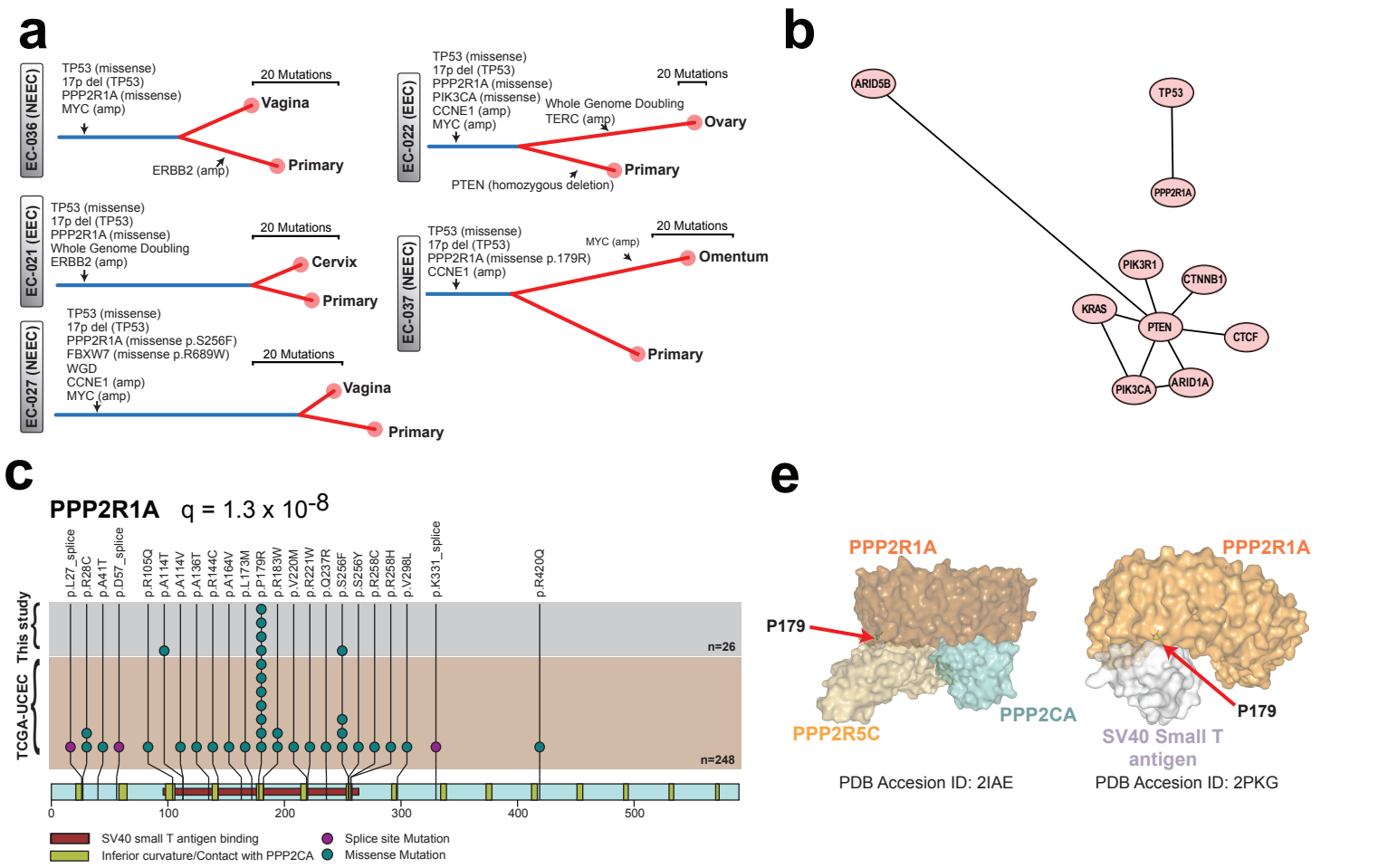
Truncating MUT	IHC level: None	Any	Tot.
No	3 (8)	34 (92)	37 (100)
Yes	9 (53)	8 (47)	17 (100)
Tot.	12 (22)	42 (78)	54 (100)

Values: n=cases (%)

Supplementary Figure 11. Immunohistochemical staining of ARID1A

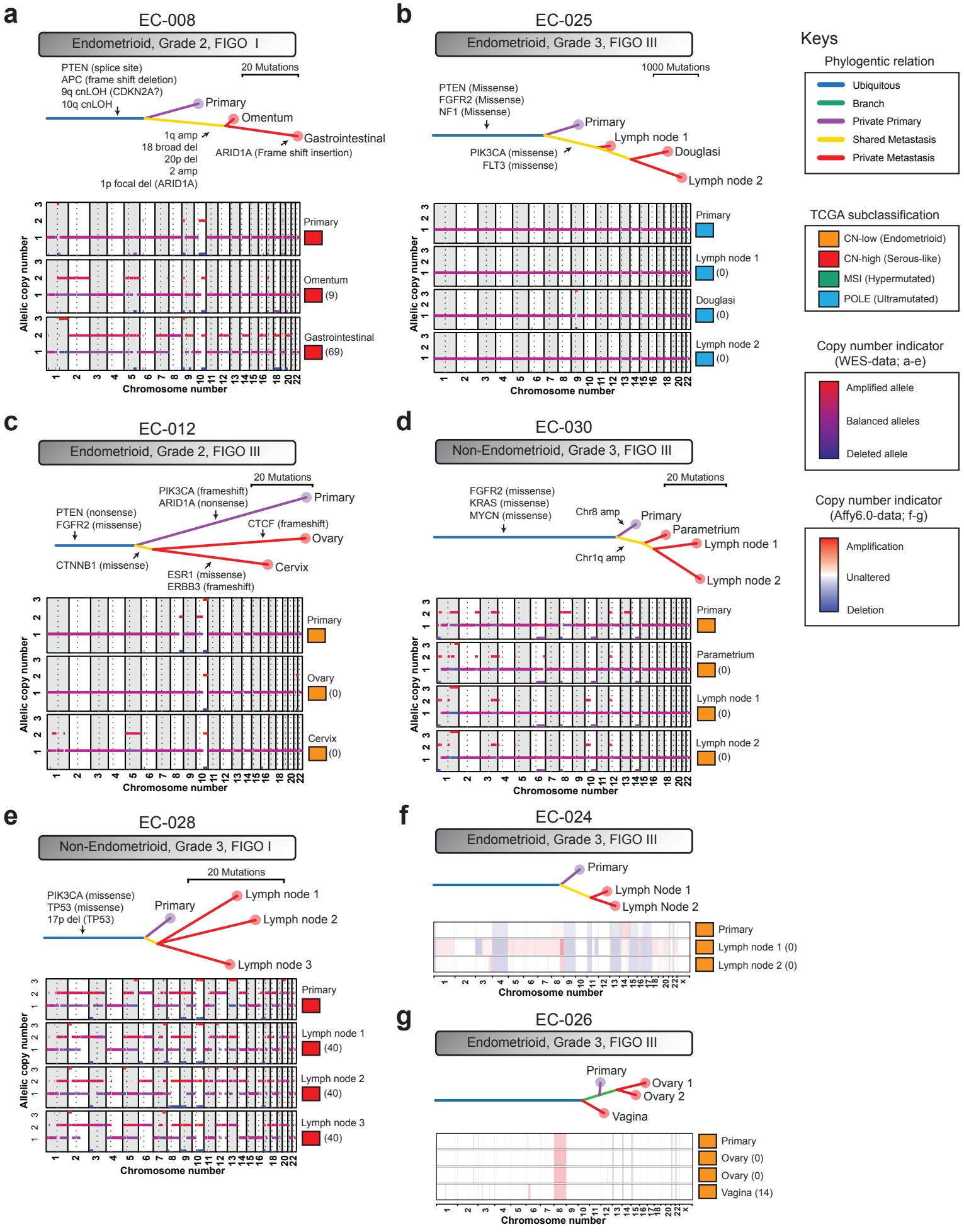
a-f, Immunohistochemical (IHC) staining of ARID1A paraffin embedded sections of selected endometrial cancer (EC) cases, demonstrating positive homogenous immunoreactivity (a-b), heterogenous immunoreactivity (c-d) and complete loss of ARID1A protein immunoreactivity (e-f) in corresponding primary tumor and metastasis. Mutations of *ARID1A* are denoted in red for each biopsy if present. Bar=20μm. **g**, Semiquantitative IHC data from TMAs evaluating anti-ARID1A antibody immunoreactivity, along with mutation status. Positions for truncating mutations are shown. Dollar sign indicate a missense mutation within potential Ab-epitope. **h,** Schematic view of the ARID1A protein with putative epitope target region of the antibody indicated (Abcam182561). **i,** Crosstab summary. Protein levels were assessed as present (TMA index 9-1) or absent (index 0). Samples with no ARID1A protein (index 0), are associated with presence of truncating mutations (p<0.001).

Supplementary Figure 12



Supplementary Figure 12. PPP2R1A and TP53 mutations cooperate to promote transformation. **a,** Phylogenies with recurrent PPP2R1A mutations. **b,** Network model of gene comutation. Distances between nodes reflect negative log q-values for correlated mutations. **c,** Stick plot of PPP2R1A mutations by data source. Due to the large number of likely passenger mutations in POLE/ultramutated tumors, these samples were removed. **d,** Comutation plot of TP53 mutation, PPP2R1A mutation, TP53 loss of heterozygosity (LOH), histology (blue: NEEC, grey: EEC) and whole genome doubling (WGD, purple). PPP2R1A hotspot mutations are positively correlated with TP53 loss ($p= 6x10^{-8}$). **e,** The P179R mutation falls at the binding site of PPP2R1A to PPP2R5C (left).The P179R mutation overlaps the SV40 small T binding site to PPP2R1A (right)

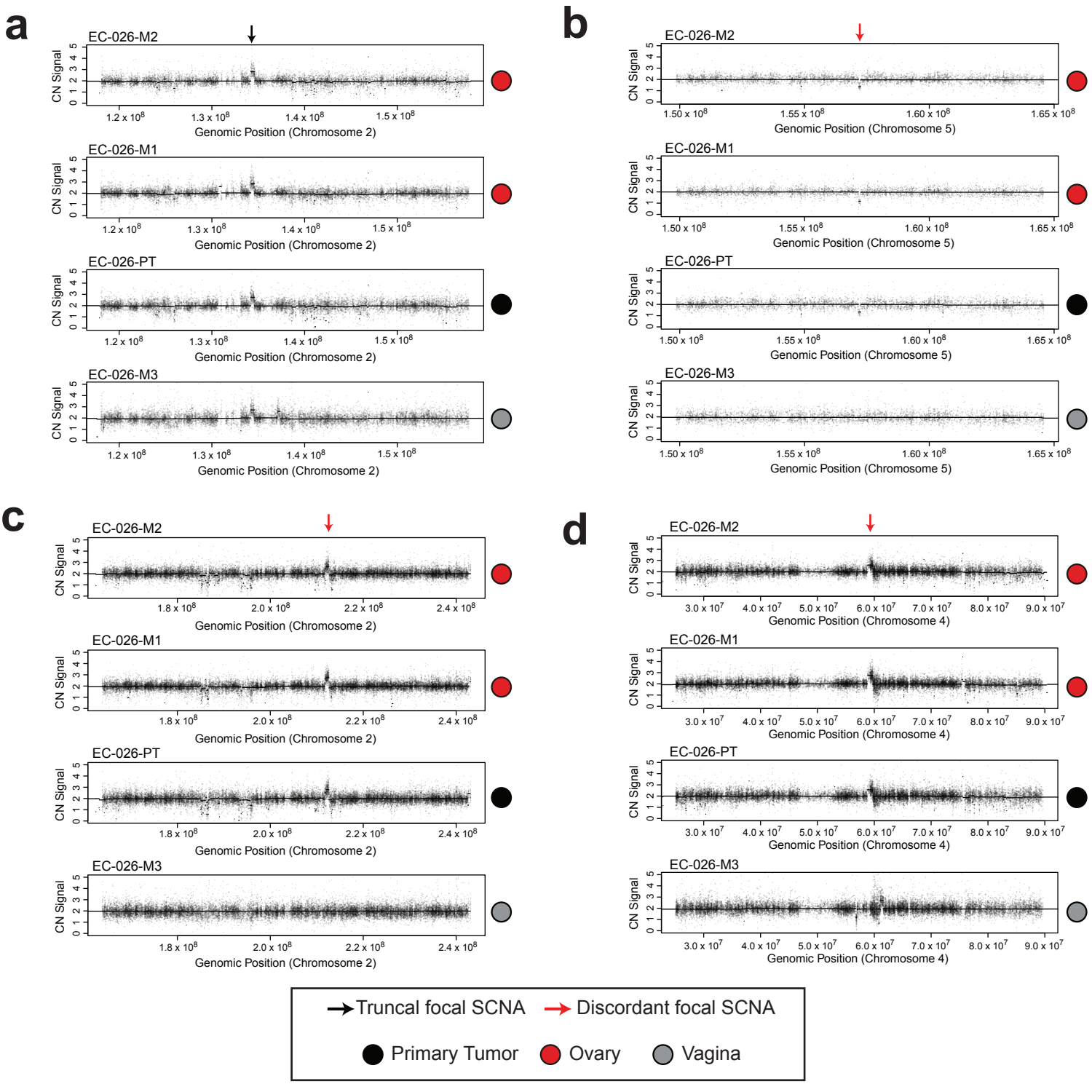
Supplementary Figure 13



Supplementary Figure 13. Phylogenetic trees and copy number alterations for trees with multiple metastases.

a-e, (Top) WES-derived tumor phylogenies from the combined mutation and copy number alteration profiles from individual biopsies. The trees display the evolutionary relationship between primary-metastasis samples for a single patient: alterations ubiquitous to all biopsies are blue, yellow indicates mutations shared by all metastases, and red and purple indicates mutations specific (private) to only one metastasis or the primary biopsy, respectively. **a-e, (Bottom)** Allelic copy number profiles. Molecular classifications for all primary and metastatic lesions are denoted in colored squares to the right of each copy number panel in line with the TCGA uterine cancer classification scheme, followed by time (months) between resection of primary and metastatic lesions. Histological type, grade and FIGO stage is given for each individual. Further clinical details are in Supplementary Table 1. Scale bars for mutations are given for each case. Non-silent driver mutations are labeled in trees a, -d, and e. Due to the extremely high burden of mutations in b and c, only mutations with amino acid changes reported as recurrent in COSMIC are shown. **f-g,** Phylogenies of multi-met tumors profiled with Affymetrix SNP 6.0 Arrays.

Supplementary Figure 14

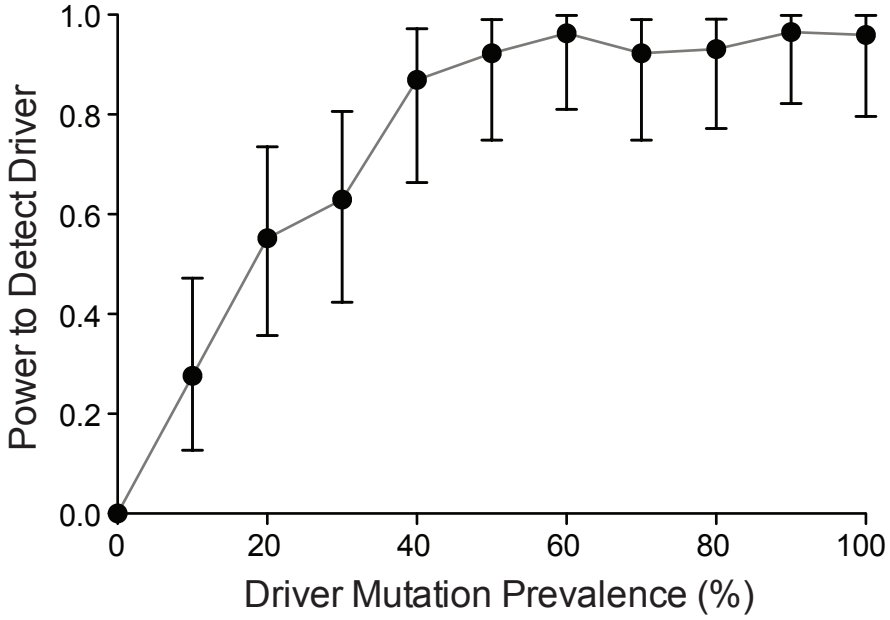


Supplementary Figure 14. Focal copy number alterations demonstrate separate origins of metastasis in patient EC-026.

a, A focal SCNA on chromosome 2 that is shared across all biopsies demonstrates that all biopsies were of comparable purity and similarly able to detect SCNAs.

b-d, Focal SCNAs that are shared by two ovarian metastases and the primary tumor but absent in the biopsy from a vaginal metastasis. These SCNAs indicate that the ovarian biopsies and the primary biopsy have a shared common ancestor that did not give rise to the vaginal biopsy.

Supplementary Figure 15



Supplementary Figure 15. Power to detect metastasis specific drivers in this study.
Hypothetical drivers of metastasis were spiked into the set of mutations detected only in metastasis biopsies that were used to search for metastasis-specific drivers. The fraction of hypothetical drivers was varied and the rate at which MUTSIGCV2 recovered the driver as significant ($q<0.25$) is plotted above. Error bars reflect binomial 95% confidence intervals on the fraction of hypothetical drivers recovered.

Supplementary Figure 16

Supplementary Figure 16. Mutations assigned to primary subclone that seeded metases in EC-007
a, 2D confidence intervals on the cancer cell fraction (CCF) of mutations assigned to the cluster of mutations that became clonal in the metastasis but was subclonal in the primary for patient EC-007. The CCF of the mutation in the primary biopsy is plotted on the x-axis and the CCF of the mutation in the metastasis biopsy (Vagina 1) is plotted on the y-axis. The outermost ellipse represents the 95% confidence interval. Candidate driver mutations are highlighted in red.