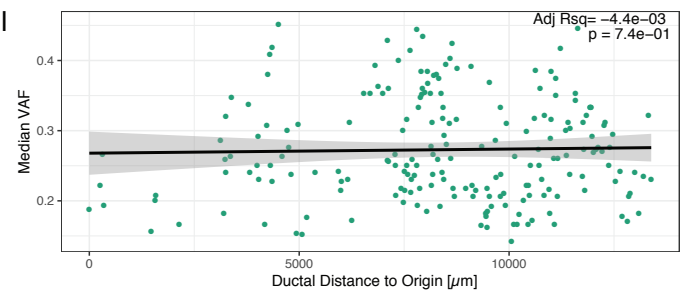
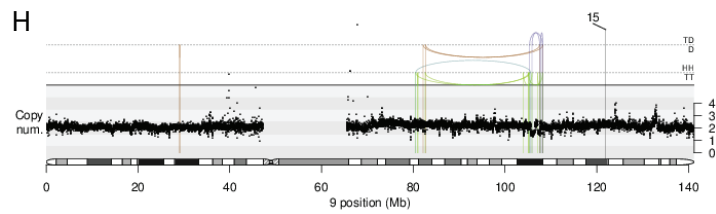
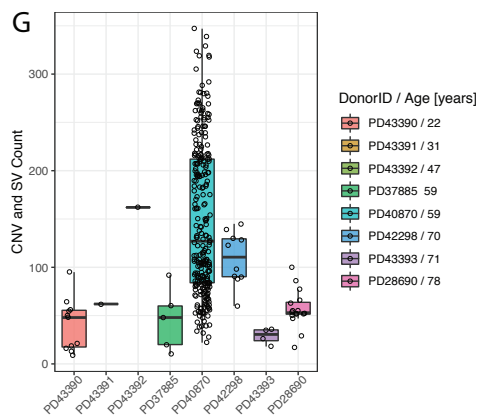
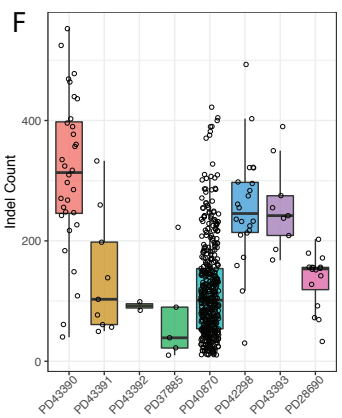
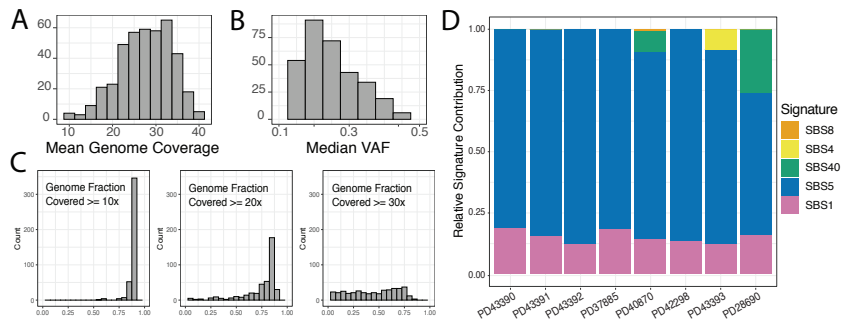


Cell Stem Cell, Volume 28

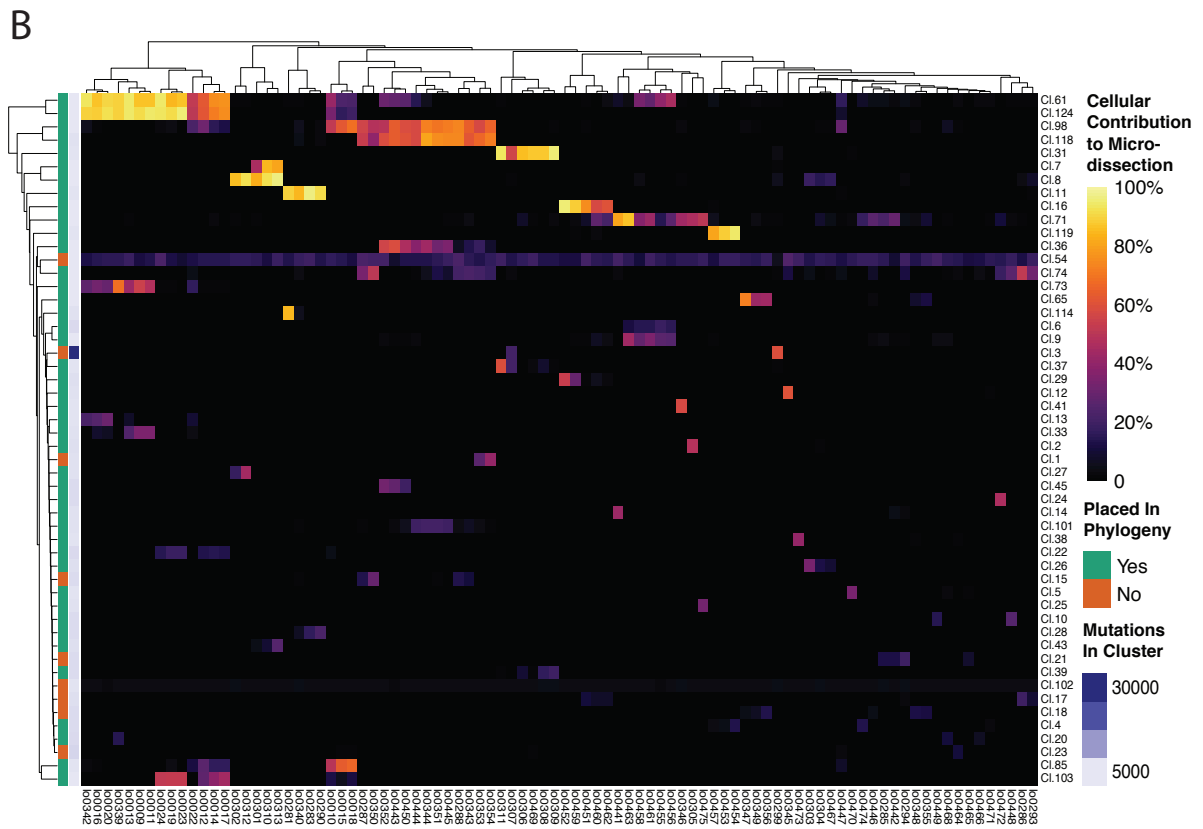
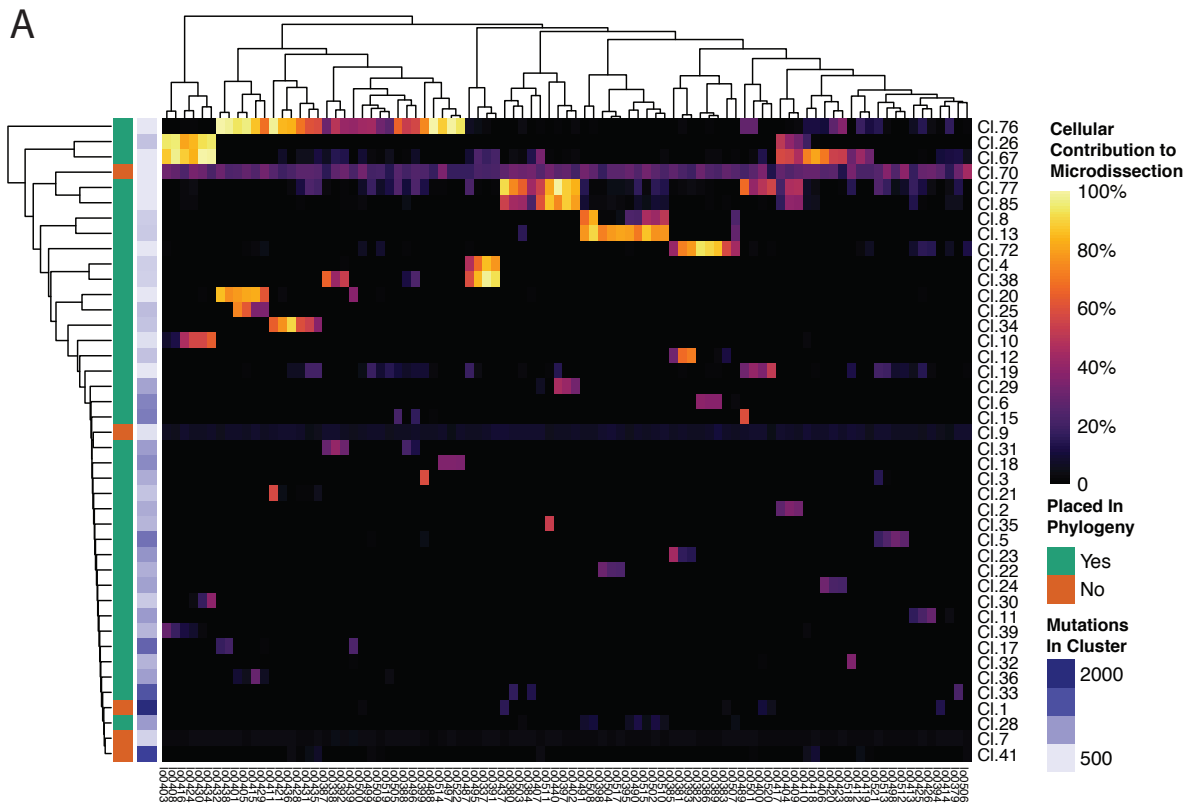
Supplemental Information

**Development, maturation, and maintenance
of human prostate inferred from somatic mutations**

Sebastian Grossmann, Yvette Hooks, Laura Wilson, Luiza Moore, Laura O'Neill, Iñigo Martincorena, Thierry Voet, Michael R. Stratton, Rakesh Heer, and Peter J. Campbell

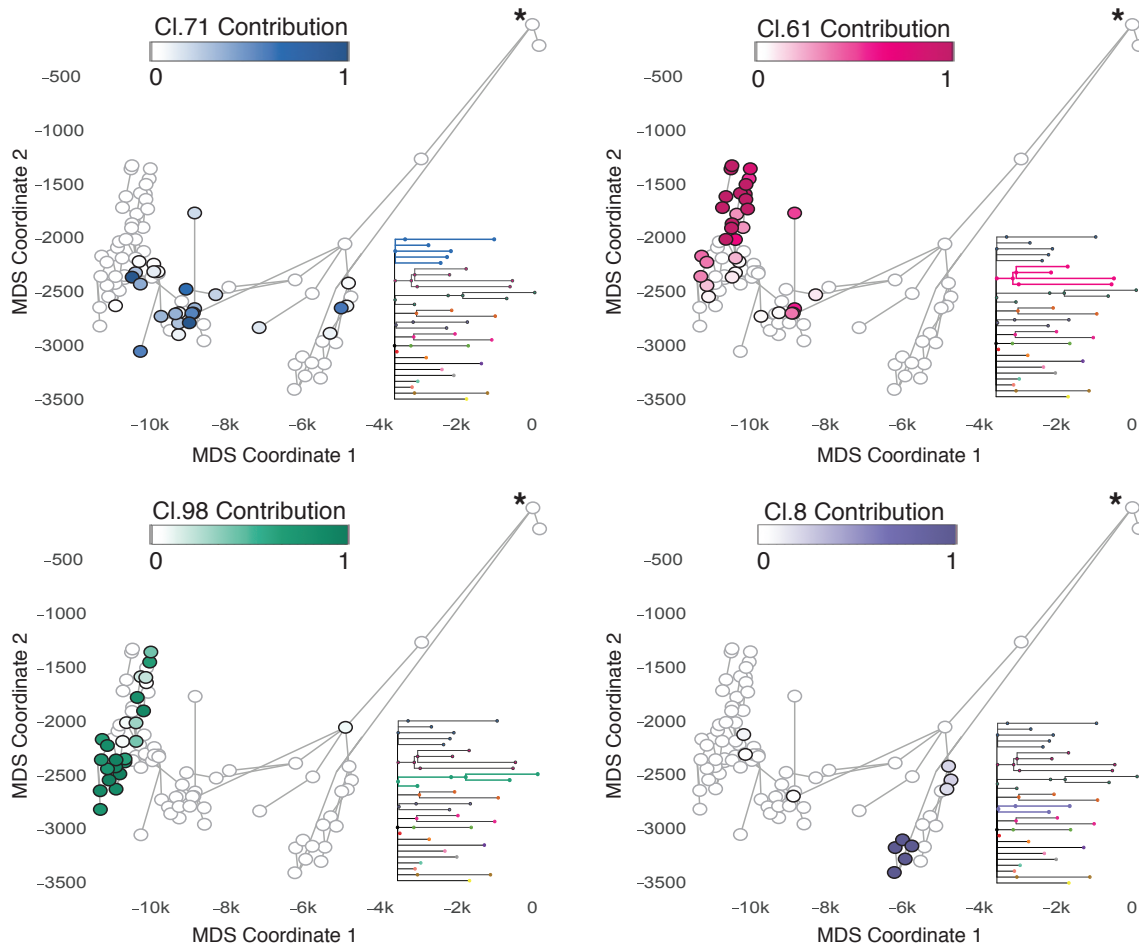


Supplementary Figure 1 (related to Figure 1): Genome Sequencing Quality Metrics and Mutational Landscape of Normal Prostate Epithelium. **(A)** Mean genome coverage per sample. **(B)** Median variant allele fraction (VAF) of single base substitutions as a measure of clonality per microdissection. **(C)** Genome Fraction per sample with coverage of at least 10x/20x/30x. **(D)** Relative contribution of deconvoluted mutational signatures per donor. The clock-like mutational signatures SBS1 and SBS5 dominate the signature landscape across all donors. **(E)** Exemplary spectral signatures of base substitutions per donor. **(F)** Count of small insertions and deletions per sample. **(G)** Only few, short copy number variants and very rare structural variants were detected in normal prostatic epithelium. **(H)** One histologically normal microdissection of prostatic epithelium contained a case of chromothripsis on chromosome 9-q. **(I)** Median VAF of base substitutions per microdissection is not correlated to the position within the ductal network.

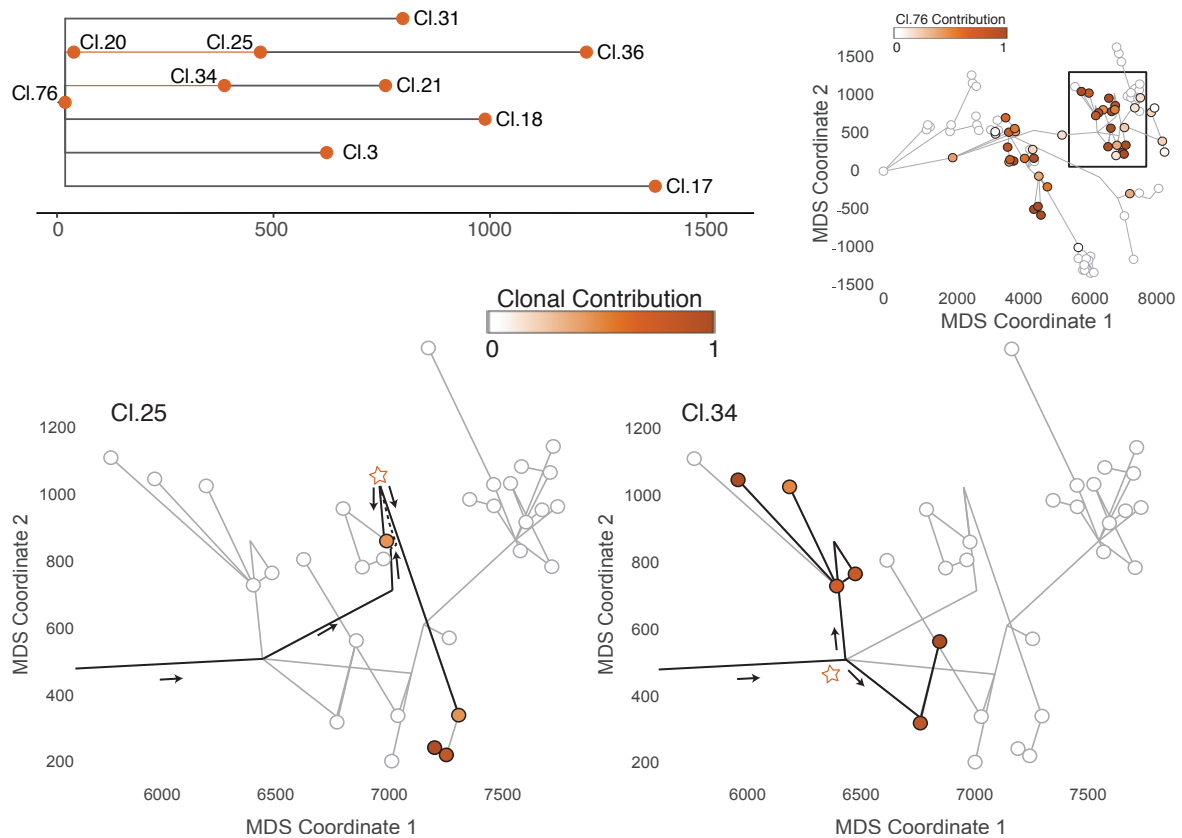


Supplementary Figure 2 (related to Figure 2): Discrete Clusters of Mutations Identify Ancestral Clones that Contribute to Current Prostatic Epithelium.

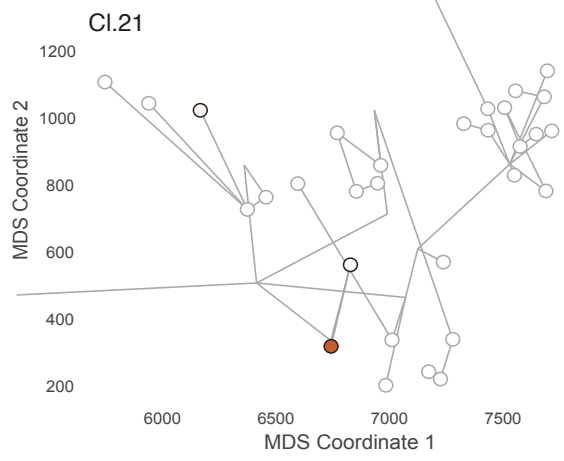
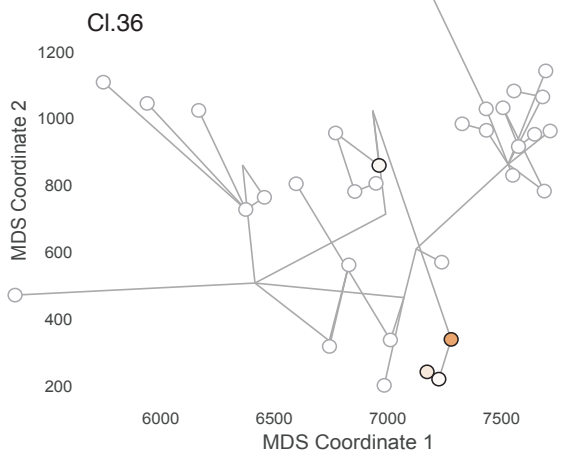
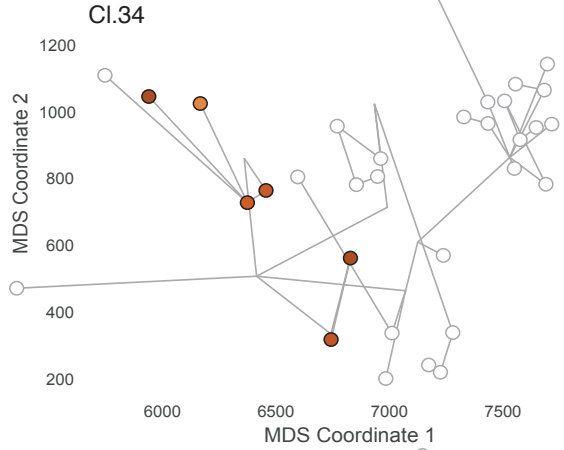
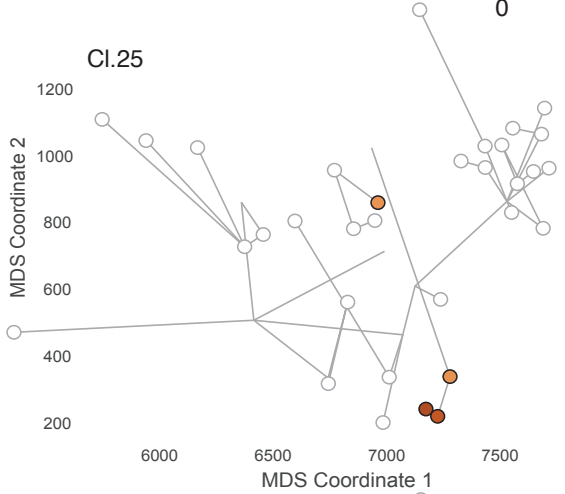
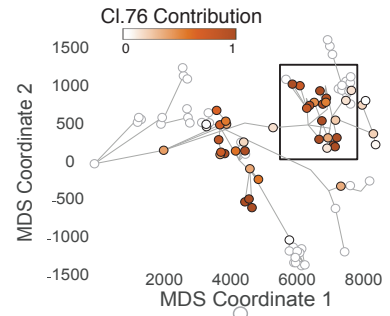
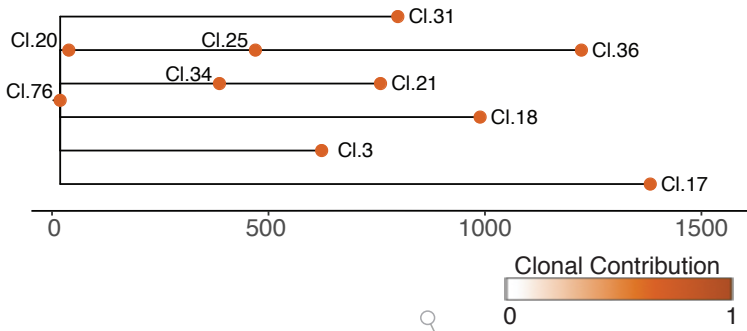
Clusters of mutations are identified based on VAF using an n-dimensional Hierarchical Dirichlet Process (n-HDP) clustering algorithm and constitute the rows of the heatmap. These clusters of mutations represent ancestral clones which contribute to current cells sampled in individual microdissections, which are shown in the columns of the heatmap. The median VAF of cells associated with an ancestral cluster is used to estimate the cellular contribution to each microdissection and is visualized using color intensities. The co-occurrence of multiple ancestral clones within the same microdissection sample and their corresponding cellular contribution was used to construct the lineage tree for individual glandular subunits. **(A) Heatmap for the left-hand side structure.** Cluster 70, 9 and 7 include nearly exclusively artefactual mutations and were excluded from the lineage tree. Furthermore, cluster 1 and 41 are insufficiently split by the clustering algorithm used and could not be placed unambiguously. **(B) Heatmap for the right-hand side structure.** Cluster 54 and 102 contain artefactual mutations. Seven additional clusters could not be placed onto the lineage tree unambiguously.



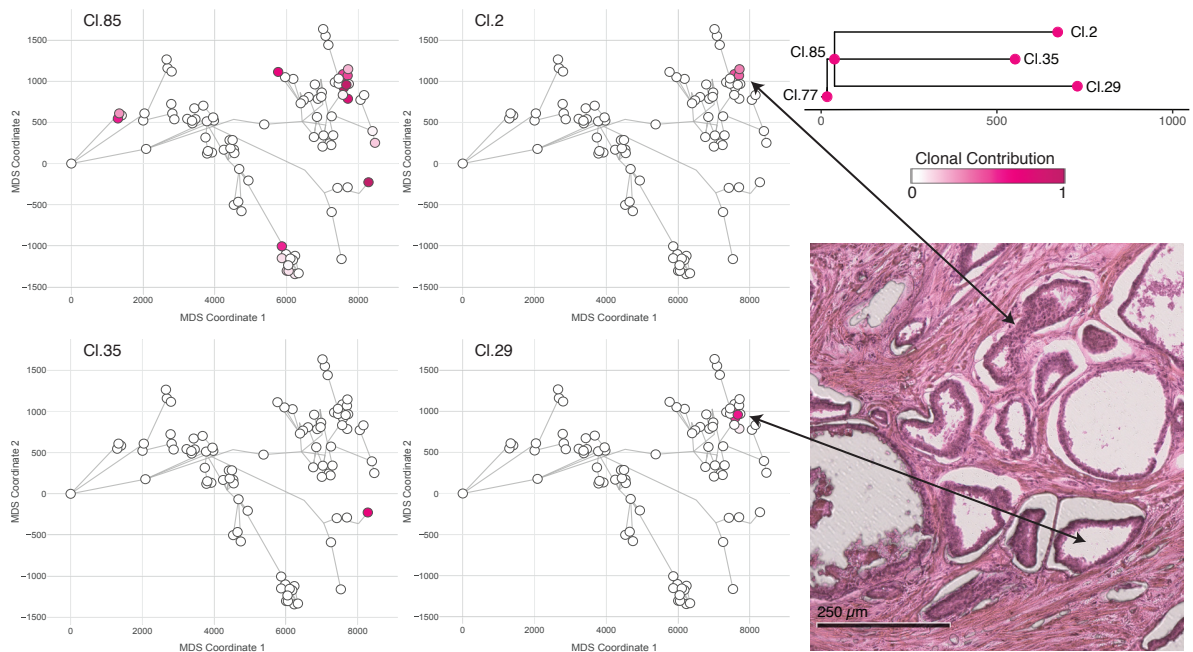
Supplementary Figure 3 (related to Figure 3): Cellular Contribution of Embryonic Clones for Glandular Subunit on the Right-Hand Side. The cellular contribution of four ancestral clones from embryonic development is displayed for the glandular subunit from the right-hand side. Each circle marks a microdissection and the ductal connection is indicated by grey lines. Microdissections with positive contribution from the clone are circled in black, while those with zero contribution are circled in grey. The most urethra-proximal microdissection is marked with an asterisk. Three embryonic clones display a wide and overlapping spatial distribution. The clone represented by cluster 8 is confined to a single sampled branch of the glandular network.



Supplementary Figure 4 (related to Figure 4): Cellular Contribution of Pubertal Clones Within the Ductal Subunit on the Left-Hand Side. The cellular contribution of pubertal clones represented by cluster 25 and 34 within the clonal territory of the embryonic clone represented by cluster 76 is highlighted. The proximal-to-peripheral direction of the ductal network is indicated by arrows. Both pubertal clones contribute to two distinct side-branches and their corresponding point from the main duct is highlighted with a star. Notably, in both cases the two branches seeded by the corresponding pubertal clones extend from the main duct directly on opposing sides and cannot be detected anywhere else. This strong spatial confinement suggests a localized progenitor cell before their respective pubertal expansions.



Supplementary Figure 5 (related to Figure 4): Embryonic, Pubertal and Adult Clones Occupy Increasingly Confined Territories. An example for embryonic, pubertal and adult clones within one clade of the left-hand side glandular subunit. The embryonic clone represented by cluster 76 occupies a wide spatial territory throughout most of the sampled branches. Pubertal clones represented by cluster 25 and 34 still display a confined spatial expansion with the embryonic territory. The adult clones represented by cluster 36 and 21 are only found within one or few microdissections overlapping their corresponding embryonic and pubertal clone. Notably, while most adult clones are found in peripheral microdissections, the highest cellular contribution of an adult clone is not necessarily associated with the most peripherally sampled microdissection as seen for both cluster 36 and cluster 21.



Supplementary Figure 6 (related to Figure 5): Phylogenetic Relatedness can be Independent of Spatial Proximity. An example for the glandular subunit of the left-hand side is shown. The three clones represented by cluster 2, 35 and 29 share a most recent common ancestor in the embryonic clone represented by cluster 85. Cluster 85 shows great spatial distribution across most sampled branches of the glandular subunit. While the clonal territories occupied by cluster 2 and 29 are spatially close as indicated by the structural overview as well as the histology image, the clonal territory occupied by cluster 35 is completely distinct. This example highlights the strongly localized nature of adult tissue maintenance that causes phylogenetic relationship to be largely independent of spatial proximity after embryonic and pubertal morphogenesis.

Supplementary tables

Table S1 (related to Figure 1): Overview of Available Prostate and WGS Samples. WGS from microdissections of histologically normal prostatic epithelium was obtained from eight different donors.

Donor ID	Sex	Donor Age [Years]	Sample Type	WGS Samples
PD43390	M	22	Post-mortem biopsy	32
PD43391	M	31	Post-mortem biopsy	9
PD43392	M	47	Post-mortem biopsy	2
PD37885	M	59	Post-mortem biopsy	5
PD40870	M	59	Whole prostate	319
PD42298	M	70	Biopsy	21
PD43393	M	71	Post-mortem biopsy	9
PD28690	M	78	Post-mortem biopsy	12

Table S2 (related to Figure 6): Overview of Non-synonymous Mutations in Significantly Mutated Genes in Prostate Cancer. Sixteen mutations called in microdissections from normal prostatic epithelium were found in previously published SMGs in prostate cancer. Most mutations were only called in a single microdissection.

Donor ID	Gene	Change	Impact	# of calls
PD42298	<i>KMT2A</i>	C1479W	Missense	1
PD42298	<i>ZNF292</i>	Y2193C	Missense	1
PD40870	<i>RPRD2</i>	D327G	Missense	1
PD40870	<i>RAG1</i>	R34W	Missense	1
PD40870	<i>RAG1</i>	V529M	Missense	1
PD40870	<i>MRE11A</i>	I647L	Missense	1
PD40870	<i>BRCA2</i>	F1336L	Missense	1
PD40870	<i>FOXA1</i>	R219S	Missense	5
PD40870	<i>NCOR1</i>	P1665T	Missense	1
PD40870	<i>NOX3</i>	V72I	Missense	1
PD40870	<i>CHD7</i>	R1155C	Missense	1
PD40870	<i>COL15A1</i>	A542D	Missense	2
PD40870	<i>AR</i>	.	Essential Splice	1
PD40870	<i>AR</i>	R832Q	Missense	1
PD40870	<i>SMARCA1</i>	W426*	Nonsense	1
PD28690	<i>PTPRC</i>	E362K	Missense	2