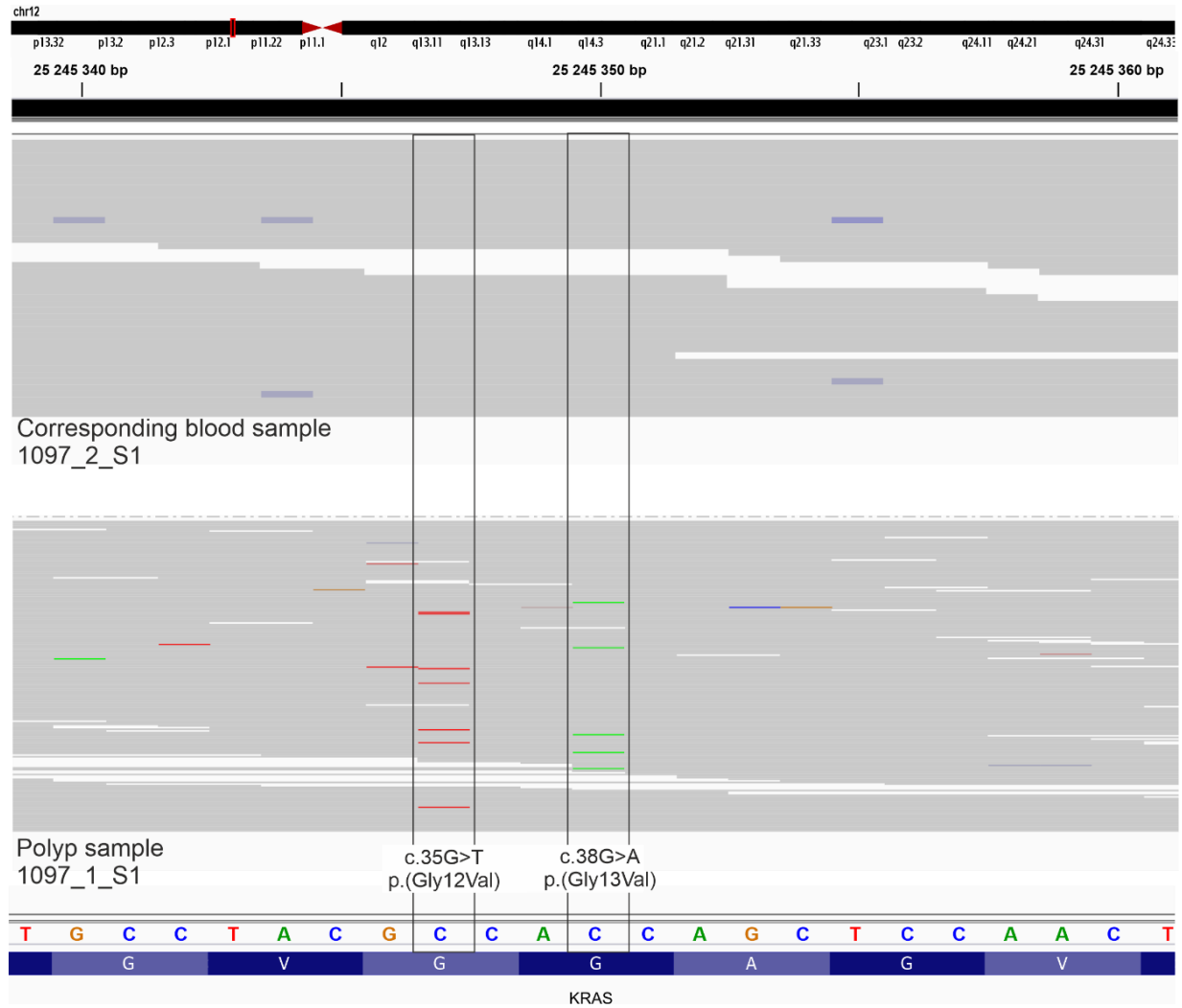
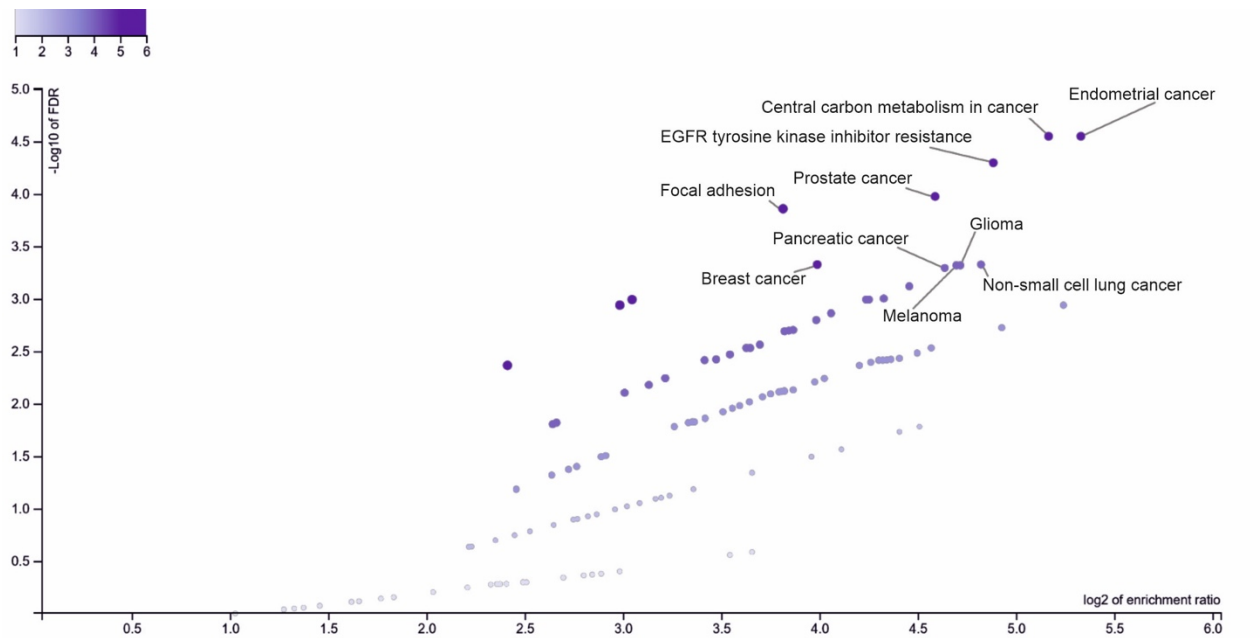


Supplementary Material

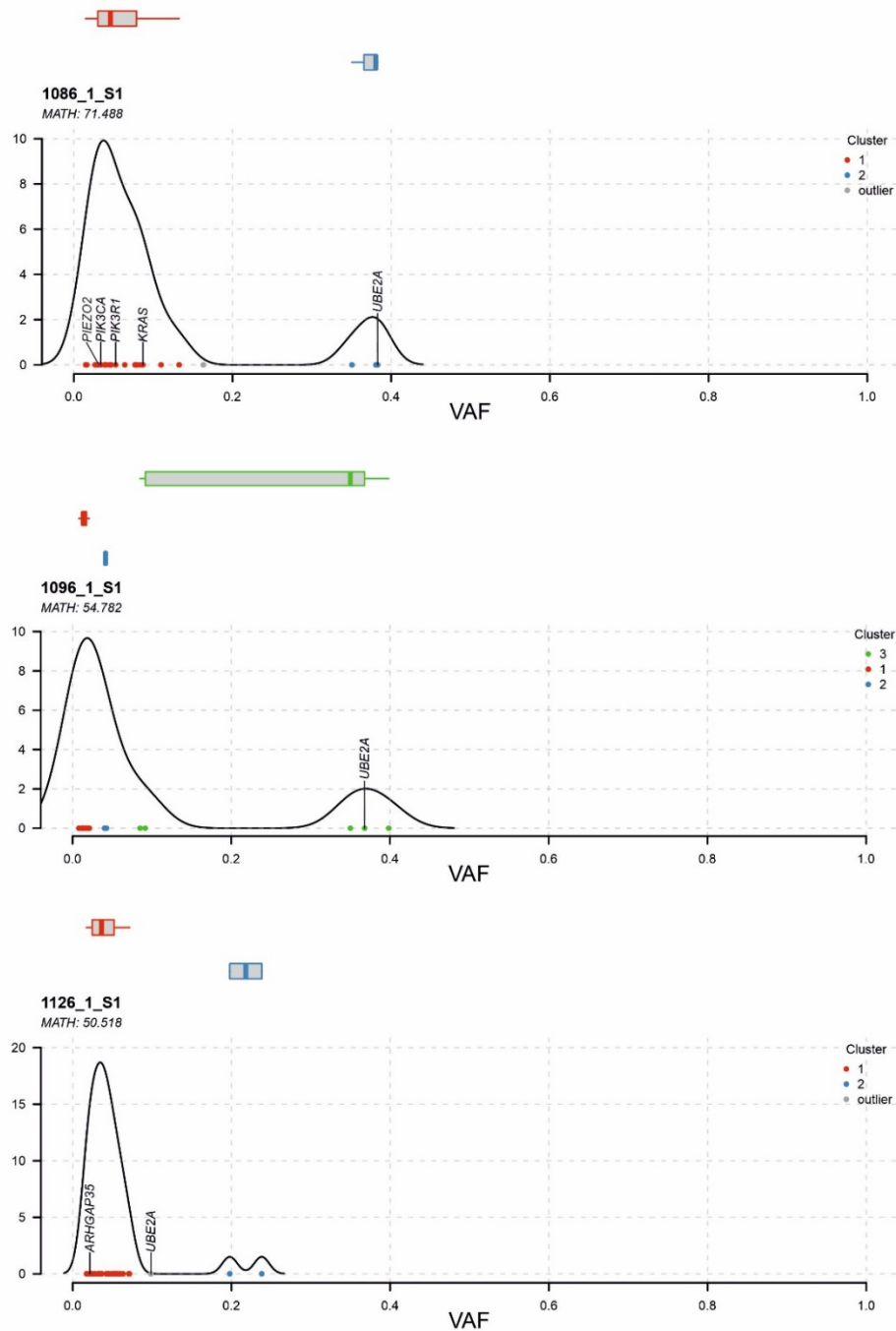
Genomic landscape of endometrial polyps



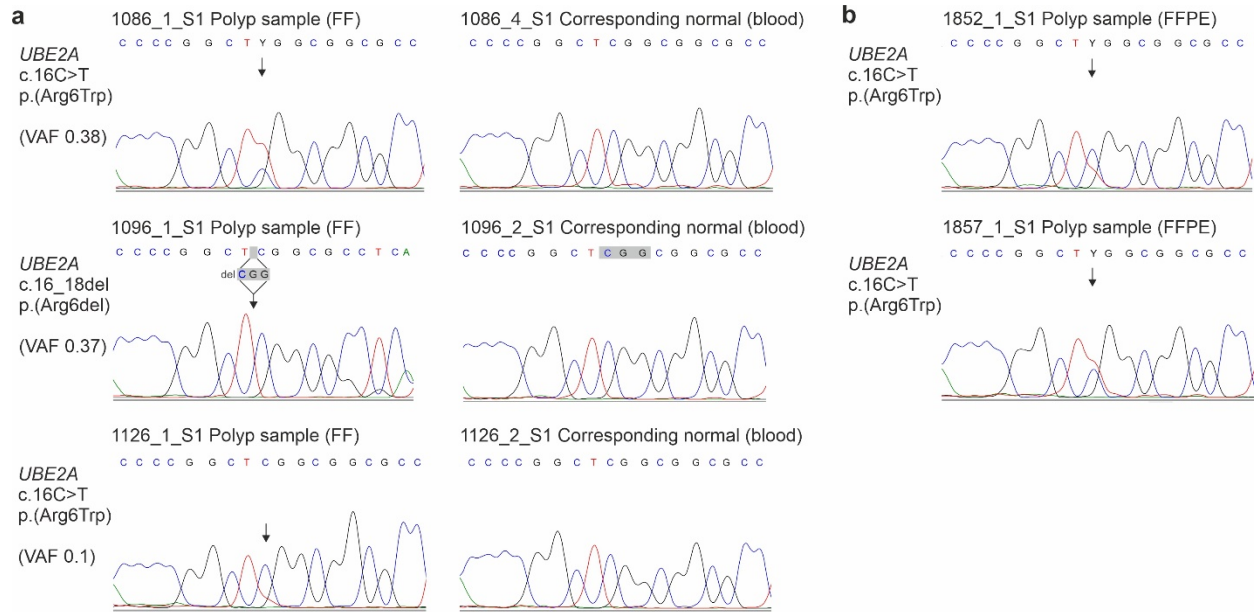
Supplementary Figure S1. Two *KRAS* hotspot mutations in one endometrial polyp. We identified one endometrial polyp to harbor two somatic hotspot mutations in *KRAS*: c.35G>T, p.(Gly12Val) and c.38G>A, p.(Gly13Val). The mutations were detected in different reads, indicating that they are present in different cell populations or on different chromosomes.



Supplementary Figure S2. Volcano plot of the most significantly mutated KEGG pathways in endometrial polyps. Pathway enrichment analysis using recurrently mutated genes identified 10 KEGG pathways that were significantly enriched ($q < 0.05$). Endometrial cancer was the most significant pathway.

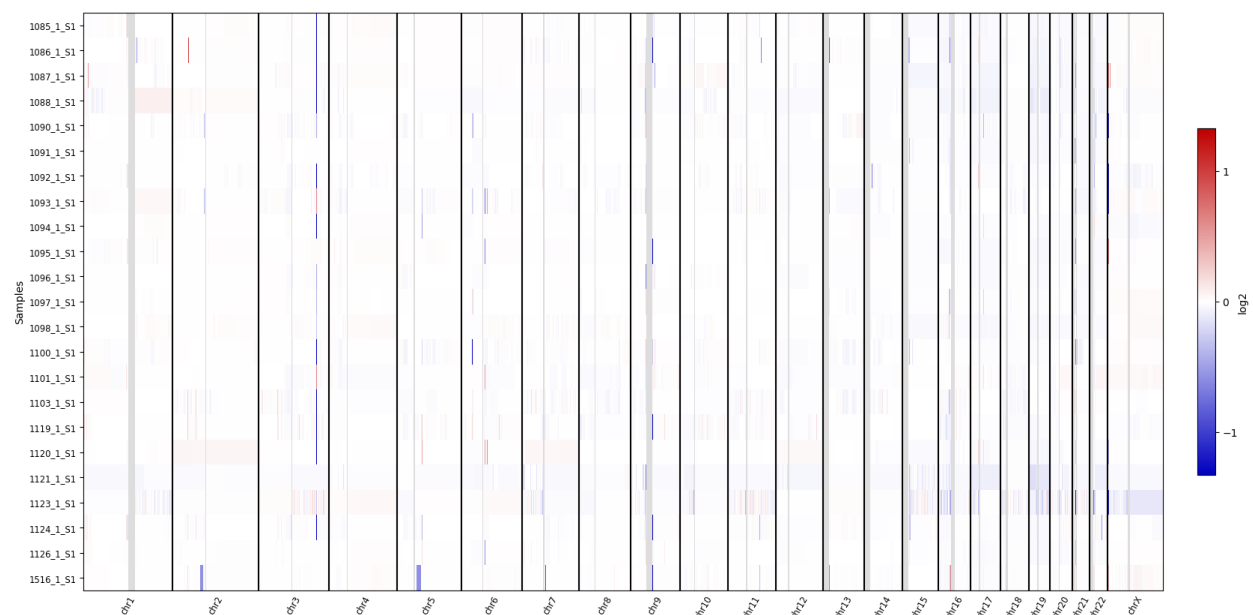


Supplementary Figure S3. Tumor heterogeneity analysis of somatic mutations identified in three endometrial polyps with a *UBE2A* mutation. Tumor heterogeneity analysis revealed that the *UBE2A* mutations were present in a more clonal cell population than the smallest cell population clone. Each mutation cluster has a different color and outliers are colored gray. VAF: Variant allele fraction. Genes mutated in more than three samples are marked in the plots.



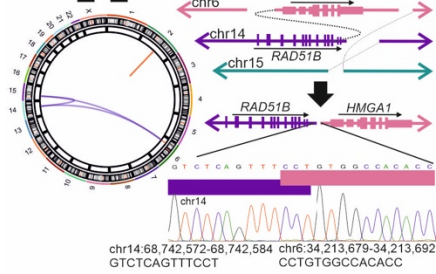
Supplementary Figure S4. Chromatograms showing a *UBE2A* mutation in five endometrial polyps.

A) Sanger sequencing confirmed a somatic *UBE2A* mutation in 3 out of 23 (13%) endometrial polyps. Two samples harbored the same missense mutation c.16C>T, p.(Arg6Trp), and the third sample an in-frame 3-bp deletion c.16_18del, p.(Arg6del). **B)** Sanger sequencing of 54 FFPE endometrial polyps revealed two additional samples with the missense mutation c.16C>T, p.(Arg6Trp). VAF: Variant allele fraction identified in high-throughput sequencing, FFPE: Formalin-fixed paraffin embedded.

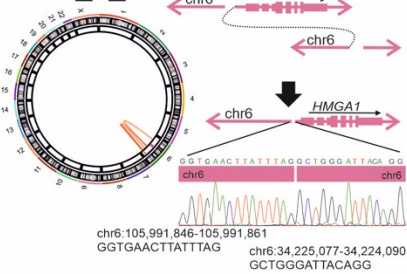


Supplementary Figure S5. Somatic copy number profiles of 23 endometrial polyps. Somatic copy number analysis revealed that endometrial polyps display very few large-scale chromosomal losses and gains. We identified no clear signs of high-level amplifications, homozygous deletions, or aneuploidy.

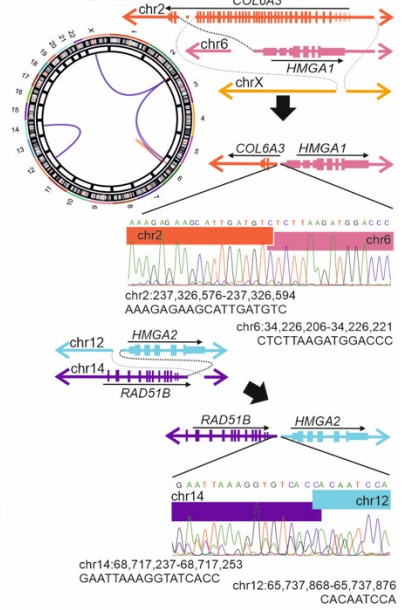
1085_1_S1



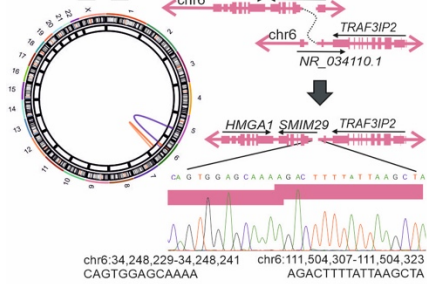
1087_1_S1



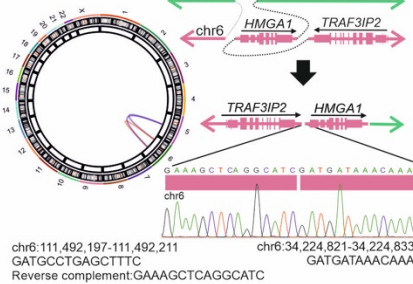
1092_1_S1



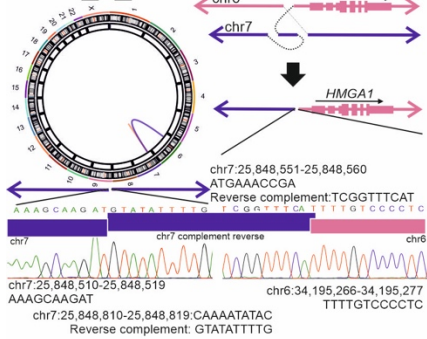
1093_1_S1



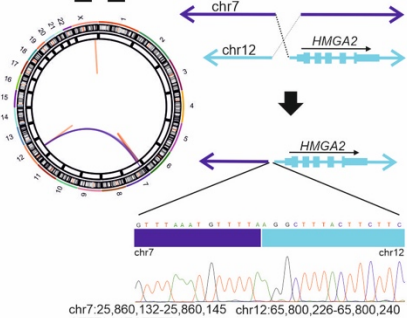
1094_1_S1



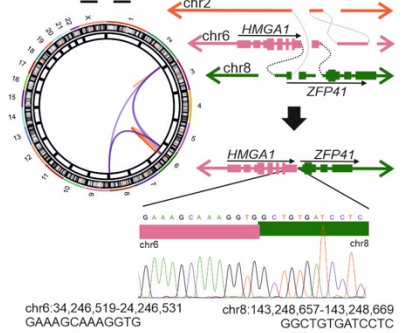
1097_1_S1



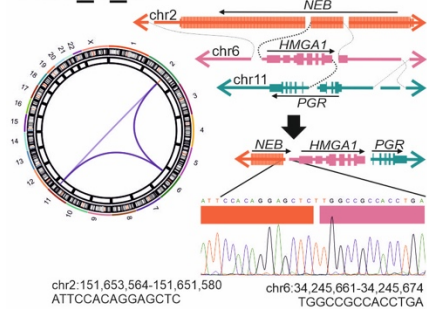
1098_1_S1



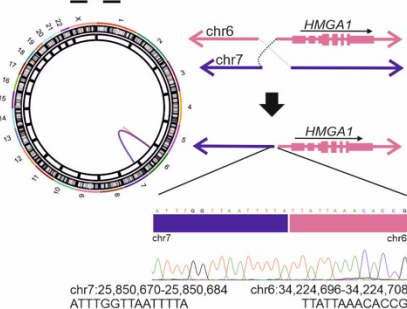
1121_1_S1



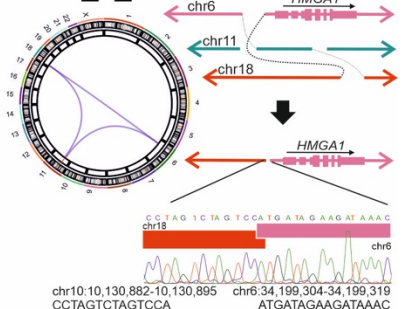
1123_1_S1



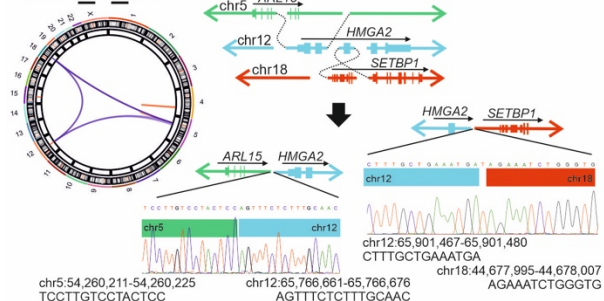
1124_1_S1

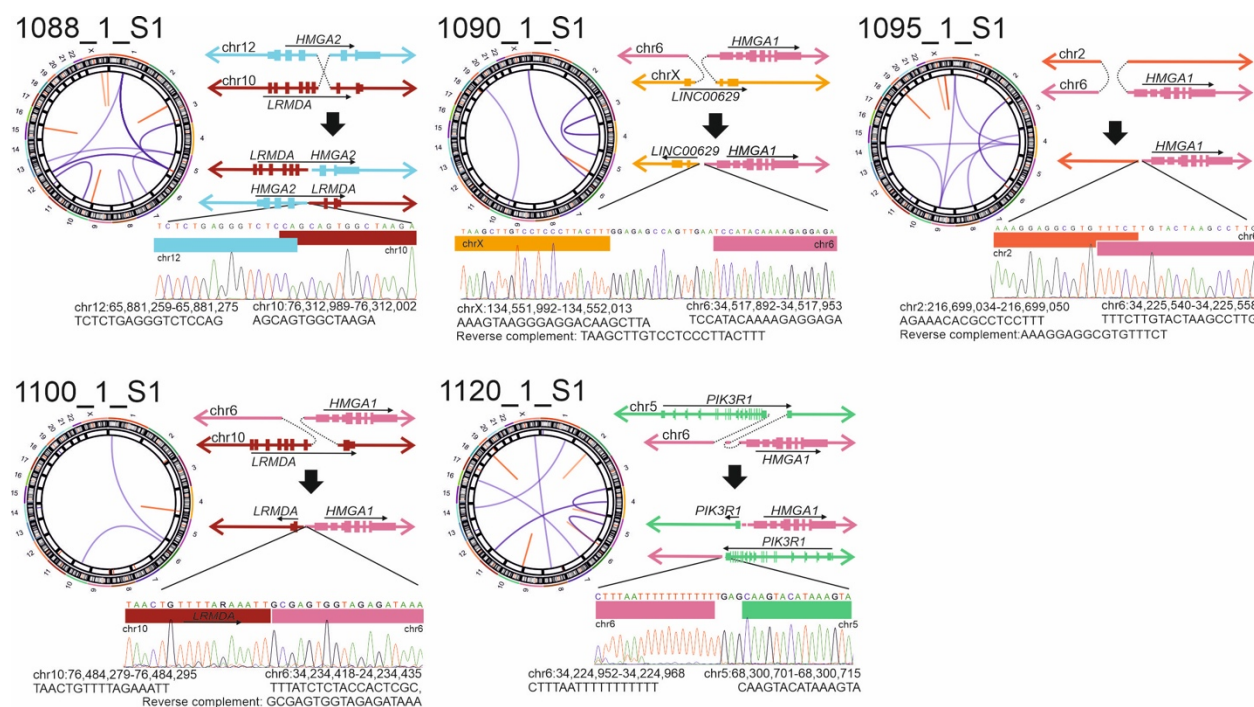


1126_1_S1



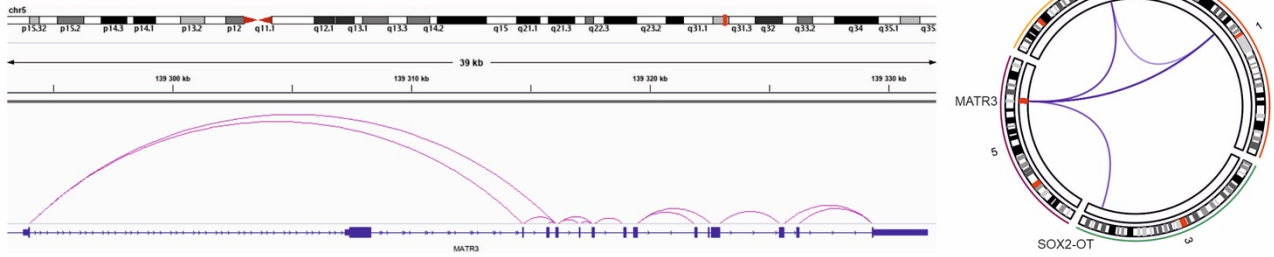
1516_1_S1



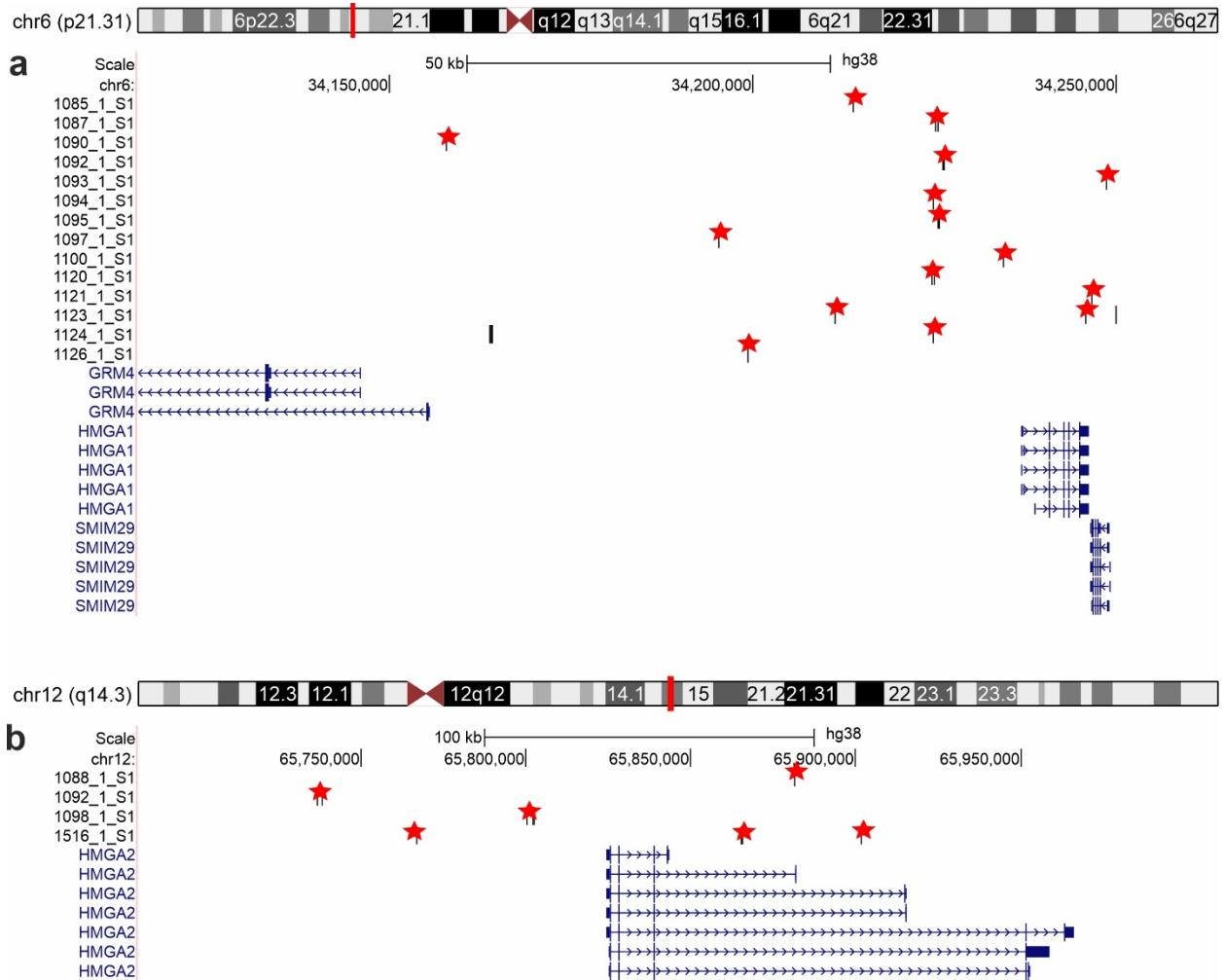


Supplementary Figure S6. Chromosomal rearrangements and candidate target genes for *HMGA1/2* rearrangements identified in 23 endometrial polyps. In total, 17/23 (74%) polyps harbored chromosomal rearrangements involving *HMGA1* or *HMGA2*. Occasionally, other independent chromosomal rearrangements were observed. Samples 1088_1_S1, 1090_1_S1, 1095_1_S1, 1100_1_S1, and 1120_1_S1 do not have a matched normal. The schematic figures are not in scale. In the circos plots, interchromosomal rearrangements are shown in blue and intrachromosomal inversions are in red. In the inner circle, deletions called as structural variants are marked with red and duplications called as structural variants in blue.

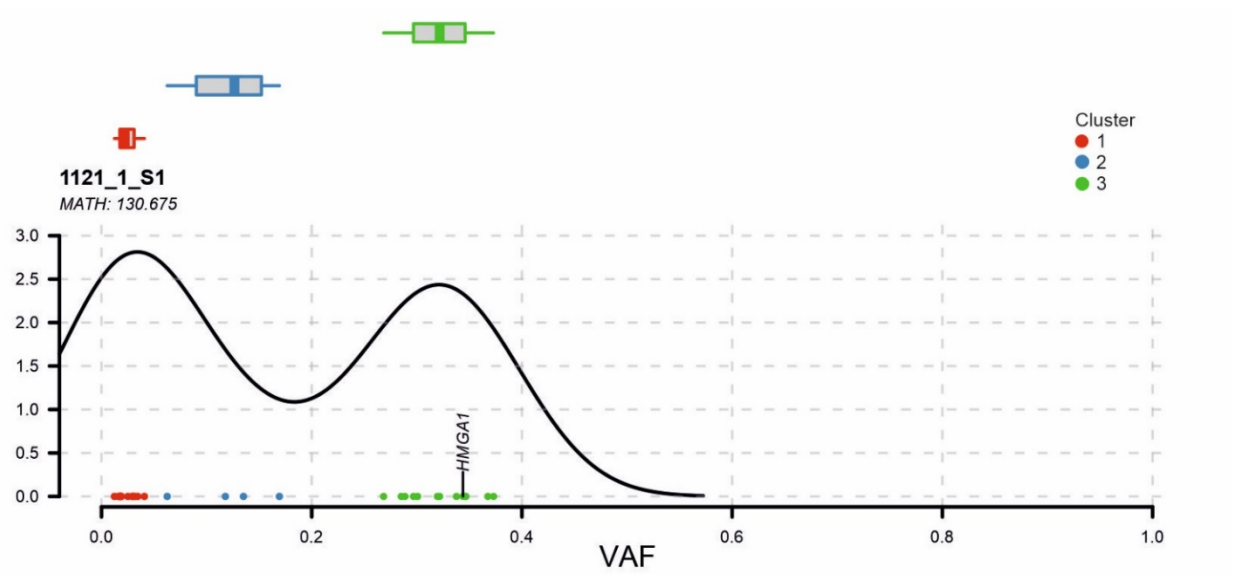
1088_1_S1



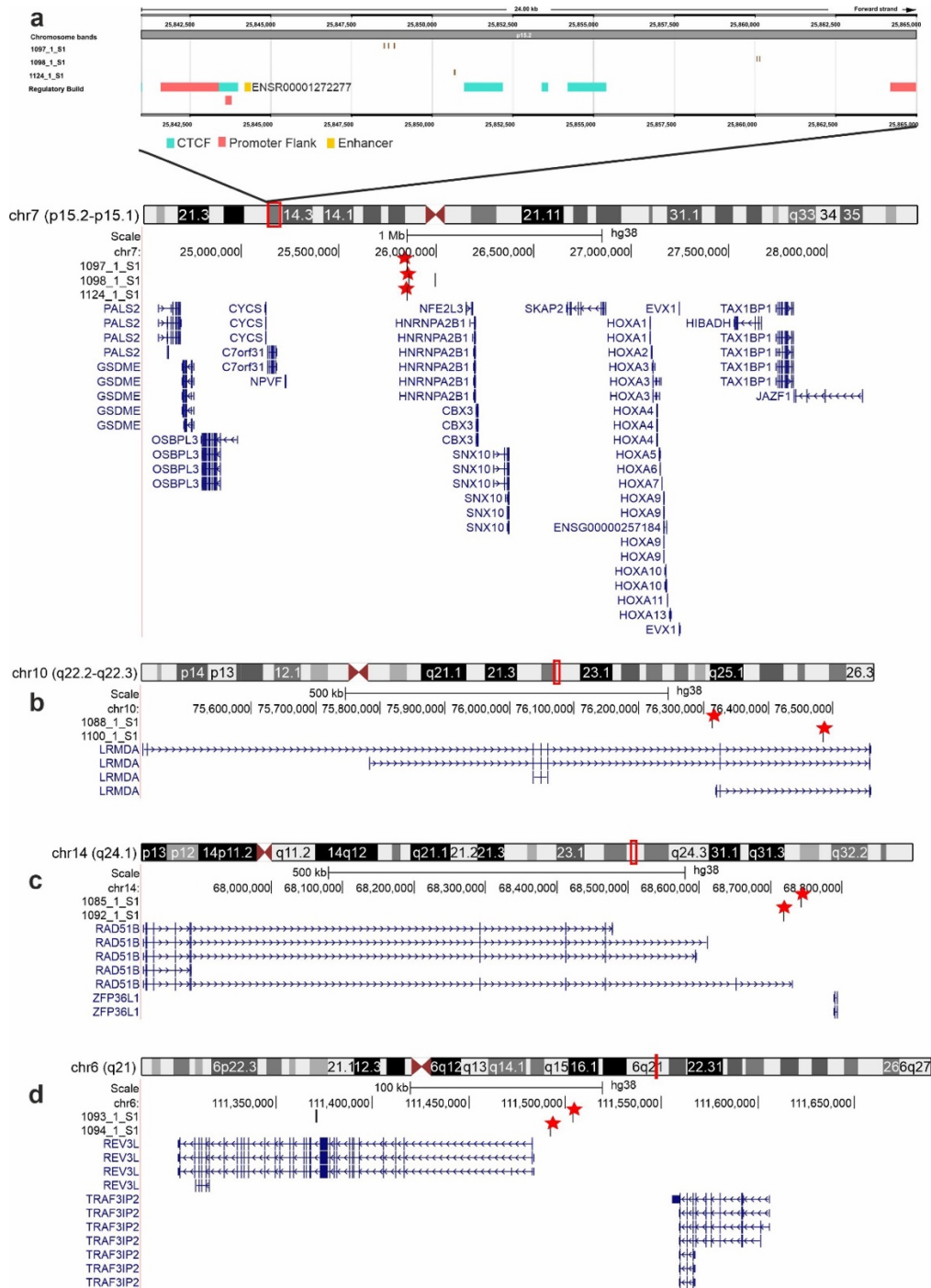
Supplementary Figure S7. Gene retrocopy insertions of *MATR3*. We identified one sample (1088_1_S1) with multiple gene retrocopy insertions. *MATR3* was inserted into chromosomes 1, 3, and 12. The circos plot displays only the calls related to the *MATR3* insertion. In the circos plot, insertion calls are shown in blue, and deletions are shown as red in the inner circle.



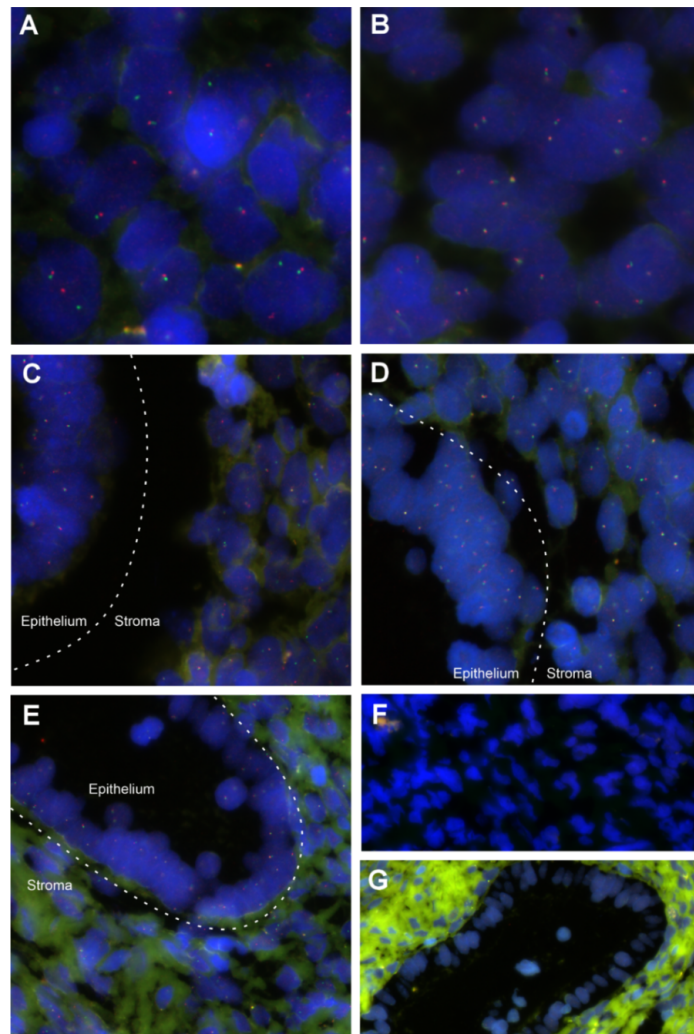
Supplementary Figure S8. Location of breakpoints upstream, within, and downstream of *HMGA1* and *HMGA2*. **A)** We identified 15 breakpoints leading to *HMGA1* rearrangements in 14 polyps (polyp 1123_1_S1 had two breakpoints). Twelve breakpoints were upstream of *HMGA1*, one downstream of *HMGA1*, and two in the 3'UTR of *HMGA1*. **B)** We identified four polyps with a total of six breakpoints causing *HMGA2* rearrangements. Three breakpoints were upstream of *HMGA2* and three were intragenic. One polyp (1516_1_S1) showed complex rearrangements including three breakpoints upstream and within *HMGA2*. Breakpoints connecting *HMGA1* or *HMGA2* to a candidate partner gene/region are marked with a star.



Supplementary Figure S9. High allelic fraction mutation in *HMGA1*. One polyp (1121_1_S1) with an *HMGA1* rearrangement also displayed a somatic missense mutation c.268C>G, p.(Leu90Val) in *HMGA1*. The variant had a relatively high allelic fraction, and it was present in a more clonal cell population.



Supplementary Figure S10. Breakpoints located upstream, within, and downstream of recurrent partner genes. We identified **A) 7p15.2**, **B) *LRMDA***, **C) *RAD51B***, and **D) *TRAF3IP2*** as recurrent candidate rearrangement partners for *HMGA1* and *HMGA2*. Breakpoints connecting the rearrangement to *HMGA1* or *HMGA2* are marked with a star.



Supplementary Figure S11. *HMGAI* break-apart FISH in formalin-fixed paraffin embedded endometrial polyp samples. **A)** Close-up of stromal cells showing break-apart signals in a sample that harbors an *HMGAI* rearrangement based on WGS. **B)** Close-up of stromal cells showing a normal signal pattern in a sample without *HMGAI* alterations based on WGS. **C)** Sample with an *HMGAI* rearrangement displays a normal signal pattern in epithelial cells, whereas a break-apart signal can be seen in stromal cells. Quality of the sample is not sufficient for formal analysis. **D)** Sample without an *HMGAI* rearrangement shows normal signals in both epithelial and stromal cells. Again, the quality of the sample is not sufficient for formal analysis. **E)** Throughout the experiment, epithelial cells performed better and exhibited a normal signal pattern with co-localized 3' and 5' probes, while stromal cells were of poor quality. **F-G)** Examples of samples that failed in the analysis due to **F)** badly damaged nuclei and **G)** excessive paraffin combined with weak signals. Overall, while some samples showed break-apart signals in some cells, the quality of the FFPE samples did not allow formal analysis with appropriate cell counting and scoring.