## Supplementary Figures:

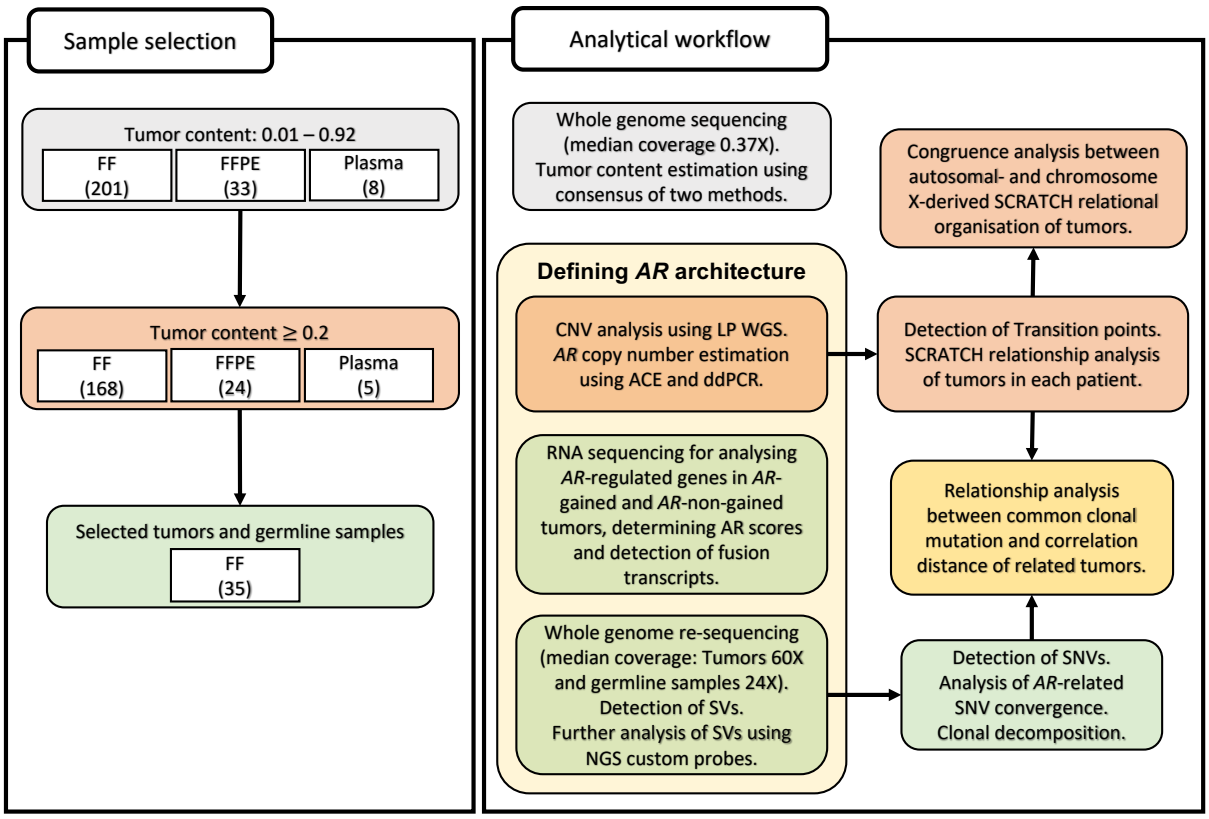



**Figure S1: Experimental Study Design**

**a) Sample summary**: Samples (FF: fresh frozen, FFPE: formalin-fixed paraffin embedded or plasma; sample numbers as shown in parentheses) were subjected to shallow whole genome sequencing (median coverage: 0.37X, range: 0.07X-5.8X). Samples with estimated tumor content ≥0.2 were selected for further analysis, including *AR* copy number determination and genome-wide copy number assessment. Two to six metastases from each patient were chosen as described in the text for re-sequencing at a higher depth of coverage (median coverage: 60X, range 27X-83X). RNA-Seq was performed on 20 samples selected from three patients (CA63, CA76 and CA83).

**b) Analysis summary**: Copy number analysis on low coverage WGS was performed using ACE package, while somatic structural variants surrounding *AR* gene (chrX:50M-80M) were detected from high coverage WGS data using Delly v-0.7.8. *AR* downstream transcriptional activity was determined as AR score. Furthermore, clonal somatic mutations were detected from 25 tumor samples using Mutect2 and Sclust. On the other hand, relationship among metastases in a patient was determined

applying SCRATCH algorithm on copy number transition points. Congruence analysis between autosomal- and chromosome X-based SCRATCH relationship of metastases from patient CA34 and CA63 showed that those relationships are not similar by chance.

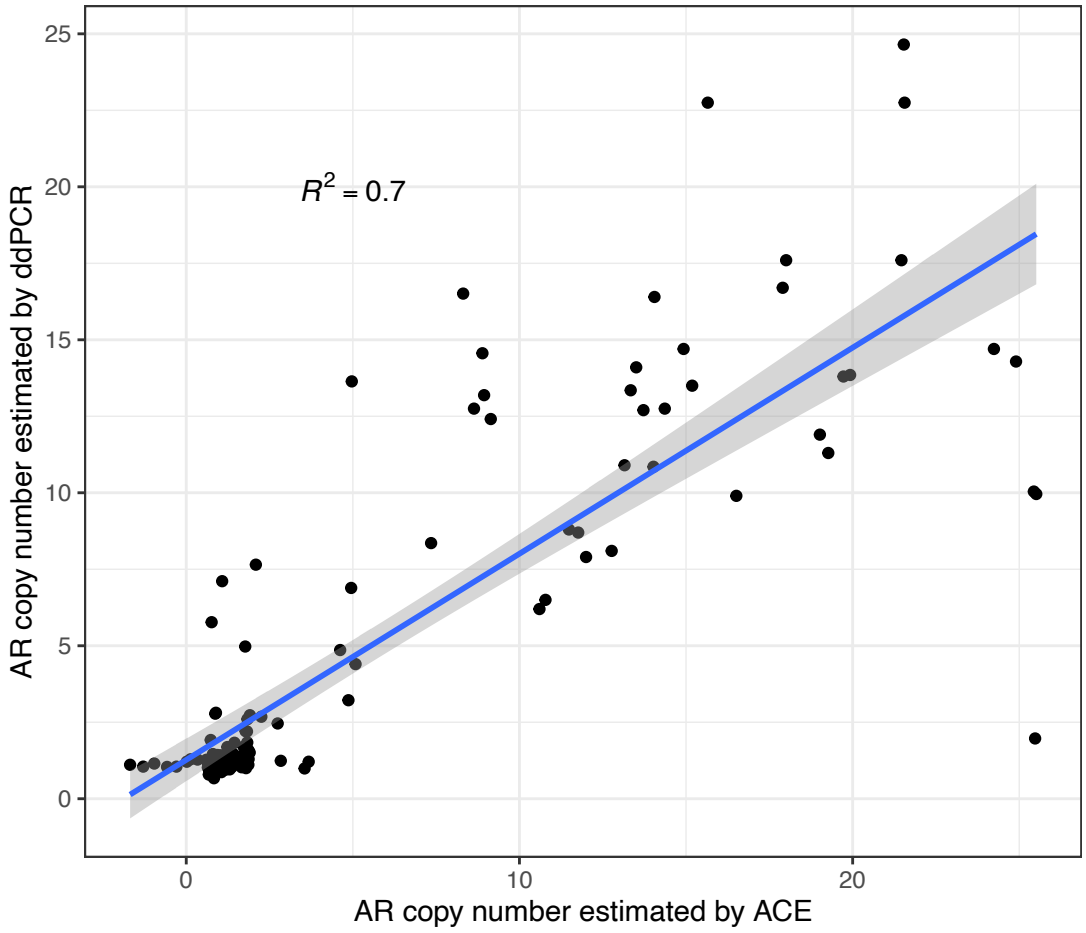



**Figure S2: *AR* copy number estimation**

*AR* copy numbers for 127 samples estimated by droplet digital PCR (ddPCR, along y-axis) correlated highly with the *AR* copy number estimates from shallow whole genome sequencing (using ACE algorithm, along x-axis). Linear regression line is in blue and confidence interval is shown in grey.

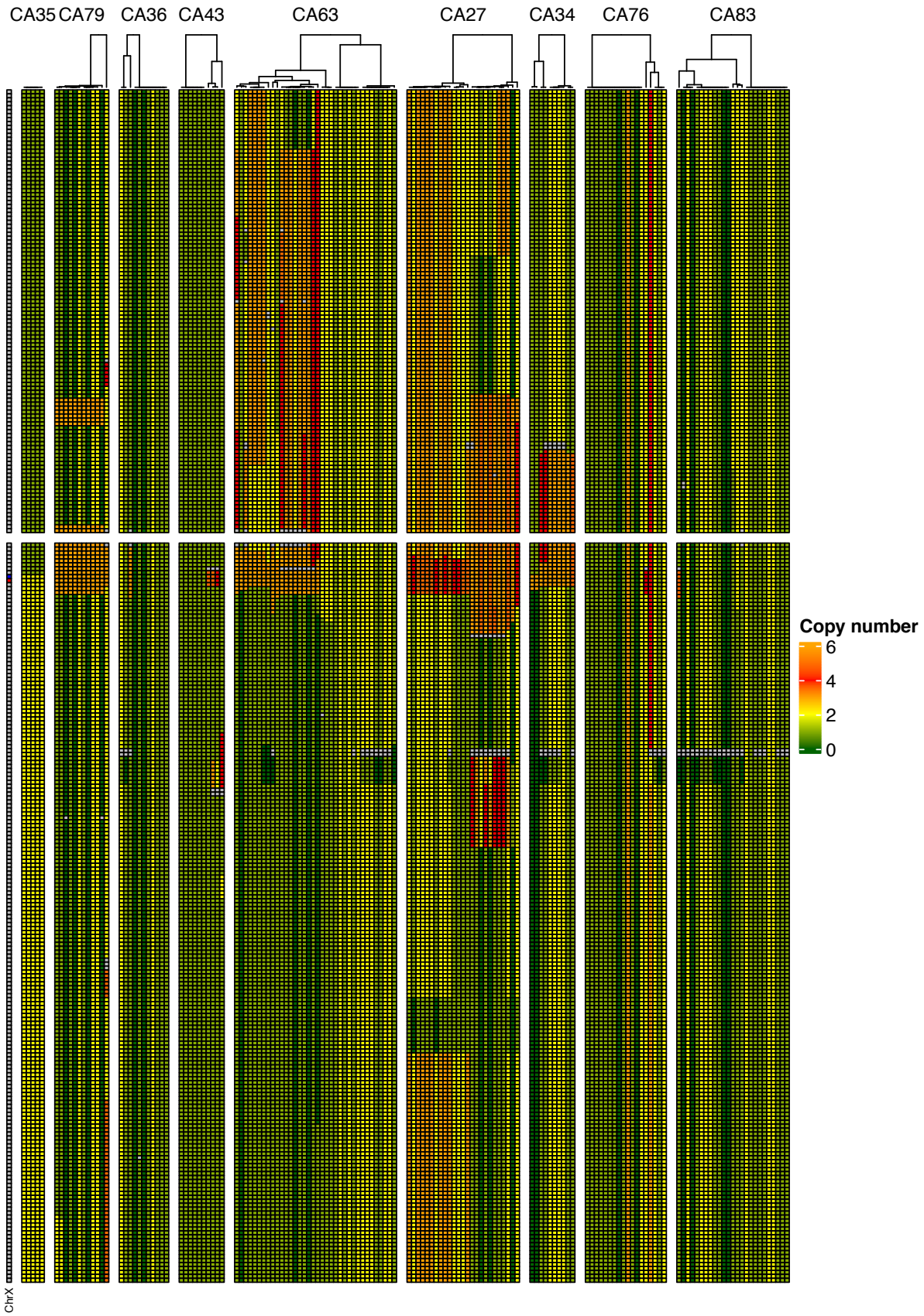



**Figure S3. Copy number gain extends over a large area on chromosome X.**

Heatmap showing the copy number of individual chromosome X bins (500 kb wide) for all 142 samples harvested post-mortem. Color scheme for the copy number shown in the legend. Chromosome X is shown on the left of the heatmap along with its bins covering *AR* gene (in red) and *AR*-associated centromeric enhancer (in blue). Metastases are ordered by patient identification and within patients, based on hierarchical clustering (shown as dendrogram) of correlation distances calculated among copy number transitions on chromosome X.

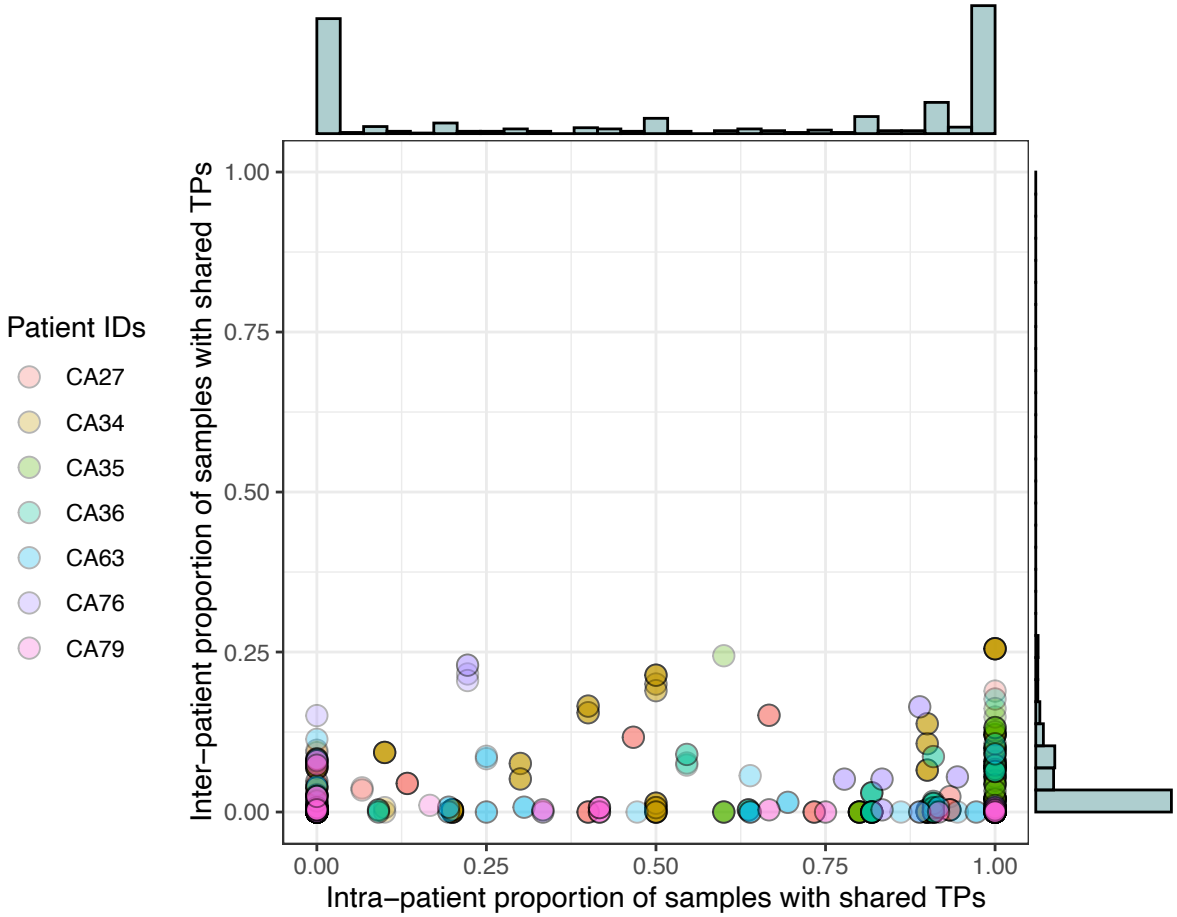
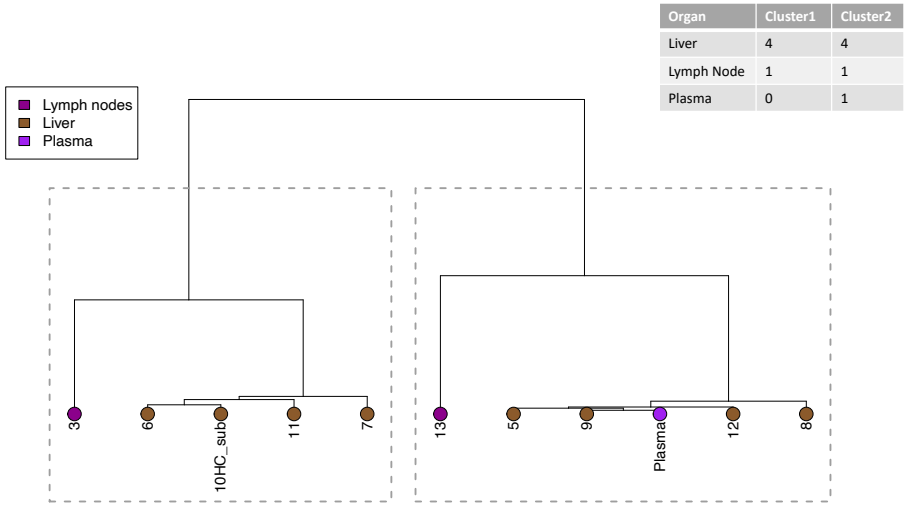



**Figure S4: Overlap of copy number transition points between post-mortem tumor samples and diagnostic biopsies within and across patients**

Circles representing proportion of transition points (TPs) detected in a diagnostic biopsy sample (n=7 patients) that are shared with samples harvested post-mortem from the same patient (x-axis) or any metastases harvested from other patients (y-axis). Marginal histograms confirm that whilst occasional TPs are shared with

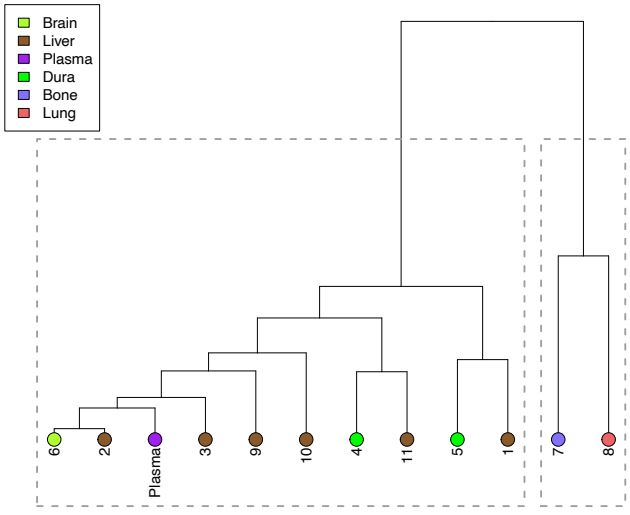
metastases from other patients, all patients' diagnostic samples share a high proportion of TPs (common or truncal TPs) with metastases from the same patient.

CA34



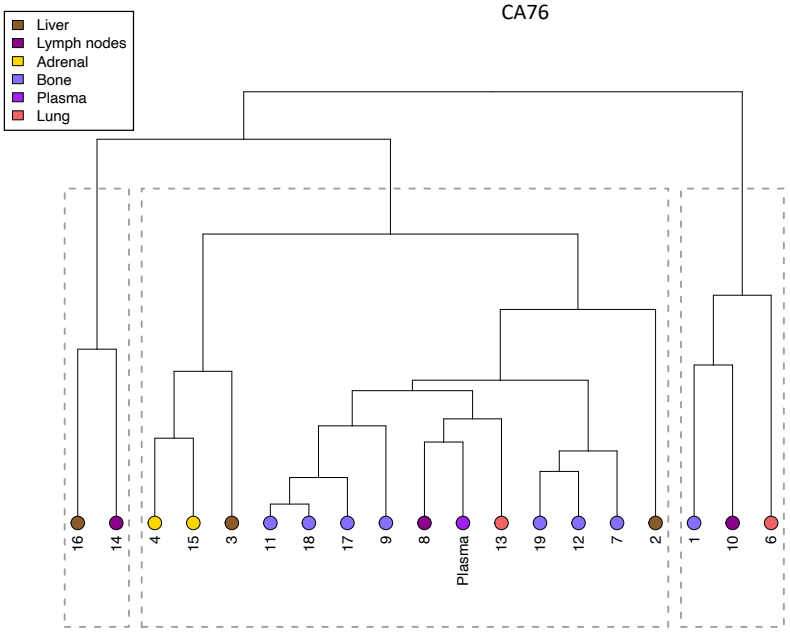

| Organ | Cluster1 | Cluster2 |
|---|---|---|
| Liver | 4 | 4 |
| Lymph Node | 1 | 1 |
| Plasma | 0 | 1 |

CA36




| Organ | Cluster1 | Cluster2 |
|---|---|---|
| Bone | 0 | 1 |
| Brain | 1 | 0 |
| Dura | 2 | 0 |
| Liver | 6 | 0 |
| Lung | 0 | 1 |
| Plasma | 1 | 0 |

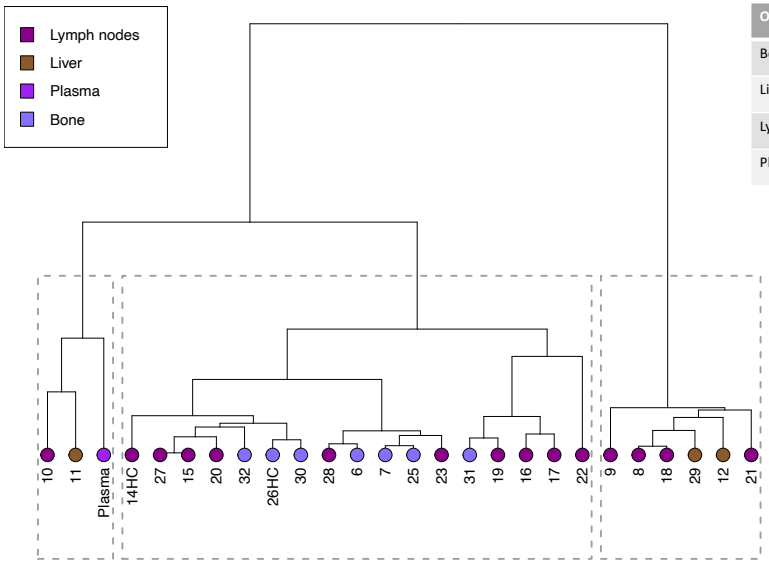
CA76




| Organ | Cluster1 | Cluster2 | Cluster3 |
|---|---|---|---|
| Adrenal | 0 | 2 | 0 |
| Bone | 0 | 7 | 1 |
| Liver | 1 | 2 | 0 |
| Lung | 0 | 1 | 1 |
| Lymph Node | 1 | 1 | 1 |
| Plasma | 0 | 1 | 0 |

54

CA83



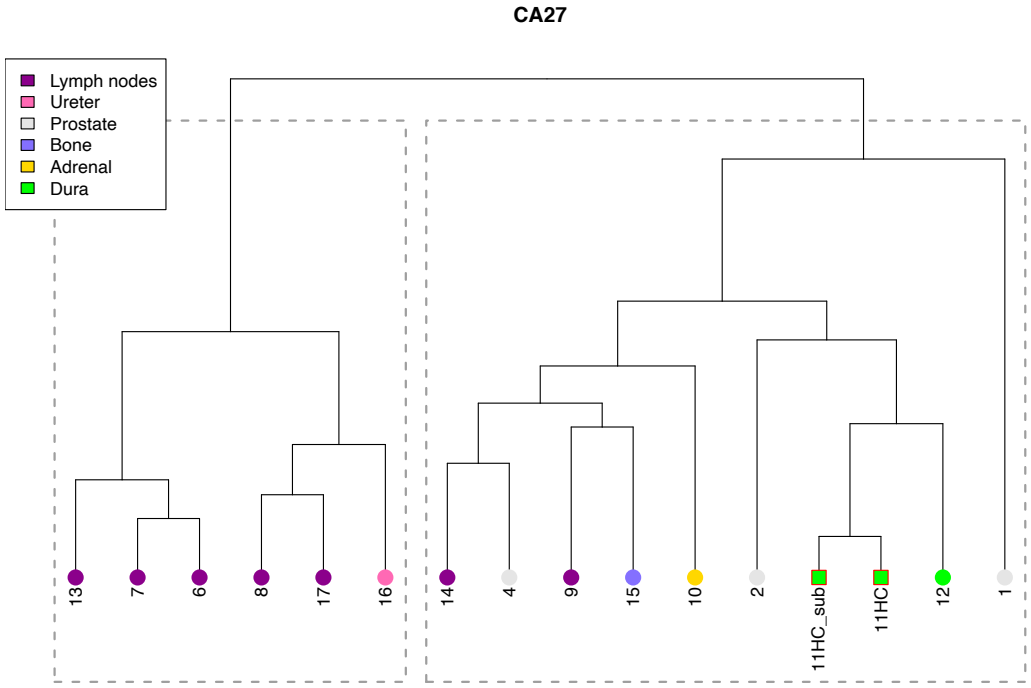

| Organ | Cluster1 | Cluster2 |
|---|---|---|
| Bone | 7 | 0 |
| Liver | 1 | 2 |
| Lymph Node | 11 | 4 |
| Plasma | 1 | 0 |

55

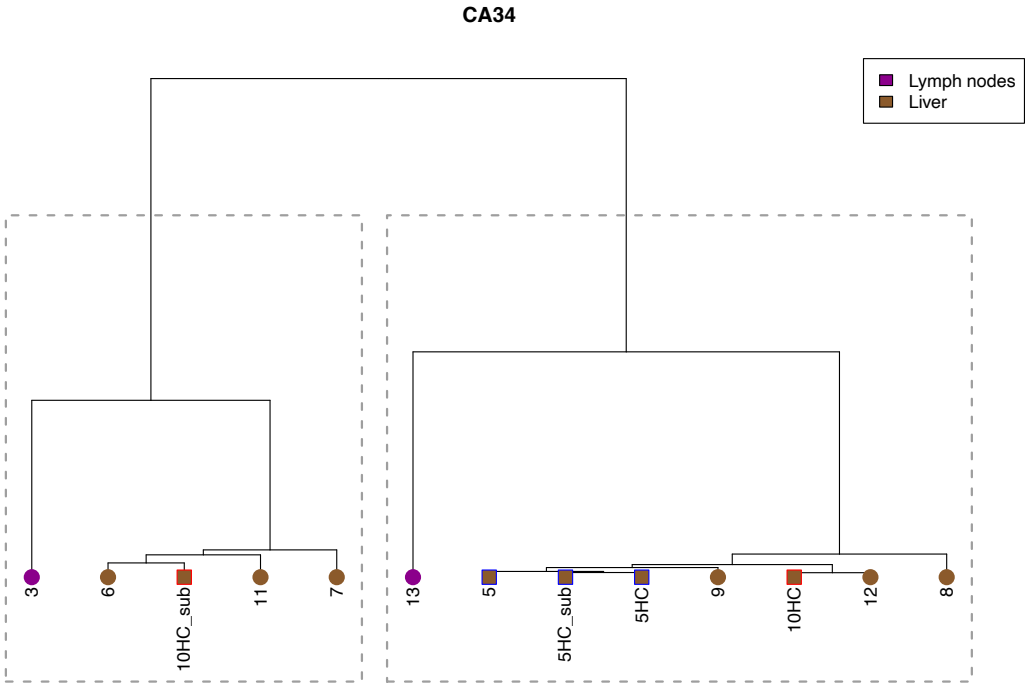
56

**Figure S5: SCRATCH relational networks of autopsy tumors for four patients with their plasma collected post-mortem**

SCRATCH relational network (described in methods), built on copy numbers at autosomal transition points, derived from the post-mortem metastatic samples and plasma (tumor content >= 0.2) collected at death from four patients (CA34, CA36, CA76 and CA83), are shown. The circles at the terminal nodes depict organ site for the metastatic samples or plasma in a patient. Metastatic samples and plasma are grouped in two to three clusters (details in methods) and separated by dashed grey boxes. Inset tables show the distribution of organ sites of metastases and plasma in individual clusters.

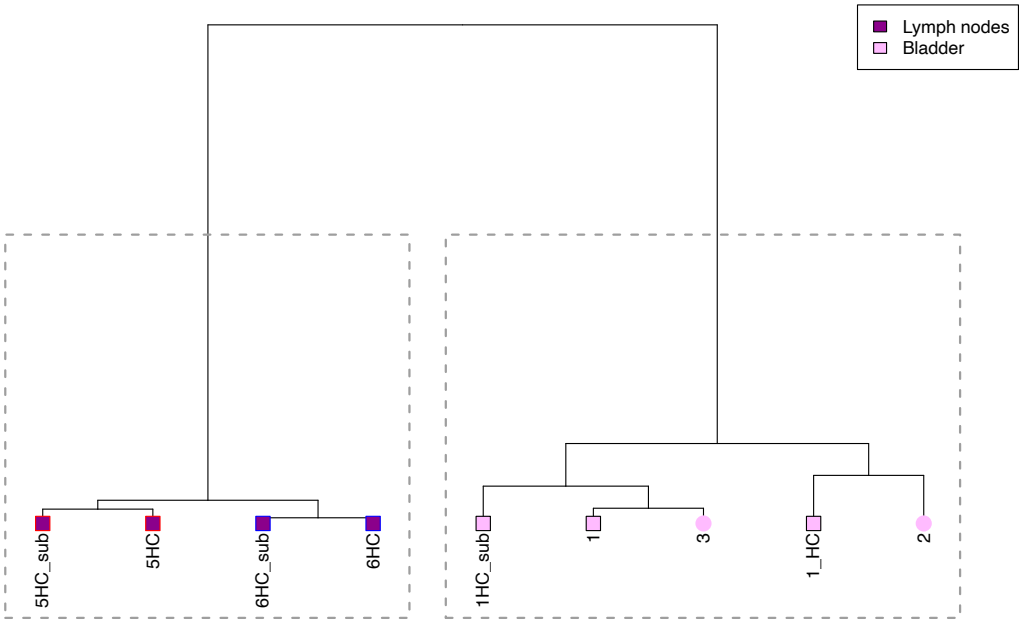

CA27
Lymph nodes
Ureter
Prostate
Bone
Adrenal
Dura
13
7
6
8
17
16
14
4
9
15
10
2
11HC_sub
11HC
12
1

68


CA34
Lymph nodes
Liver
3
6
10HC_sub
11
7
13
5
5HC_sub
5HC
9
10HC
12
8

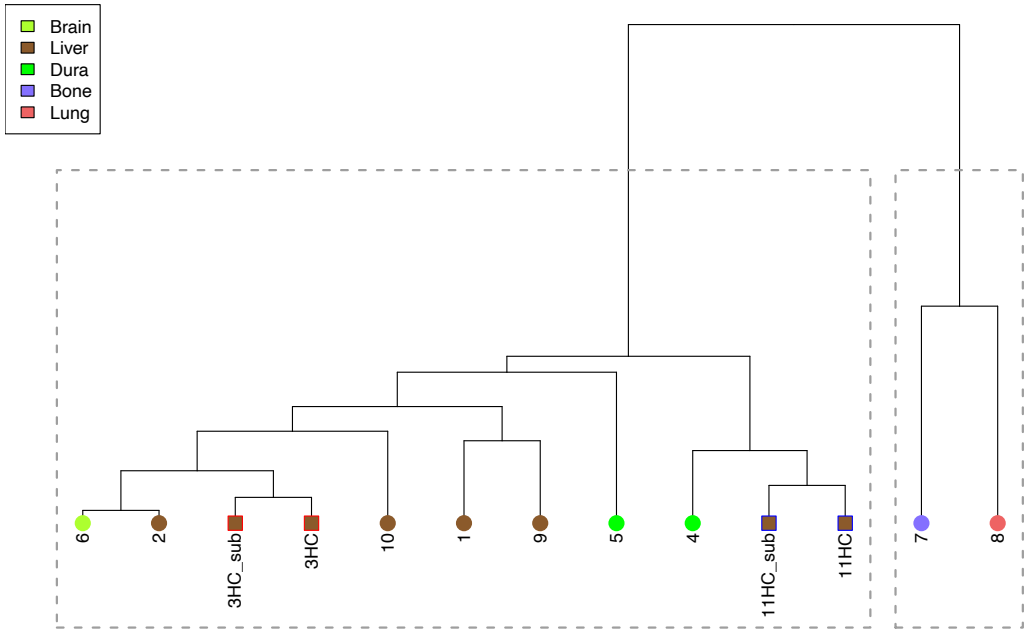
69

**CA35**



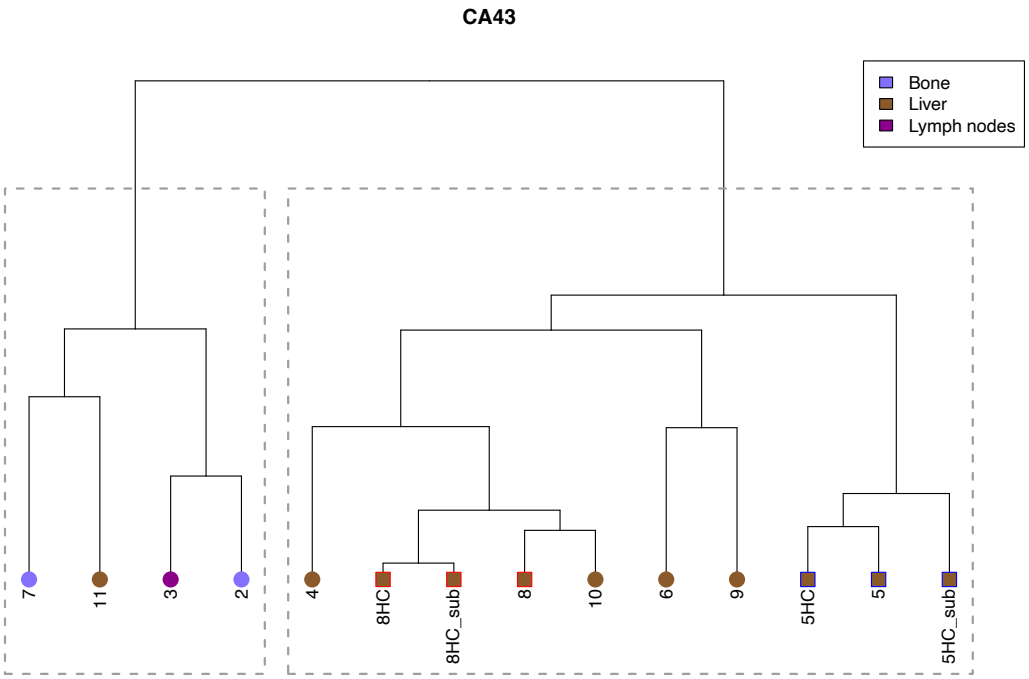

70

**CA36**

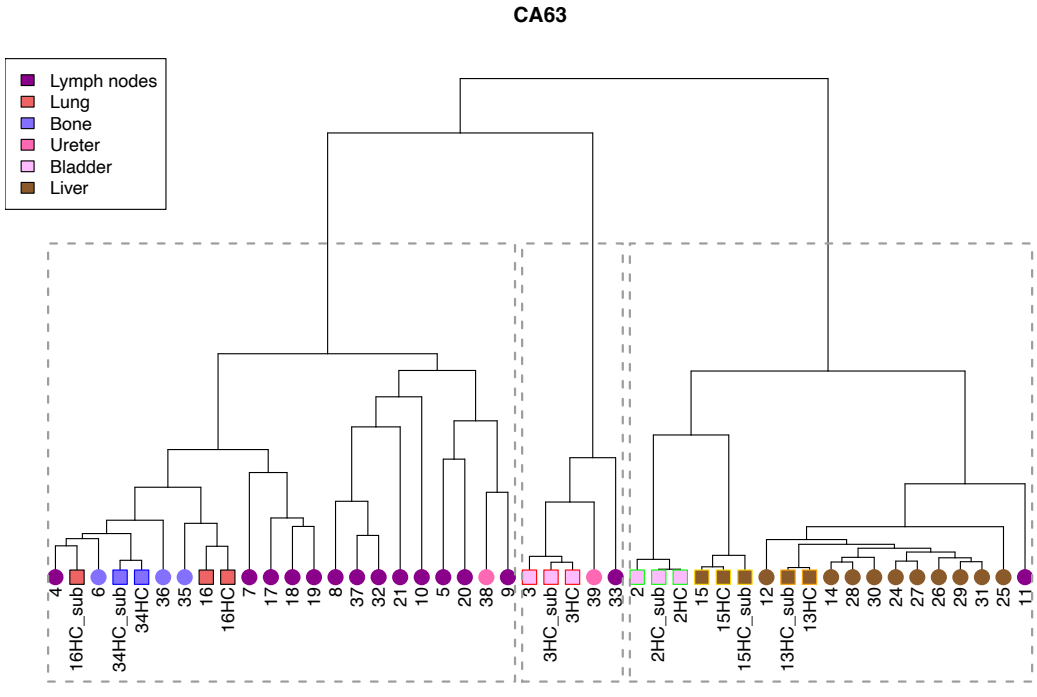



71

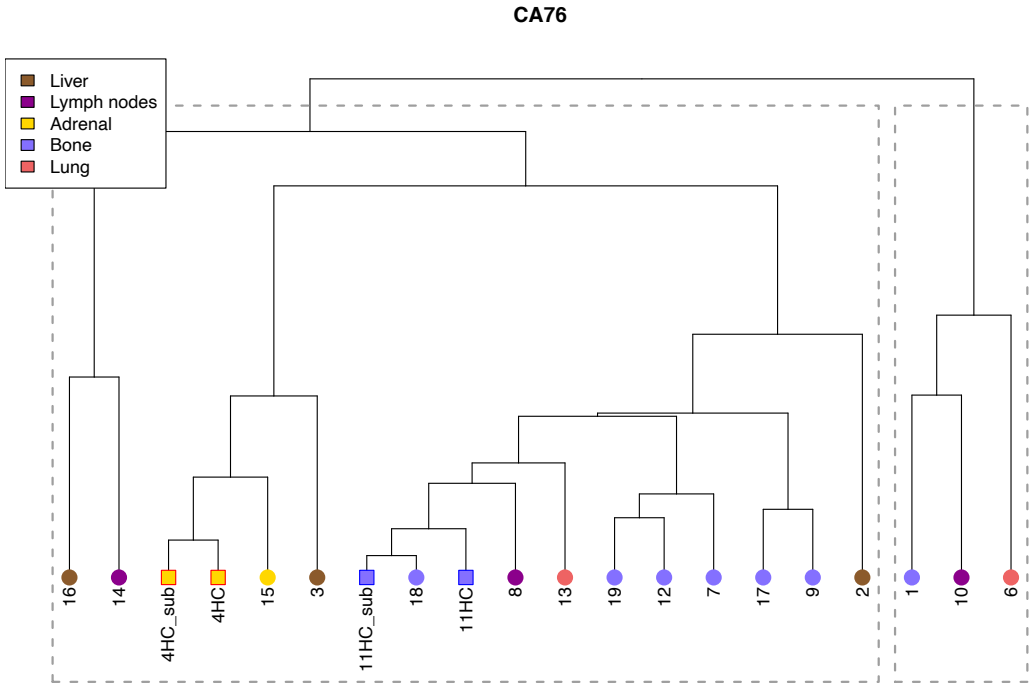
72


CA43
Bone
Liver
Lymph nodes
7
11
3
2
4
8HC
8HC_sub
8
10
6
9
5HC
5
5HC_sub

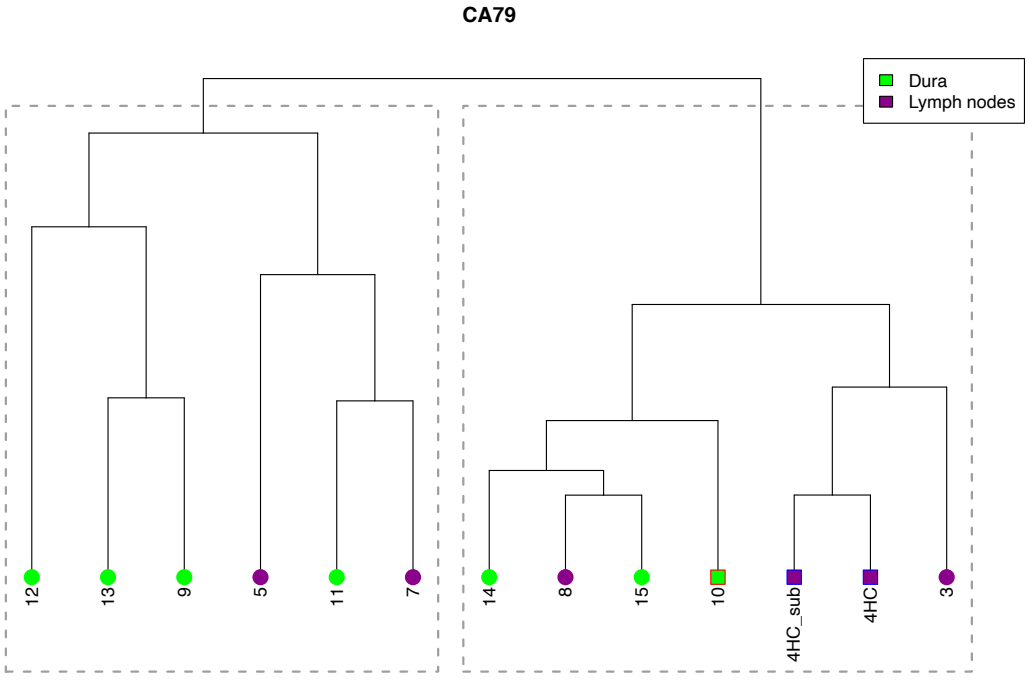
73


CA63
Lymph nodes
Lung
Bone
Ureter
Bladder
Liver
4
16HC_sub
6
34HC_sub
34HC
36
35
16
16HC
7
17
18
19
8
37
32
21
10
5
20
38
9
3
3HC_sub
3HC
39
33
2
2HC_sub
2HC
15
15HC
15HC_sub
12
13HC_sub
13HC
14
28
30
24
27
26
29
31
25
1

74


CA76
Liver
Lymph nodes
Adrenal
Bone
Lung
16
14
4HC_sub
4HC
15
3
11HC_sub
18
11HC
8
13
19
12
7
17
9
2
1
10
6

75


CA79
Dura
Lymph nodes
12
13
9
5
11
7
14
8
15
10
4HC_sub
4HC
3

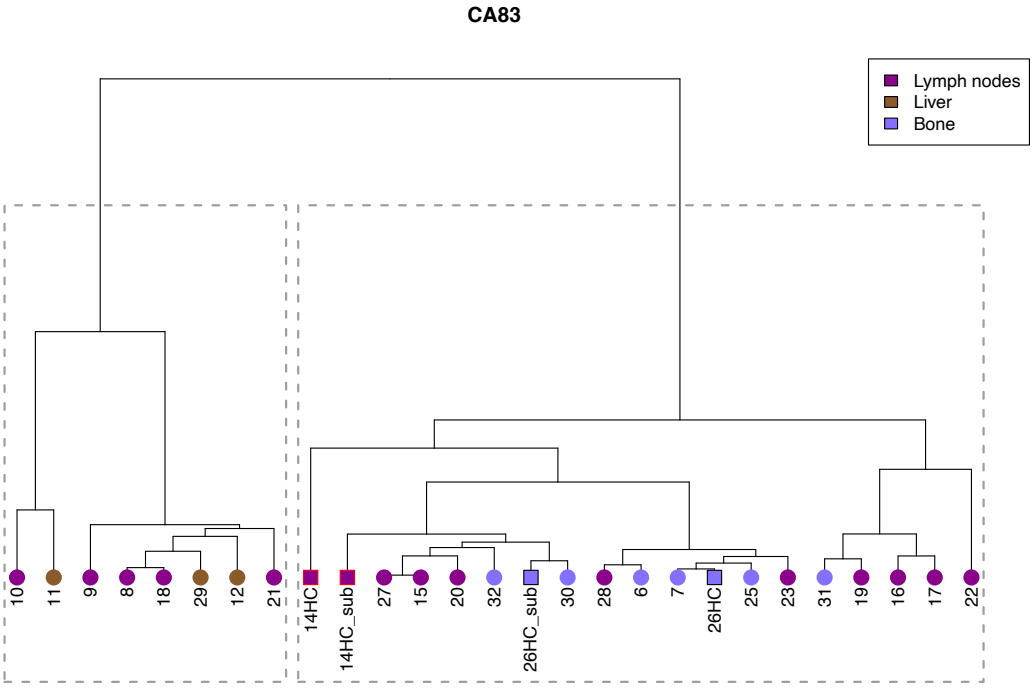



**Figure S6: Relationship assignment of metastases on a SCRATCH relational network is not influenced by sequencing depth**

In SCRATCH relationship networks, tumors with high coverage sequencing ("HC") and/or subsampled version of them ("HC_sub") were assigned predominantly close (or in the same cluster, denoted by grey dashed boxes) to the same samples with low coverage sequencing data. The organ sites are depicted with color-coded circles (color scheme in the legend) on the terminal nodes and the samples in comparison (outlined with same-colored lines) are depicted in same-colored squares.

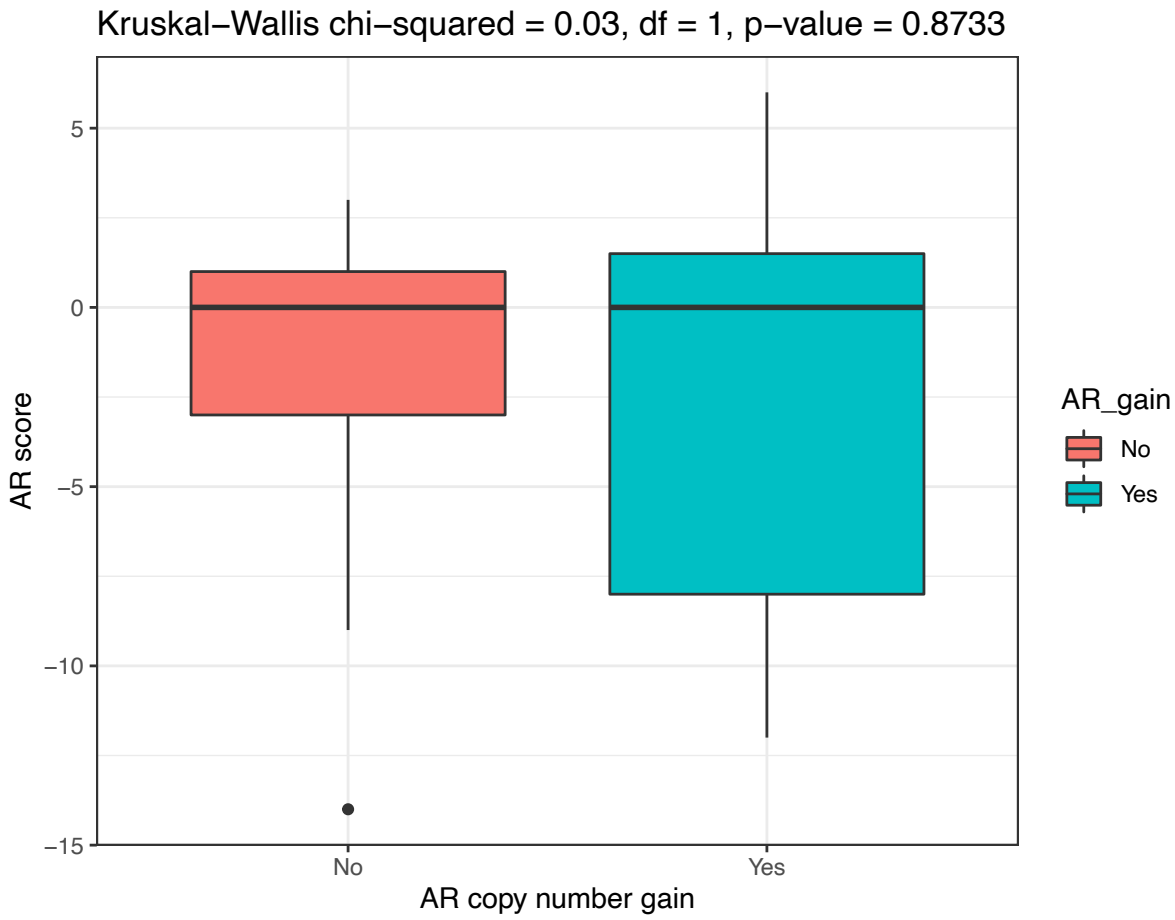



**Figure S7: Expression of AR-target genes are not directly correlated with *AR* expression**

Boxplots of the expression of AR-regulated genes (calculated as AR-score) in metastases with *AR* copy number gain or no gain (significance of difference calculated using Kruskal-Wallis test).

**Supplementary tables**

**Table S1:** Clinical characteristics of patients.

**Table S2:** Detailed sample summary and coverage statistics.

**Table S3**: Allele frequency and copy number of *AR* pE710G and wild-type allele in CA43 (N=6 metastases).

**Table S4:** Clonal somatic mutations detected from high coverage (>30X) next-generation sequencing data for clonality assessments.

**Table S5**: Custom capture panel design. Chromosomal positions are based on human genome assembly GRCh37.