Electronic Supplementary Material (ESI) for Molecular Omics.

Metabolic Deregulation in Prostate Cancer

Srihari S et al. (2018)

Supplementary Figures

Figure S1: Overview of the discovery and validation process for the six metabolic clusters of PCa (a) RNASeq data from 498 PCa patients from TCGA dataset were clustered using hierarchical clustering to identify the best set of clusters (number of clusters k) that gave the most significant separation in terms of disease-free survival outcomes (here, k=6 and logrank-test $p<0.0001$); and **(b)** bionomal classifiers were then trained to separate each such cluster from the others clusters, and the genes that were most significantly ($p<0.05$) associated with these classifications were combined into a multinomial (multi-class) classifier. The multinomial classifer was validated by five-fold cross-validation (80-20%) on TCGA and then retrained on 100% TCGA data and validated on independent datasets including from the Taylors et al. (2010) ($p=0.00088$). Clusters C5 and C3 consistently showed poor prognosis in these datasets.

Figure S2: Curves showing days to complete remission / response for the 6 TCGA clusters upon **(a)** primary drug therapy and followup drug therapy; clusters reproducible from the **(b)** Hieronymus et al. (2014) (overall and/or disease-free survival), **(c)** Ross-Adams et al. (2015) (biochemical relapse), and **(d)** Jain et al. (metastasis) datasets.

Figure S3: Overall and/or disease-free survival (if alive, disease-free survival) of patients divided by percentiles of their metabolic deregulation scores.

Figure S4: 'Oncoprint' genetic alteration profiles for the six clusters using key known genes in prostate cancer.

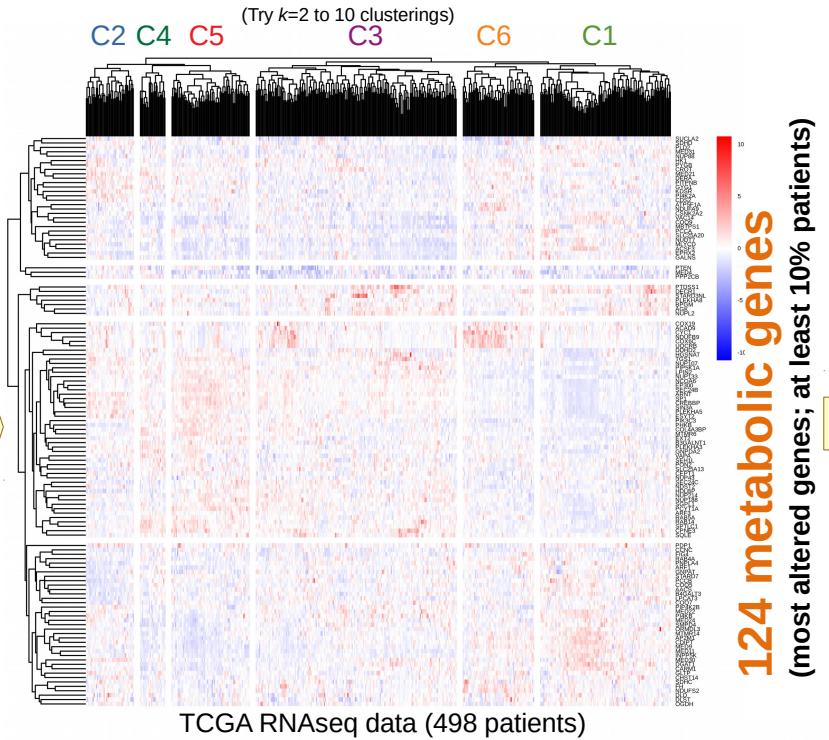
Figure S5: Deregulation of the homologous recombination DNA-damage response pathway in the six clusters.

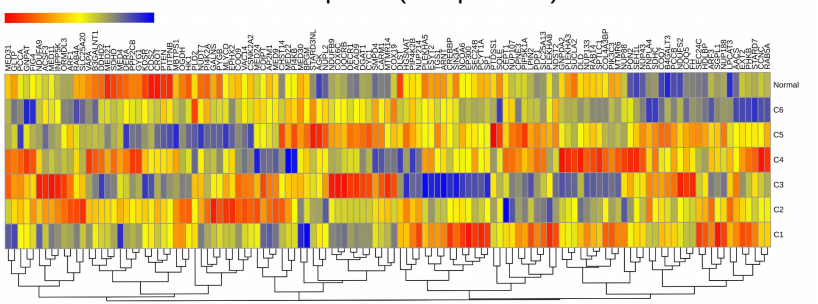
Figure S6: Predicted Sensitive and Resistant subgroups within the six metabolic subtypes for Olaparib response – TCGA and Taylors et al. (2010) and Hieronymus et al. (2014) datasets.

Figure S7: Curves showing **(a)** days to biochemical relapse and **(b)** actual relapse for the 6 TCGA clusters.

Metabolism pathways (>1,000 genes)

[see main text for source]

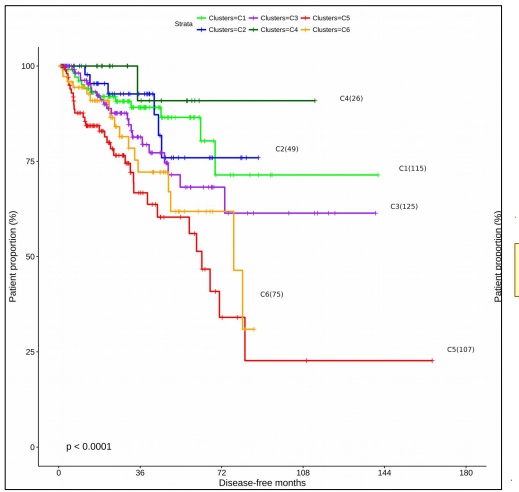




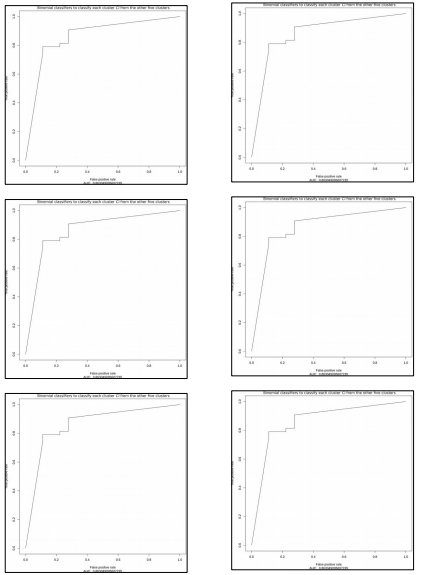
Average expression levels of genes in the six clusters vis-a-vis normal tissues (just to give a comparison; not used in the model)

(a)

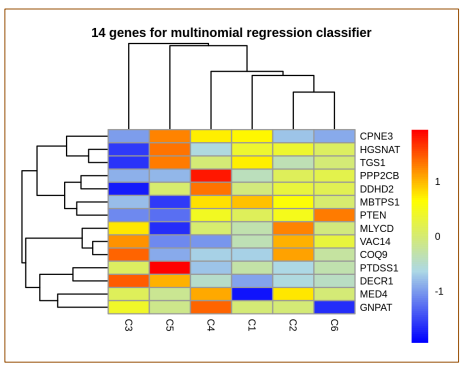
Select the most significant separation (logrank test p-value) in terms of disease-free survival outcomes
(here, k=6 clusters, C1, C2, C3, C4, C5, and C6)



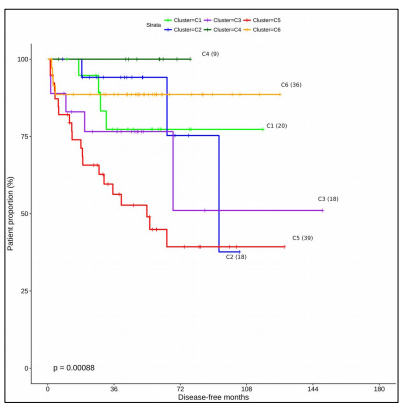
Binomial classifiers to classify each cluster C_i from the remaining five clusters



Select genes into multinomial classifier



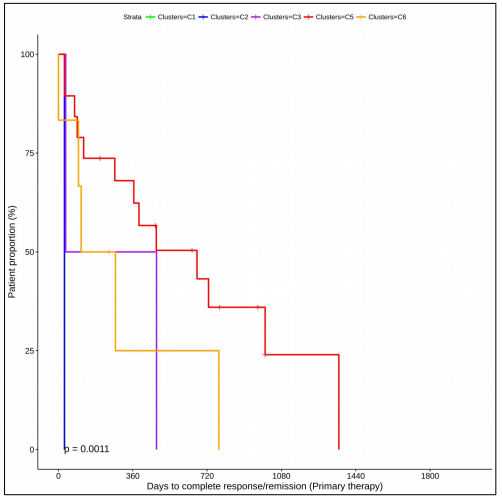
(b)



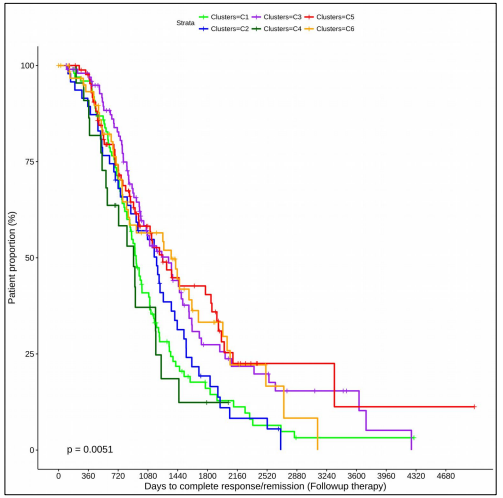
Validate multinomial classifier on TCGA by five-fold cross-validation

Validate on independent datasets

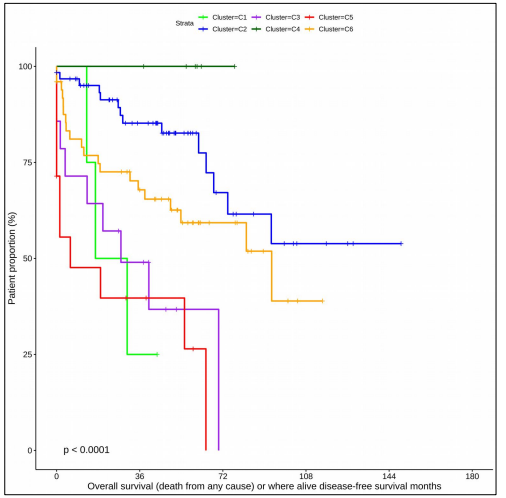
Figure S1



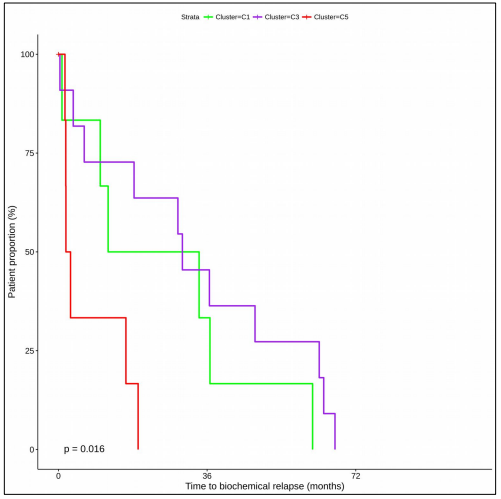
(a) Response to primary therapy (TCGA)



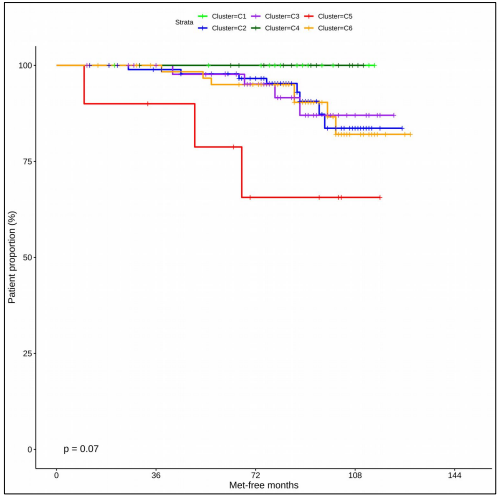
Response to followup therapy (TCGA)



(b) Hieronymus et al. (2014) – OS / DF



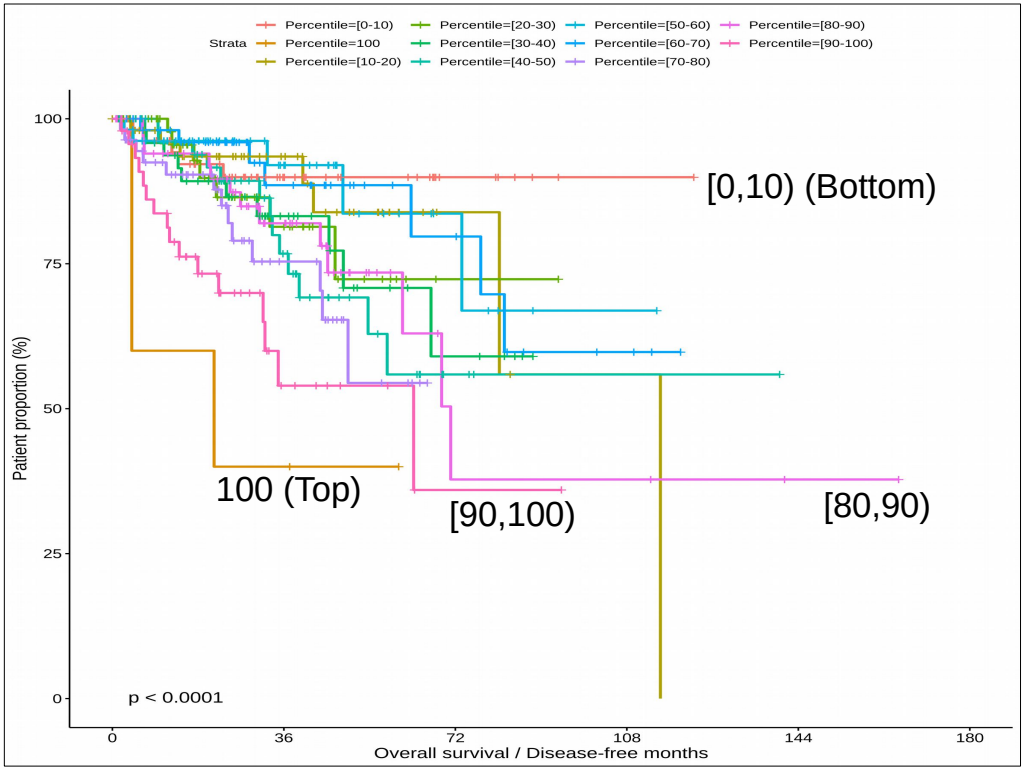
(c) Ross-Adams et al. (2015) - BCR



(d) Jain et al. (2017) – Met

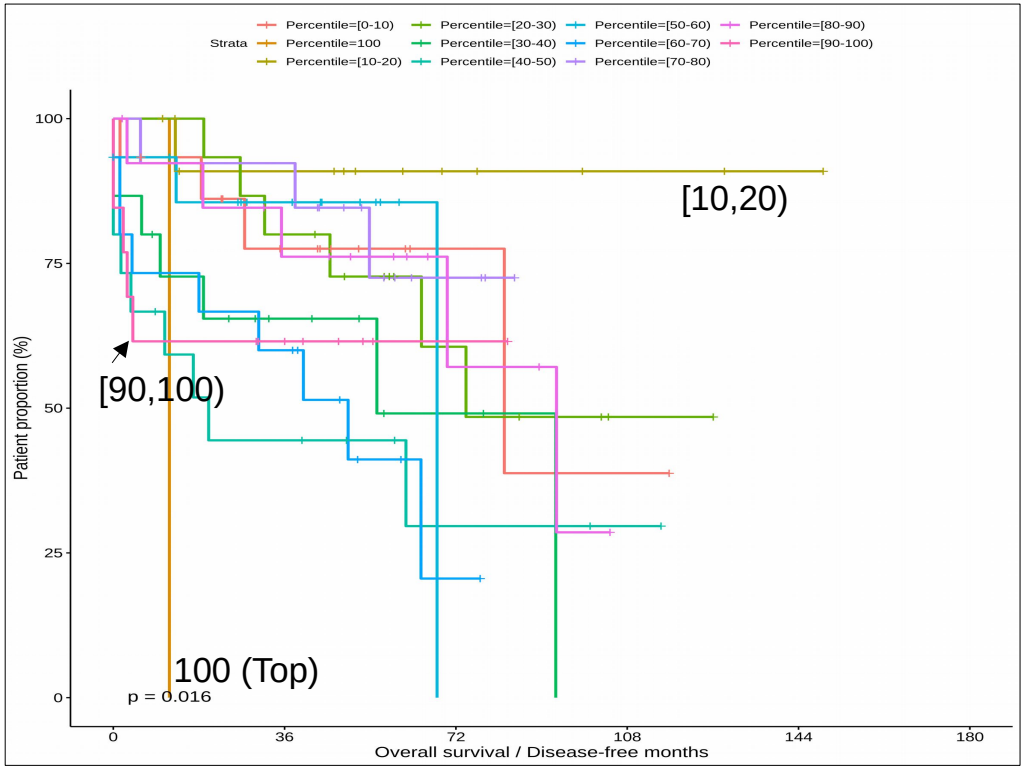
OS = Overall survival
DF = Disease-free survival
OS/DF = OS but if alive, DF
Met = Metastasis

Figure S2



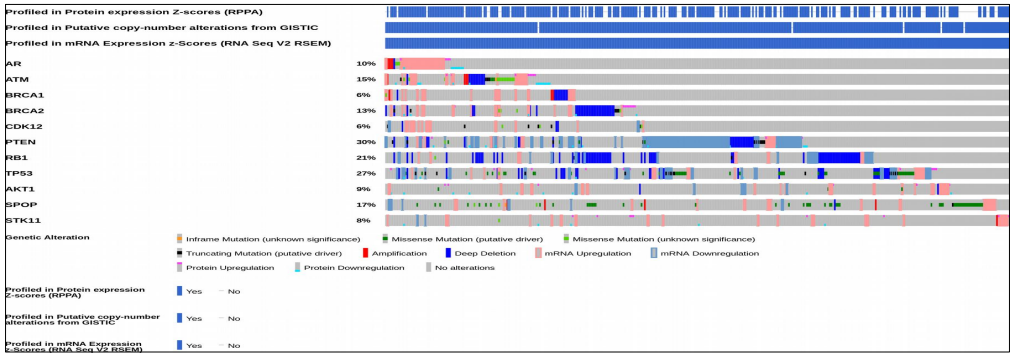
TCGA (498 patients)

OS/DFS months

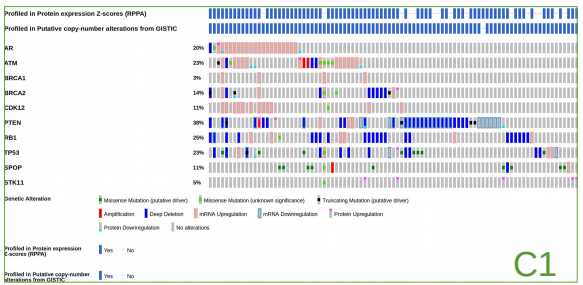


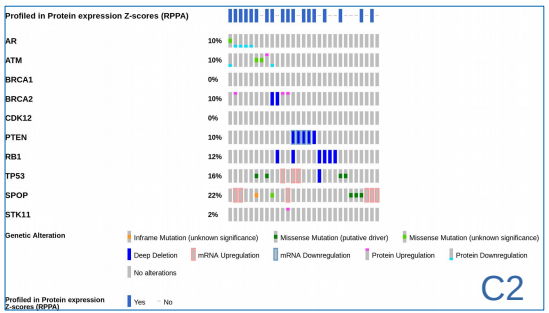
MSKCC Hieronymus et al. (150 patients)

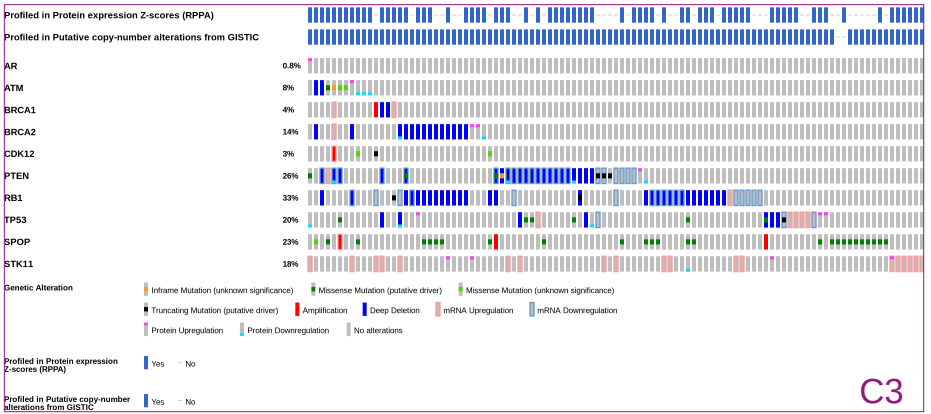
Figure S3

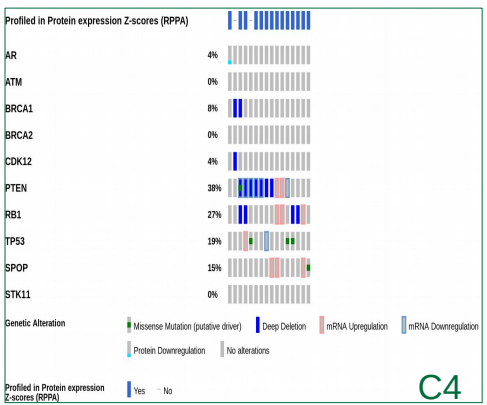


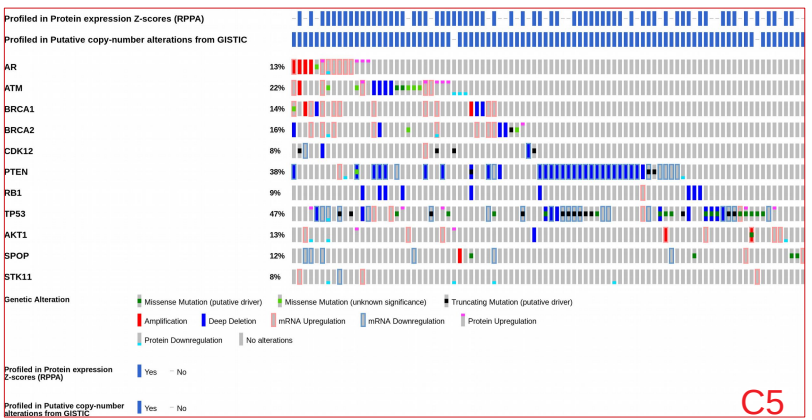
All TCGA tumours











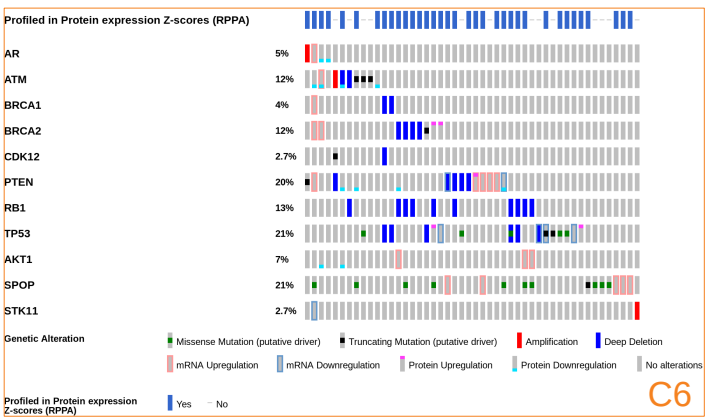
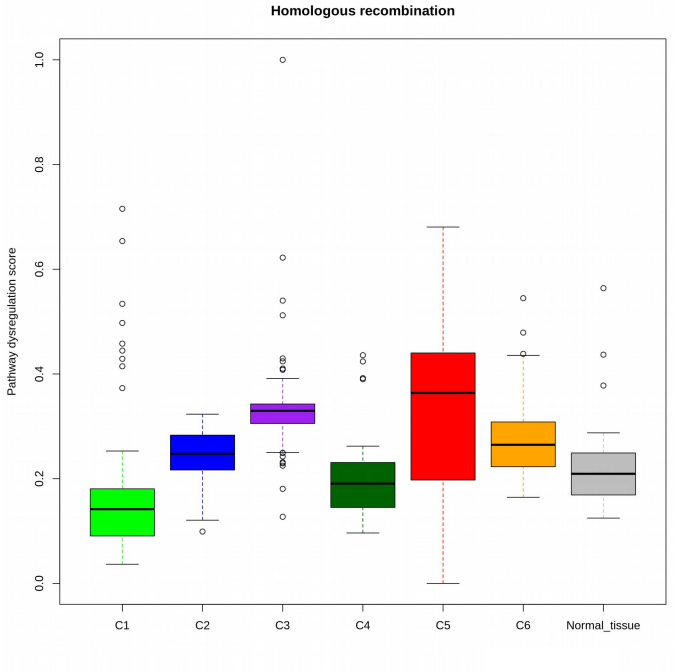
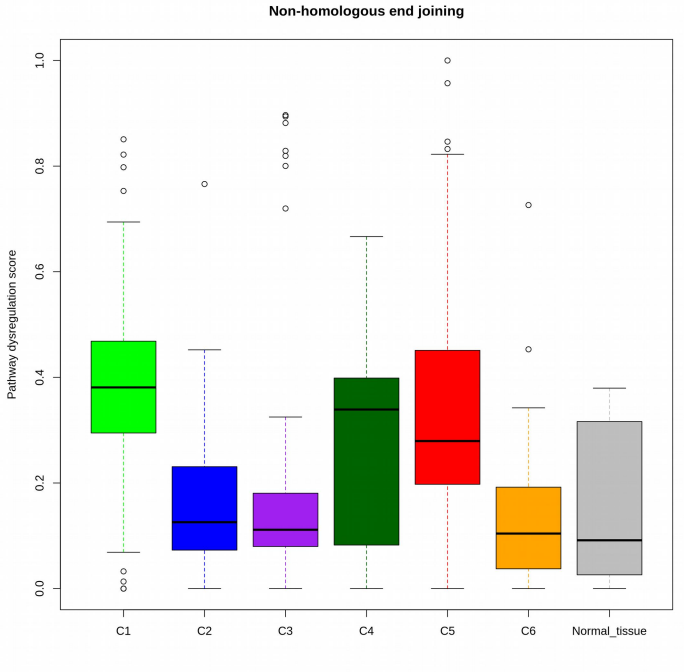


Figure S4


Homologous recombination
Pathway dysregulation score
0.0
0.2
0.4
0.6
0.8
1.0
C1
C2
C3
C4
C5
C6
Normal_tissue


Non-homologous end joining
Pathway dysregulation score
0.0
0.2
0.4
0.6
0.8
1.0
C1
C2
C3
C4
C5
C6
Normal_tissue

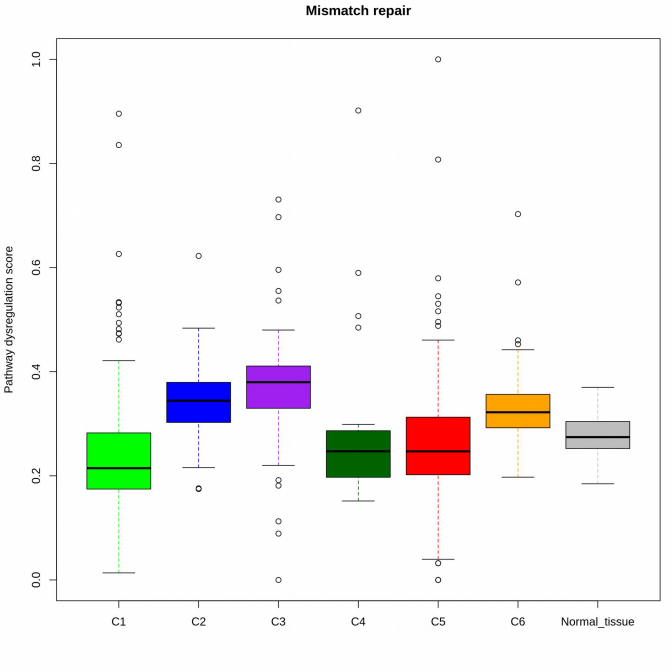
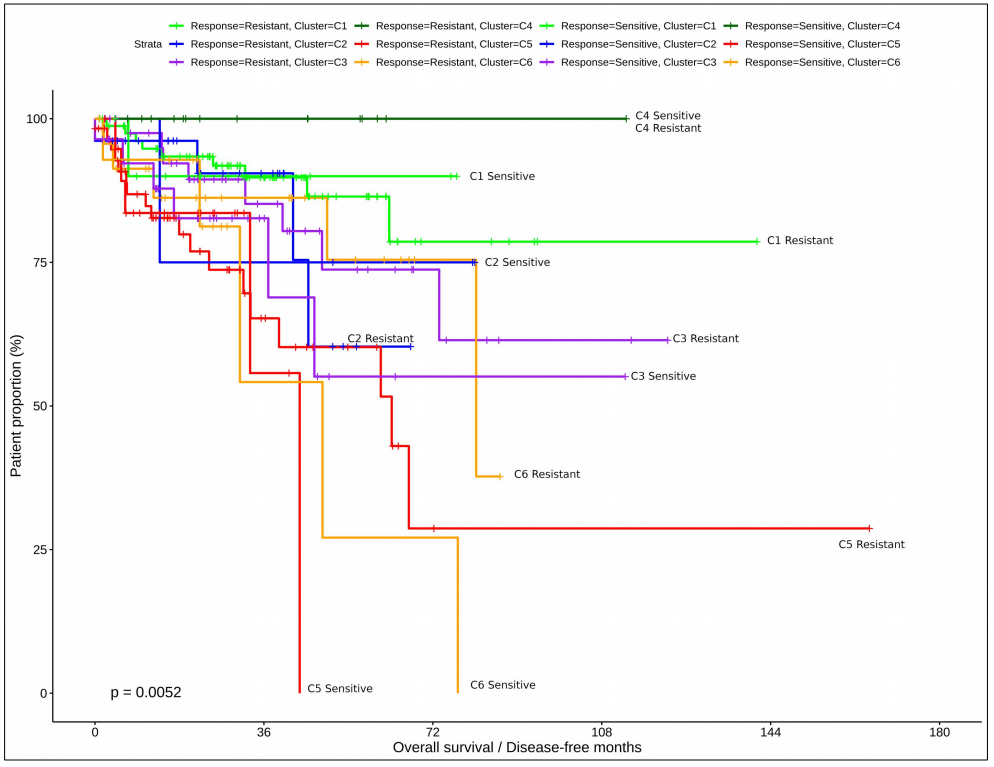
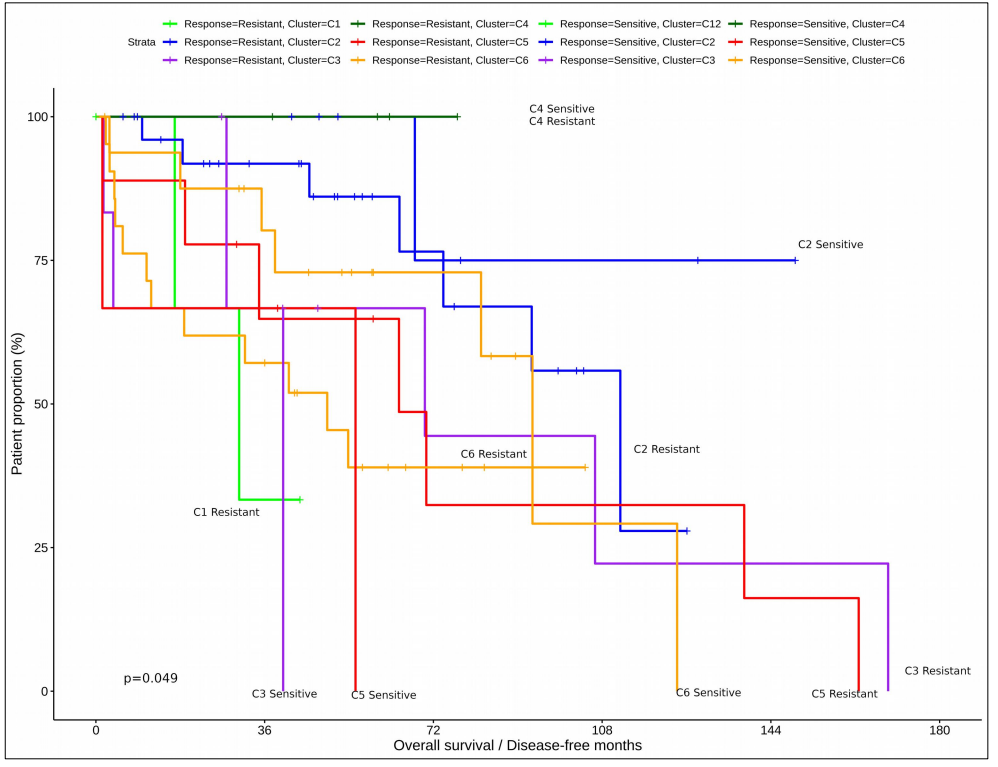

Mismatch repair
Pathway dysregulation score
0.0
0.2
0.4
0.6
0.8
1.0
C1
C2
C3
C4
C5
C6
Normal_tissue

Figure S5

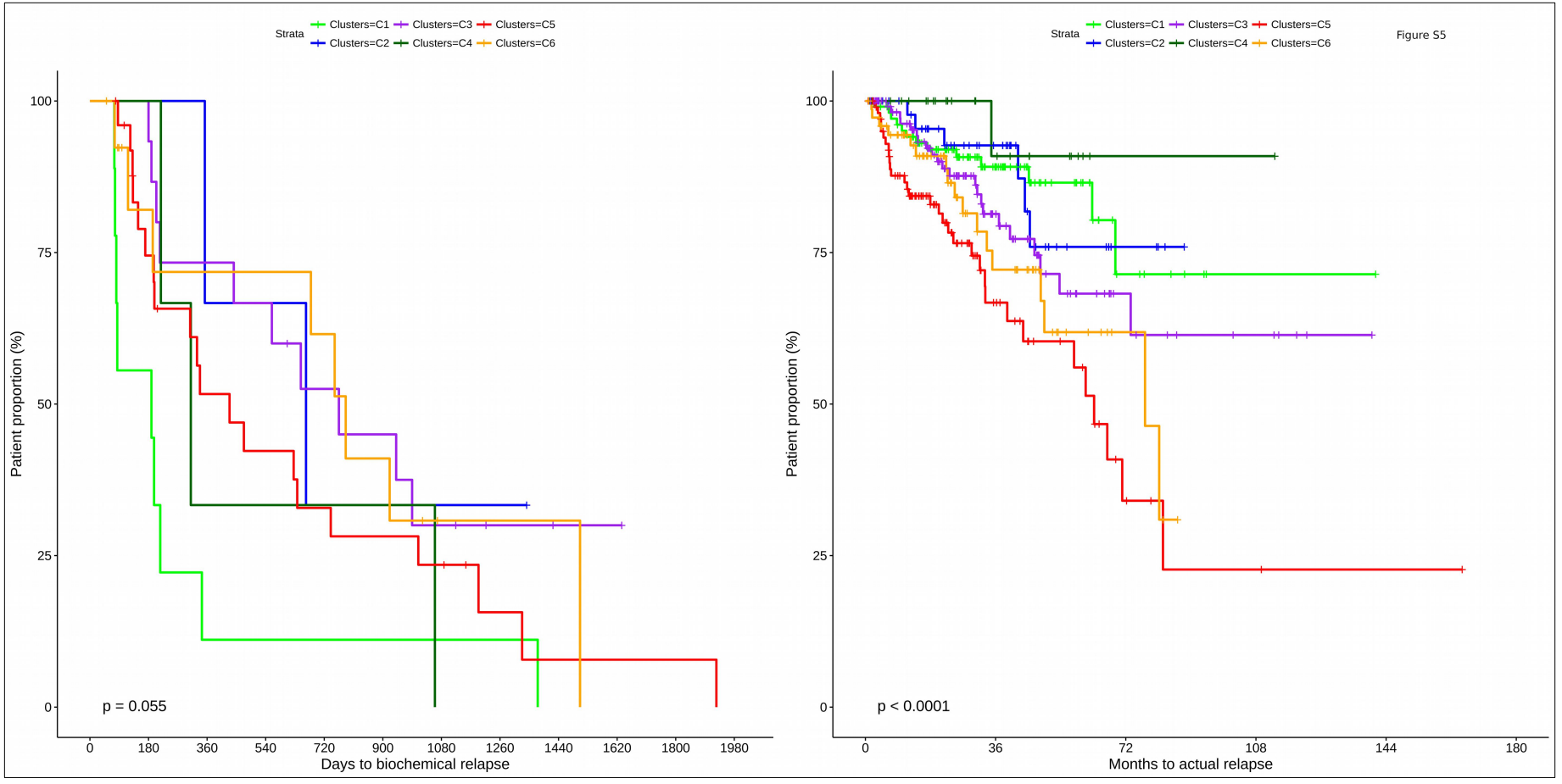


TCGA – DF / OS



MSKCC (Hieronymus et al. (2014)) – DF / OS

Figure S6


Strata
Clusters=C1
Clusters=C2
Clusters=C3
Clusters=C4
Clusters=C5
Clusters=C6
Patient proportion (%)
100
75
50
25
0
p = 0.055
0
180
360
540
720
900
1080
1260
1440
1620
1800
1980
Days to biochemical relapse
Figure S5
p < 0.0001
36
72
108
144
180
Months to actual relapse

(a)

(b)

Figure S7

Supplementary Tables

Table S1: List of genes involved in the 20 metabolic pathways studied in this work.

Table S2: Patient clusters identified from TCGA and other datasets – Taylors et al. (2010), Hieronymus et al. (2014), Ross-Adams et al. (2015), and Jain et al. (2018) datasets along with relevant clinical information.

Table S3: Enrichments (hypergeometric test p-values) for genetic alterations in the six clusters.

Table S4: proportions of ACRPC and AS patients from the Olmos et al. (2012) dataset that were predicted to be in the six PCa clusters.