

Summary of the curatedPCaData-package

Teemu Daniel Laajala

August 31, 2021

Contents

1	Summary of all MAE-objects	2
1.1	End-points	2
1.1.1	Gleason grades	2
1.1.2	Overall Survival	3
1.1.3	Recurrence	3
1.1.4	Other end-points	4
1.1.5	MAE-slots	4
1.1.6	MAE-slot classes	5
2	Individual MAE colData-summaries	8
2.1	abida	8
2.2	barbieri	10
2.3	chandran	11
2.4	friedrich	13
2.5	hieronymus	14
2.6	icgcca	15
2.7	igc	17
2.8	kim	18
2.9	kunderfranco	19
2.10	ren	21
2.11	sun	22
2.12	taylor	23
2.13	tcga	25
2.14	true	26
2.15	wallace	28
2.16	wang	29
2.17	weiner	30
3	Omics sample overlap within MAE-objects	32
3.1	abida	32
3.2	barbieri	33
3.3	chandran	34
3.4	friedrich	35
3.5	hieronymus	36
3.6	icgcca	37
3.7	igc	38
3.8	kim	39

3.9	kunderfranco	40
3.10	ren	41
3.11	sun	42
3.12	taylor	43
3.13	tcga	44
3.14	true	45
3.15	wallace	46
3.16	wang	47
3.17	weiner	48

This is a summary of the MultiAssayExperiment-objects (MAE) offered for curated prostate cancer datasets in the *curatedPCaData*-package version 0.6.8. The following objects are currently available in the package:

```
> utils::data(package="curatedPCaData")$result[,c("Item", "Title")]

      Item
[1,] "mae_abida"
[2,] "mae_barbieri"
[3,] "mae_chandran"
[4,] "mae_friedrich"
[5,] "mae_hieronymus"
[6,] "mae_icgcca"
[7,] "mae_igc"
[8,] "mae_kim"
[9,] "mae_kunderfranco"
[10,] "mae_ren"
[11,] "mae_sun"
[12,] "mae_taylor"
[13,] "mae_tcga"
[14,] "mae_true"
[15,] "mae_wallace"
[16,] "mae_wang"
[17,] "mae_weiner"
      Title
[1,] "Abida et al. MAE-object"
[2,] "Barbieri MAE-object"
[3,] "Barwick MAE-object"
[4,] "Friedrich et al. MAE-object"
[5,] "Hieronymus et al. MAE-object"
[6,] "ICGC CA MAE-object"
[7,] ""
[8,] ""
[9,] "Kunderfranco et al. MAE-object"
[10,] "Ren 2017 MAE-object"
[11,] "Sun et al. MAE-object"
[12,] "Taylor et al. MAE-object"
[13,] "TCGA MAE-object"
[14,] "True et al. MAE-object"
[15,] "Wallace et al. MAE-object"
[16,] ""
[17,] "MultiAssayExperiment object containing gene expression (gex),"
```

They can be accessed via syntax `curatedPCaData::mae_name` or by first calling `library("curatedPCaData")` and then calling directly the objects `mae_name` in the workspace environment.

1 Summary of all MAE-objects

Summary table of all MAE-objects and availability of most important variables.

1.1 End-points

1.1.1 Gleason grades

Gleason grade distribution over studies:

	5	6	7	8	9	10	Other	N/A
abida	-	-	-	-	-	-	0 (0%)	444 (100%)
barbieri	-	13 (12%)	84 (77%)	8 (7%)	4 (4%)	-	0 (0%)	0 (0%)
chandran	6 (1%)	46 (9%)	85 (17%)	21 (4%)	35 (7%)	-	3 (1%)	307 (61%)
friedrich	2 (1%)	47 (18%)	54 (21%)	68 (27%)	43 (17%)	2 (1%)	0 (0%)	39 (15%)
hieronymus	-	16 (15%)	78 (75%)	4 (4%)	6 (6%)	-	0 (0%)	0 (0%)
icgcca	-	12 (6%)	58 (27%)	5 (2%)	-	-	0 (0%)	138 (65%)
igc	-	27 (33%)	40 (48%)	13 (16%)	-	-	0 (0%)	3 (4%)
kim	-	-	-	-	-	-	266 (100%)	0 (0%)
kunderfranco	1 (1%)	9 (13%)	32 (48%)	6 (9%)	5 (7%)	-	0 (0%)	14 (21%)
ren	-	-	-	-	-	-	65 (100%)	0 (0%)
sun	-	-	-	-	-	-	0 (0%)	79 (100%)
taylor	2 (1%)	104 (48%)	77 (35%)	19 (9%)	15 (7%)	-	0 (0%)	1 (0%)
tcga	-	45 (9%)	244 (49%)	64 (13%)	137 (28%)	4 (1%)	0 (0%)	0 (0%)
true	-	4 (14%)	20 (69%)	1 (3%)	4 (14%)	-	0 (0%)	0 (0%)
wallace	2 (2%)	21 (25%)	57 (69%)	1 (1%)	2 (2%)	-	0 (0%)	0 (0%)
wang	-	-	-	-	-	-	0 (0%)	148 (100%)
weiner	-	-	-	-	-	-	0 (0%)	838 (100%)

1.1.2 Overall Survival

Quantiles [min, 25%, median, 75%, max] for follow-up times and 0 (no event) and 1 (event) for survival:

	0 (no event)	1 (event)	N/A (event)	Time (d)	N/A (time)
abida	52 (12%)	84 (19%)	308 (69%)	[50,321,595,883,2070]	308 (69%)
barbieri	-	-	109 (100%)	-	109 (100%)
chandran	-	-	503 (100%)	-	503 (100%)
friedrich	230 (90%)	25 (10%)	0 (0%)	[641,3005,3614,4301,6771]	91 (36%)
hieronymus	96 (92%)	8 (8%)	0 (0%)	[295,1575,2139,2895,3758]	0 (0%)
icgcca	198 (93%)	8 (4%)	7 (3%)	[1460,2190,2920,3650,4745]	1 (0%)
igc	-	-	83 (100%)	-	83 (100%)
kim	-	-	266 (100%)	-	266 (100%)
kunderfranco	-	-	67 (100%)	-	67 (100%)
ren	-	-	65 (100%)	-	65 (100%)
sun	-	-	79 (100%)	-	79 (100%)
taylor	-	-	218 (100%)	-	218 (100%)
tcga	484 (98%)	10 (2%)	0 (0%)	[23,525,926,1466,5034]	0 (0%)
true	-	-	29 (100%)	-	29 (100%)
wallace	-	-	83 (100%)	-	83 (100%)
wang	-	-	148 (100%)	-	148 (100%)
weiner	-	-	838 (100%)	-	838 (100%)

1.1.3 Recurrence

Quantiles [min, 25%, median, 75%, max] for follow-up times and counts for 0 (no event) and 1 (event) for recurrence:

	0 (no event)	1 (event)	N/A (event)	Time (d)	N/A (time)
abida	-	-	444 (100%)	-	444 (100%)
barbieri	-	-	109 (100%)	-	109 (100%)
chandran	-	-	503 (100%)	-	503 (100%)
friedrich	-	-	255 (100%)	-	255 (100%)
hieronymus	-	-	104 (100%)	-	104 (100%)
icgcca	-	-	213 (100%)	-	213 (100%)
igc	-	-	83 (100%)	-	83 (100%)
kim	-	-	266 (100%)	-	266 (100%)
kunderfranco	-	-	67 (100%)	-	67 (100%)
ren	-	-	65 (100%)	-	65 (100%)
sun	40 (51%)	39 (49%)	0 (0%)	-	79 (100%)
taylor	137 (63%)	61 (28%)	20 (9%)	[3,717,1386,1974,4909]	20 (9%)
tcga	397 (80%)	91 (18%)	6 (1%)	[23,427,823,1376,5034]	6 (1%)
true	-	-	29 (100%)	-	29 (100%)
wallace	-	-	83 (100%)	-	83 (100%)
wang	-	-	148 (100%)	-	148 (100%)
weiner	-	-	838 (100%)	-	838 (100%)

1.1.4 Other end-points

TODO, for example the newer grading system available in Weiner et al.

1.1.5 MAE-slots

The following MAE-slots are available in each study:

	mae_abida	mae_barbieri	mae_chandran	mae_friedrich	mae_hieronymus	mae_icgcca	mae_igc	mae_kim	mae_kunderfranco	mae_ren	mae_sun	mae_taylor	mae_tcga	mae_true	mae_wallace	mae_wang	mae_weiner
cibersort		x	x	x		x	x	x	x	x	x	x	x		x	x	
cibersort_polyA	x																
cna	x				x												
cna_gistic		x								x							
cna_log												x	x				
epic		x	x			x	x	x	x	x		x	x	x		x	x
epic_polyA	x																
gex		x	x	x		x	x	x	x	x	x	x	x	x	x	x	x
gex_polyA	x																
mcp		x	x			x	x	x	x	x	x	x	x	x	x	x	x
mcp_polyA	x																
mut	x									x		x	x				
quantiseq		x	x			x	x	x	x	x	x	x	x	x	x	x	x
quantiseq_polyA	x																
scores	x	x	x			x	x	x	x	x	x	x		x	x	x	x
xcell		x	x			x	x	x	x	x		x	x			x	x
xcell_polyA	x																

1.1.6 MAE-slot classes

The classes for each slot for datasets

abida

```

cibersort_polyA: matrix, array
cna: matrix, array
gex_polyA: matrix, array
mut: matrix, array
xcell_polyA: matrix, array
epic_polyA: matrix, array
quantiseq_polyA: matrix, array
mcp_polyA: matrix, array
scores: matrix, array

```

barbieri

```

cibersort: matrix, array
cna_gistic: matrix, array
gex: matrix, array
xcell: matrix, array
epic: matrix, array
quantiseq: matrix, array
mcp: matrix, array
scores: matrix, array

```

chandran

```

gex: matrix, array
xcell: matrix, array
epic: matrix, array
mcp: matrix, array
scores: matrix, array
cibersort: matrix, array
quantiseq: matrix, array

friedrich
  cibersort: matrix, array
  gex: matrix, array

hieronymus
  cna: matrix, array

icgcca
  gex: matrix, array
  xcell: matrix, array
  epic: matrix, array
  quantiseq: matrix, array
  mcp: matrix, array
  scores: matrix, array
  cibersort: matrix, array

igc
  gex: matrix, array
  scores: matrix, array
  cibersort: matrix, array
  xcell: matrix, array
  epic: matrix, array
  quantiseq: matrix, array
  mcp: matrix, array

kim
  gex: matrix, array
  scores: matrix, array
  cibersort: matrix, array
  xcell: matrix, array
  epic: matrix, array
  quantiseq: matrix, array
  mcp: matrix, array

kunderfranco
  cibersort: matrix, array
  gex: data.frame
  scores: matrix, array
  xcell: matrix, array
  epic: matrix, array
  quantiseq: matrix, array
  mcp: matrix, array

```

```

ren
  cna_gistic: matrix, array
  gex: matrix, array
  mut: matrix, array
  xcell: matrix, array
  epic: matrix, array
  quantiseq: matrix, array
  mcp: matrix, array
  scores: matrix, array
  cibersort: matrix, array

sun
  gex: matrix, array
  mcp: matrix, array
  quantiseq: matrix, array
  scores: matrix, array
  cibersort: matrix, array

taylor
  cibersort: matrix, array
  cna_log: matrix, array
  gex: matrix, array
  mut: matrix, array
  scores: matrix, array
  mcp: matrix, array
  quantiseq: matrix, array
  epic: matrix, array
  xcell: matrix, array

tcga
  cna_log: matrix, array
  gex: matrix, array
  mut: matrix, array
  cibersort: matrix, array
  epic: matrix, array
  mcp: matrix, array
  quantiseq: matrix, array
  xcell: matrix, array

true
  gex: data.frame
  scores: matrix, array
  epic: matrix, array
  mcp: matrix, array
  quantiseq: matrix, array

wallace
  gex: matrix, array
  quantiseq: matrix, array
  mcp: matrix, array
  scores: matrix, array
  cibersort: matrix, array

```

```
wang
  gex: matrix, array
  scores: matrix, array
  cibersort: matrix, array
  xcell: matrix, array
  epic: matrix, array
  quantiseq: matrix, array
  mcp: matrix, array
```

```
weiner
  epic: matrix, array
  gex: matrix, array
  mcp: matrix, array
  quantiseq: matrix, array
  xcell: matrix, array
  scores: matrix, array
```

2 Individual MAE colData-summaries

Looping over and summarizing the datasets:

2.1 abida

	Instances	Missingness
overall_survival_status	NA: 308 (69%);1: 84 (19%);0: 52 (12%)	n=308 (69%)
days_to_overall_survival	[50.26,321.04,594.78,882.8,2069.65], NA n=308	n=308 (69%)
age_at_initial_diagnosis	[38,56,61,66,89], NA n=51	n=51 (11%)
year_diagnosis	N/A	n=444 (100%)
gleason_grade	N/A	n=444 (100%)
gleason_major	N/A	n=444 (100%)
gleason_minor	N/A	n=444 (100%)
source_of_gleason	N/A	n=444 (100%)
grade_group	N/A	n=444 (100%)
T_pathological	N/A	n=444 (100%)
T_substage_pathological	N/A	n=444 (100%)
T_clinical	N/A	n=444 (100%)
T_substage_clinical	N/A	n=444 (100%)
ERG_fusion_CNA	N/A	n=444 (100%)
ERG_fusion_IHC	N/A	n=444 (100%)
ERG_fusion_GEX	N/A	n=444 (100%)
disease_specific_recurrence_status	N/A	n=444 (100%)
days_to_disease_specific_recurrence	N/A	n=444 (100%)
metastasis_occurrence_status	N/A	n=444 (100%)
days_to_metastatic_occurrence	N/A	n=444 (100%)
psa	[0.2,5.7,14.04,52.84,3118], NA n=91	-
race	N/A	n=444 (100%)
smoking_status	N/A	n=444 (100%)
extraprostatic_extension	N/A	n=444 (100%)

perineural_invasion	N/A	n=444 (100%)
seminal_vesicle_invasion	N/A	n=444 (100%)
angiolymphatic_invasion	N/A	n=444 (100%)
androgen_ablation	N/A	n=444 (100%)
capsule	N/A	n=444 (100%)
M_stage	N/A	n=444 (100%)
M_substage	N/A	n=444 (100%)
other_patient	N/A	n=444 (100%)
sample_type	N/A	n=444 (100%)
genomic_alterations	N/A	n=444 (100%)
tumor_margins_positive	N/A	n=444 (100%)
tissue_source	LN: 167 (38%);Bone: 160 (36%);Liver: 64 (14%); ...	-
metastatic_site	N/A	n=444 (100%)
microdissected	N/A	n=444 (100%)
frozen_ffpe	N/A	n=444 (100%)
other_feature	N/A	n=444 (100%)
batch	N/A	n=444 (100%)
other_sample	N/A	n=444 (100%)
tumor_purity_pathology	N/A	n=444 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=444 (100%)
zone_of_origin	N/A	n=444 (100%)
zone_of_origin_estimated	N/A	n=444 (100%)
mutational_signatures	N/A	n=444 (100%)
neoantigen_load	N/A	n=444 (100%)
AR_activity	N/A	n=444 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=444 (100%)
N_substage	N/A	n=444 (100%)
therapy_radiation_initial	N/A	n=444 (100%)
therapy_radiation_salvage	N/A	n=444 (100%)
therapy_surgery_initial	N/A	n=444 (100%)
therapy_hormonal_initial	N/A	n=444 (100%)
other_treatment	N/A	n=444 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.2 barbieri

	Instances	Missingness
overall_survival_status	N/A	n=109 (100%)
days_to_overall_survival	N/A	n=109 (100%)
age_at_initial_diagnosis	[34, 58, 63, 68, 77]	-
year_diagnosis	N/A	n=109 (100%)
gleason_grade	[6, 7, 7, 7, 9]	-
gleason_major	3: 70 (64%);4: 39 (36%)	-

gleason_minor	4: 65 (60%);3: 40 (37%);5: 4 (4%)	-
source_of_gleason	N/A	n=109 (100%)
grade_group	3+4: 57 (52%);4+3: 27 (25%);<=6: 13 (12%); ...	-
T_pathological	N/A	n=109 (100%)
T_substage_pathological	N/A	n=109 (100%)
T_clinical	3: 67 (61%);2: 42 (39%)	-
T_substage_clinical	a: 53 (49%);c: 37 (34%);b: 19 (17%)	-
ERG_fusion_CNA	N/A	n=109 (100%)
ERG_fusion_IHC	N/A	n=109 (100%)
ERG_fusion_GEX	N/A	n=109 (100%)
disease_specific_recurrence_status	N/A	n=109 (100%)
days_to_disease_specific_recurrence	N/A	n=109 (100%)
metastasis_occurrence_status	N/A	n=109 (100%)
days_to_metastatic_occurrence	N/A	n=109 (100%)
psa	[2.7,5.57,7.8,10.85,31.5], NA n=1	n=1 (1%)
race	N/A	n=109 (100%)
smoking_status	N/A	n=109 (100%)
extraprostatic_extension	N/A	n=109 (100%)
perineural_invasion	N/A	n=109 (100%)
seminal_vesicle_invasion	N/A	n=109 (100%)
angiolympathic_invasion	N/A	n=109 (100%)
androgen_ablation	N/A	n=109 (100%)
capsule	N/A	n=109 (100%)
M_stage	N/A	n=109 (100%)
M_substage	N/A	n=109 (100%)
other_patient	N/A	n=109 (100%)
sample_type	N/A	n=109 (100%)
genomic_alterations	N/A	n=109 (100%)
tumor_margins_positive	N/A	n=109 (100%)
tissue_source	N/A	n=109 (100%)
metastatic_site	N/A	n=109 (100%)
microdissected	N/A	n=109 (100%)
frozen_ffpe	N/A	n=109 (100%)
other_feature	N/A	n=109 (100%)
batch	N/A	n=109 (100%)
other_sample	N/A	n=109 (100%)
tumor_purity_pathology	N/A	n=109 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=109 (100%)
zone_of_origin	N/A	n=109 (100%)
zone_of_origin_estimated	N/A	n=109 (100%)
mutational_signatures	N/A	n=109 (100%)
neoantigen_load	N/A	n=109 (100%)
AR_activity	N/A	n=109 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=109 (100%)
N_substage	N/A	n=109 (100%)

therapy_radiation_initial	N/A	n=109 (100%)
therapy_radiation_salvage	N/A	n=109 (100%)
therapy_surgery_initial	N/A	n=109 (100%)
therapy_hormonal_initial	N/A	n=109 (100%)
other_treatment	N/A	n=109 (100%)
psa_category	N/A	n=109 (100%)
genome_altered	Feature not found in MAE	-

2.3 chandran

	Instances	Missingness
overall_survival_status	N/A	n=503 (100%)
days_to_overall_survival	N/A	n=503 (100%)
age_at_initial_diagnosis	[13,20,24,48,63], NA n=452	n=452 (90%)
year_diagnosis	N/A	n=503 (100%)
gleason_grade	[4,6,7,8,9], NA n=307	n=307 (61%)
gleason_major	N/A	n=503 (100%)
gleason_minor	N/A	n=503 (100%)
source_of_gleason	N/A	n=503 (100%)
grade_group	N/A	n=503 (100%)
T_pathological	[2,2,3,3,4], NA n=307	n=307 (61%)
T_substage_pathological	NA: 310 (62%);b: 102 (20%);a: 91 (18%)	n=310 (62%)
T_clinical	N/A	n=503 (100%)
T_substage_clinical	N/A	n=503 (100%)
ERG_fusion_CNA	N/A	n=503 (100%)
ERG_fusion_IHC	N/A	n=503 (100%)
ERG_fusion_GEX	N/A	n=503 (100%)
disease_specific_recurrence_status	N/A	n=503 (100%)
days_to_disease_specific_recurrence	N/A	n=503 (100%)
metastasis_occurrence_status	N/A	n=503 (100%)
days_to_metastatic_occurrence	N/A	n=503 (100%)
psa	N/A	n=503 (100%)
race	NA: 452 (90%);caucasian: 45 (9%);african_american: 6 (1%)	n=452 (90%)
smoking_status	N/A	n=503 (100%)
extraprostatic_extension	N/A	n=503 (100%)
perineural_invasion	N/A	n=503 (100%)
seminal_vesicle_invasion	N/A	n=503 (100%)
angiolymphatic_invasion	N/A	n=503 (100%)
androgen_ablation	N/A	n=503 (100%)
capsule	N/A	n=503 (100%)
M_stage	N/A	n=503 (100%)
M_substage	N/A	n=503 (100%)
other_patient	N/A	n=503 (100%)
sample_type	primary: 196 (39%);adjacentnormal: 181 (36%);metastatic: 75 (15%); ...	-
genomic_alterations	N/A	n=503 (100%)
tumor_margins_positive	N/A	n=503 (100%)
tissue_source	N/A	n=503 (100%)

metastatic_site	NA: 428 (85%);lymph_node: 45 (9%);liver: 15 (3%); ...	n=428 (85%)
microdissected	N/A	n=503 (100%)
frozen_ffpe	N/A	n=503 (100%)
other_feature	N/A	n=503 (100%)
batch	N/A	n=503 (100%)
other_sample	N/A	n=503 (100%)
tumor_purity_pathology	N/A	n=503 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=503 (100%)
zone_of_origin	N/A	n=503 (100%)
zone_of_origin_estimated	N/A	n=503 (100%)
mutational_signatures	N/A	n=503 (100%)
neoantigen_load	N/A	n=503 (100%)
AR_activity	N/A	n=503 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=503 (100%)
N_substage	N/A	n=503 (100%)
therapy_radiation_initial	N/A	n=503 (100%)
therapy_radiation_salvage	N/A	n=503 (100%)
therapy_surgery_initial	N/A	n=503 (100%)
therapy_hormonal_initial	N/A	n=503 (100%)
other_treatment	N/A	n=503 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.4 friedrich

	Instances	Missingness
overall_survival_status	0: 230 (90%);1: 25 (10%)	-
days_to_overall_survival	[641,3004.75,3614.5,4301,6771], NA n=91	n=91 (36%)
age_at_initial_diagnosis	N/A	n=255 (100%)
year_diagnosis	N/A	n=255 (100%)
gleason_grade	[5,7,8,8,10], NA n=39	n=39 (15%)
gleason_major	N/A	n=255 (100%)
gleason_minor	N/A	n=255 (100%)
source_of_gleason	N/A	n=255 (100%)
grade_group	N/A	n=255 (100%)
T_pathological	N/A	n=255 (100%)
T_substage_pathological	N/A	n=255 (100%)
T_clinical	N/A	n=255 (100%)
T_substage_clinical	N/A	n=255 (100%)
ERG_fusion_CNA	N/A	n=255 (100%)
ERG_fusion_IHC	N/A	n=255 (100%)
ERG_fusion_GEX	N/A	n=255 (100%)
disease_specific_recurrence_status	N/A	n=255 (100%)
days_to_disease_specific_recurrence	N/A	n=255 (100%)

metastasis_occurrence_status	N/A	n=255 (100%)
days_to_metastatic_occurrence	N/A	n=255 (100%)
psa	N/A	n=255 (100%)
race	caucasian: 255 (100%)	-
smoking_status	N/A	n=255 (100%)
extraprostatic_extension	N/A	n=255 (100%)
perineural_invasion	N/A	n=255 (100%)
seminal_vesicle_invasion	N/A	n=255 (100%)
angiolymphatic_invasion	N/A	n=255 (100%)
androgen_ablation	N/A	n=255 (100%)
capsule	N/A	n=255 (100%)
M_stage	N/A	n=255 (100%)
M_substage	N/A	n=255 (100%)
other_patient	N/A	n=255 (100%)
sample_type	primary: 164 (64%);adjacentnormal: 52 (20%);BPH: 39 (15%)	-
genomic_alterations	N/A	n=255 (100%)
tumor_margins_positive	N/A	n=255 (100%)
tissue_source	prostatectomy: 216 (85%);cystoprostatectomy: 28 (11%);TURP: 11 (4%)	-
metastatic_site	Feature not found in MAE	-
microdissected	1: 255 (100%)	-
frozen_ffpe	frozen: 255 (100%)	-
other_feature	N/A	n=255 (100%)
batch	N/A	n=255 (100%)
other_sample	N/A	n=255 (100%)
tumor_purity_pathology	N/A	n=255 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=255 (100%)
zone_of_origin	N/A	n=255 (100%)
zone_of_origin_estimated	N/A	n=255 (100%)
mutational_signatures	N/A	n=255 (100%)
neoantigen_load	N/A	n=255 (100%)
AR_activity	N/A	n=255 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=255 (100%)
N_substage	N/A	n=255 (100%)
therapy_radiation_initial	N/A	n=255 (100%)
therapy_radiation_salvage	N/A	n=255 (100%)
therapy_surgery_initial	N/A	n=255 (100%)
therapy_hormonal_initial	N/A	n=255 (100%)
other_treatment	N/A	n=255 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.5 hieronymus

	Instances	Missingness
overall_survival_status	0: 96 (92%);1: 8 (8%)	-
days_to_overall_survival	[294.83, 1575.33, 2139.07, 2895.21, 3757.6]	-
age_at_initial_diagnosis	[41.07, 53.25, 58.19, 63.06, 75.63]	-
year_diagnosis	N/A	n=104 (100%)
gleason_grade	[6, 7, 7, 7, 9]	-
gleason_major	3: 72 (69%);4: 30 (29%);5: 2 (2%)	-
gleason_minor	4: 60 (58%);3: 39 (38%);5: 5 (5%)	-
source_of_gleason	prostatectomy: 104 (100%)	-
grade_group	3+4: 56 (54%);4+3: 22 (21%);<=6: 16 (15%); ...	n=4 (4%)
T_pathological	2: 54 (52%);3: 49 (47%);4: 1 (1%)	-
T_substage_pathological	b: 54 (52%);a: 44 (42%);c: 4 (4%); ...	n=2 (2%)
T_clinical	1: 61 (59%);2: 41 (39%);3: 2 (2%)	-
T_substage_clinical	c: 67 (64%);a: 18 (17%);b: 18 (17%); ...	n=1 (1%)
ERG_fusion_CNA	N/A	n=104 (100%)
ERG_fusion_IHC	N/A	n=104 (100%)
ERG_fusion_GEX	N/A	n=104 (100%)
disease_specific_recurrence_status	N/A	n=104 (100%)
days_to_disease_specific_recurrence	N/A	n=104 (100%)
metastasis_occurrence_status	0: 101 (97%);1: 3 (3%)	-
days_to_metastatic_occurrence	[121.25, 1455.26, 1989.62, 2752.45, 3694.65]	-
psa	[1,4.35,5.3,8,56.24], NA n=1	n=1 (1%)
race	N/A	n=104 (100%)
smoking_status	N/A	n=104 (100%)
extraprostatic_extension	0: 52 (50%);1: 52 (50%)	-
perineural_invasion	N/A	n=104 (100%)
seminal_vesicle_invasion	0: 90 (87%);1: 14 (13%)	-
angiolymphatic_invasion	N/A	n=104 (100%)
androgen_ablation	N/A	n=104 (100%)
capsule	N/A	n=104 (100%)
M_stage	N/A	n=104 (100%)
M_substage	N/A	n=104 (100%)
other_patient	N/A	n=104 (100%)
sample_type	N/A	n=104 (100%)
genomic_alterations	N/A	n=104 (100%)
tumor_margins_positive	N/A	n=104 (100%)
tissue_source	N/A	n=104 (100%)
metastatic_site	Feature not found in MAE	-
microdissected	N/A	n=104 (100%)
frozen_ffpe	N/A	n=104 (100%)
other_feature	N/A	n=104 (100%)
batch	N/A	n=104 (100%)
other_sample	N/A	n=104 (100%)
tumor_purity_pathology	N/A	n=104 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=104 (100%)
zone_of_origin	N/A	n=104 (100%)
zone_of_origin_estimated	N/A	n=104 (100%)

mutational_signatures	N/A	n=104 (100%)
neoantigen_load	N/A	n=104 (100%)
AR_activity	N/A	n=104 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=104 (100%)
N_substage	N/A	n=104 (100%)
therapy_radiation_initial	N/A	n=104 (100%)
therapy_radiation_salvage	N/A	n=104 (100%)
therapy_surgery_initial	N/A	n=104 (100%)
therapy_hormonal_initial	N/A	n=104 (100%)
other_treatment	N/A	n=104 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.6 icgcca

	Instances	Missingness
overall_survival_status	0: 198 (93%);1: 8 (4%);NA: 7 (3%)	n=7 (3%)
days_to_overall_survival	[1460,2190,2920,3650,4745], NA n=1	n=1 (0%)
age_at_initial_diagnosis	[42, 58, 62, 66, 74]	-
year_diagnosis	N/A	n=213 (100%)
gleason_grade	[6,7,7,7,8], NA n=138	n=138 (65%)
gleason_major	NA: 138 (65%);3: 49 (23%);4: 26 (12%)	n=138 (65%)
gleason_minor	NA: 138 (65%);4: 42 (20%);3: 33 (15%)	n=138 (65%)
source_of_gleason	N/A	n=213 (100%)
grade_group	NA: 138 (65%);3+4: 37 (17%);4+3: 21 (10%); ...	n=138 (65%)
T_pathological	N/A	n=213 (100%)
T_substage_pathological	N/A	n=213 (100%)
T_clinical	NA: 104 (49%);1: 56 (26%);2: 53 (25%)	n=104 (49%)
T_substage_clinical	NA: 104 (49%);c: 55 (26%);a: 30 (14%); ...	n=104 (49%)
ERG_fusion_CNA	N/A	n=213 (100%)
ERG_fusion_IHC	N/A	n=213 (100%)
ERG_fusion_GEX	N/A	n=213 (100%)
disease_specific_recurrence_status	N/A	n=213 (100%)
days_to_disease_specific_recurrence	N/A	n=213 (100%)
metastasis_occurrence_status	N/A	n=213 (100%)
days_to_metastatic_occurrence	N/A	n=213 (100%)
psa	N/A	n=213 (100%)
race	N/A	n=213 (100%)
smoking_status	N/A	n=213 (100%)
extraprostatic_extension	N/A	n=213 (100%)
perineural_invasion	N/A	n=213 (100%)
seminal_vesicle_invasion	N/A	n=213 (100%)
angiolymphatic_invasion	N/A	n=213 (100%)
androgen_ablation	N/A	n=213 (100%)
capsule	N/A	n=213 (100%)

M_stage	N/A	n=213 (100%)
M_substage	N/A	n=213 (100%)
other_patient	N/A	n=213 (100%)
sample_type	N/A	n=213 (100%)
genomic_alterations	N/A	n=213 (100%)
tumor_margins_positive	N/A	n=213 (100%)
tissue_source	N/A	n=213 (100%)
metastatic_site	Feature not found in MAE	-
microdissected	N/A	n=213 (100%)
frozen_ffpe	N/A	n=213 (100%)
other_feature	N/A	n=213 (100%)
batch	N/A	n=213 (100%)
other_sample	N/A	n=213 (100%)
tumor_purity_pathology	N/A	n=213 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=213 (100%)
zone_of_origin	N/A	n=213 (100%)
zone_of_origin_estimated	N/A	n=213 (100%)
mutational_signatures	N/A	n=213 (100%)
neoantigen_load	N/A	n=213 (100%)
AR_activity	N/A	n=213 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=213 (100%)
N_substage	N/A	n=213 (100%)
therapy_radiation_initial	N/A	n=213 (100%)
therapy_radiation_salvage	N/A	n=213 (100%)
therapy_surgery_initial	N/A	n=213 (100%)
therapy_hormonal_initial	N/A	n=213 (100%)
other_treatment	N/A	n=213 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.7 igc

	Instances	Missingness
overall_survival_status	N/A	n=83 (100%)
days_to_overall_survival	N/A	n=83 (100%)
age_at_initial_diagnosis	NA: 83 (100%)	-
year_diagnosis	N/A	n=83 (100%)
gleason_grade	[6,6,7,7,8], NA n=3	n=3 (4%)
gleason_major	N/A	n=83 (100%)
gleason_minor	N/A	n=83 (100%)
source_of_gleason	N/A	n=83 (100%)
grade_group	N/A	n=83 (100%)
T_pathological	[2,2,2,3,4], NA n=26	n=26 (31%)
T_substage_pathological	c: 33 (40%);NA: 29 (35%);a: 11 (13%); ...	n=29 (35%)
T_clinical	NA: 59 (71%);2: 20 (24%);3: 4 (5%)	n=59 (71%)

T_substage_clinical	NA: 59 (71%);c: 15 (18%);a: 6 (7%); ...	n=59 (71%)
ERG_fusion_CNA	N/A	n=83 (100%)
ERG_fusion_IHC	N/A	n=83 (100%)
ERG_fusion_GEX	N/A	n=83 (100%)
disease_specific_recurrence_status	N/A	n=83 (100%)
days_to_disease_specific_recurrence	N/A	n=83 (100%)
metastasis_occurrence_status	N/A	n=83 (100%)
days_to_metastatic_occurrence	N/A	n=83 (100%)
psa	N/A	n=83 (100%)
race	Caucasian: 74 (89%); African-American: 7 (8%); American Indian: 1 (1%); ...	-
smoking_status	1: 42 (51%);0: 41 (49%)	-
extraprostatic_extension	N/A	n=83 (100%)
perineural_invasion	N/A	n=83 (100%)
seminal_vesicle_invasion	N/A	n=83 (100%)
angiolymphatic_invasion	N/A	n=83 (100%)
androgen_ablation	N/A	n=83 (100%)
capsule	N/A	n=83 (100%)
M_stage	N/A	n=83 (100%)
M_substage	N/A	n=83 (100%)
other_patient	N/A	n=83 (100%)
sample_type	N/A	n=83 (100%)
genomic_alterations	N/A	n=83 (100%)
tumor_margins_positive	N/A	n=83 (100%)
tissue_source	N/A	n=83 (100%)
metastatic_site	N/A	n=83 (100%)
microdissected	N/A	n=83 (100%)
frozen_ffpe	N/A	n=83 (100%)
other_feature	N/A	n=83 (100%)
batch	N/A	n=83 (100%)
other_sample	N/A	n=83 (100%)
tumor_purity_pathology	N/A	n=83 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=83 (100%)
zone_of_origin	N/A	n=83 (100%)
zone_of_origin_estimated	N/A	n=83 (100%)
mutational_signatures	N/A	n=83 (100%)
neoantigen_load	N/A	n=83 (100%)
AR_activity	N/A	n=83 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=83 (100%)
N_substage	N/A	n=83 (100%)
therapy_radiation_initial	N/A	n=83 (100%)
therapy_radiation_salvage	N/A	n=83 (100%)
therapy_surgery_initial	N/A	n=83 (100%)
therapy_hormonal_initial	N/A	n=83 (100%)
other_treatment	N/A	n=83 (100%)

psa_category	Elevated: 74 (89%);Normal: 6 (7%);NA: 3 (4%)	n=3 (4%)
genome_altered	Feature not found in MAE	-

2.8 kim

	Instances	Missingness
overall_survival_status	N/A	n=266 (100%)
days_to_overall_survival	N/A	n=266 (100%)
age_at_initial_diagnosis	[43.41, 57.59, 61.53, 66.91, 76]	-
year_diagnosis	N/A	n=266 (100%)
gleason_grade	NA: 266 (100%)	-
gleason_major	3: 264 (99%);NA: 2 (1%)	-
gleason_minor	[2,3,3,3,4], NA n=2	-
source_of_gleason	N/A	n=266 (100%)
grade_group	<=6: 198 (74%);3+4: 65 (24%);NA: 3 (1%)	n=3 (1%)
T_pathological	N/A	n=266 (100%)
T_substage_pathological	N/A	n=266 (100%)
T_clinical	1: 225 (85%);2: 41 (15%)	-
T_substage_clinical	c: 266 (100%)	-
ERG_fusion_CNA	N/A	n=266 (100%)
ERG_fusion_IHC	N/A	n=266 (100%)
ERG_fusion_GEX	N/A	n=266 (100%)
disease_specific_recurrence_status	N/A	n=266 (100%)
days_to_disease_specific_recurrence	N/A	n=266 (100%)
metastasis_occurrence_status	N/A	n=266 (100%)
days_to_metastatic_occurrence	N/A	n=266 (100%)
psa	[0, 4.16, 5.4, 7.19, 20]	-
race	N/A	n=266 (100%)
smoking_status	N/A	n=266 (100%)
extraprostatic_extension	N/A	n=266 (100%)
perineural_invasion	N/A	n=266 (100%)
seminal_vesicle_invasion	N/A	n=266 (100%)
angiolymphatic_invasion	N/A	n=266 (100%)
androgen_ablation	N/A	n=266 (100%)
capsule	N/A	n=266 (100%)
M_stage	N/A	n=266 (100%)
M_substage	N/A	n=266 (100%)
other_patient	N/A	n=266 (100%)
sample_type	N/A	n=266 (100%)
genomic_alterations	N/A	n=266 (100%)
tumor_margins_positive	N/A	n=266 (100%)
tissue_source	prostate cancer biopsy: 266 (100%)	-
metastatic_site	Feature not found in MAE	-
microdissected	N/A	n=266 (100%)
frozen_ffpe	N/A	n=266 (100%)
other_feature	N/A	n=266 (100%)
batch	N/A	n=266 (100%)
other_sample	N/A	n=266 (100%)

tumor_purity_pathology	N/A	n=266 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=266 (100%)
zone_of_origin	N/A	n=266 (100%)
zone_of_origin_estimated	N/A	n=266 (100%)
mutational_signatures	N/A	n=266 (100%)
neoantigen_load	N/A	n=266 (100%)
AR_activity	N/A	n=266 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=266 (100%)
N_substage	N/A	n=266 (100%)
therapy_radiation_initial	N/A	n=266 (100%)
therapy_radiation_salvage	N/A	n=266 (100%)
therapy_surgery_initial	N/A	n=266 (100%)
therapy_hormonal_initial	N/A	n=266 (100%)
other_treatment	N/A	n=266 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.9 kunderfranco

	Instances	Missingness
overall_survival_status	N/A	n=67 (100%)
days_to_overall_survival	N/A	n=67 (100%)
age_at_initial_diagnosis	[50,62,67,70,74], NA n=14	n=14 (21%)
year_diagnosis	N/A	n=67 (100%)
gleason_grade	[5,7,7,7,9], NA n=14	n=14 (21%)
gleason_major	[2,3,3,4,5], NA n=14	n=14 (21%)
gleason_minor	[3,4,4,4,5], NA n=14	n=14 (21%)
source_of_gleason	prostatectomy: 53 (79%);biopsy: 14 (21%)	-
grade_group	3+4: 29 (43%);NA: 14 (21%);>=8: 11 (16%);	-
	...	
T_pathological	N/A	n=67 (100%)
T_substage_pathological	N/A	n=67 (100%)
T_clinical	N/A	n=67 (100%)
T_substage_clinical	N/A	n=67 (100%)
ERG_fusion_CNA	N/A	n=67 (100%)
ERG_fusion_IHC	N/A	n=67 (100%)
ERG_fusion_GEX	N/A	n=67 (100%)
disease_specific_recurrence_status	N/A	n=67 (100%)
days_to_disease_specific_recurrence	N/A	n=67 (100%)
metastasis_occurrence_status	N/A	n=67 (100%)
days_to_metastatic_occurrence	N/A	n=67 (100%)
psa	N/A	n=67 (100%)
race	N/A	n=67 (100%)
smoking_status	N/A	n=67 (100%)
extraprostatic_extension	N/A	n=67 (100%)

perineural_invasion	N/A	n=67 (100%)
seminal_vesicle_invasion	N/A	n=67 (100%)
angiolymphatic_invasion	N/A	n=67 (100%)
androgen_ablation	N/A	n=67 (100%)
capsule	N/A	n=67 (100%)
M_stage	N/A	n=67 (100%)
M_substage	N/A	n=67 (100%)
other_patient	N/A	n=67 (100%)
sample_type	primary: 53 (79%);BPH: 14 (21%)	-
genomic_alterations	N/A	n=67 (100%)
tumor_margins_positive	N/A	n=67 (100%)
tissue_source	prostatectomy: 53 (79%);biopsy: 14 (21%)	-
metastatic_site	N/A	n=67 (100%)
microdissected	0: 67 (100%)	-
frozen_ffpe	FFPE: 67 (100%)	-
other_feature	N/A	n=67 (100%)
batch	N/A	n=67 (100%)
other_sample	N/A	n=67 (100%)
tumor_purity_pathology	N/A	n=67 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=67 (100%)
zone_of_origin	N/A	n=67 (100%)
zone_of_origin_estimated	N/A	n=67 (100%)
mutational_signatures	N/A	n=67 (100%)
neoantigen_load	N/A	n=67 (100%)
AR_activity	N/A	n=67 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=67 (100%)
N_substage	N/A	n=67 (100%)
therapy_radiation_initial	N/A	n=67 (100%)
therapy_radiation_salvage	N/A	n=67 (100%)
therapy_surgery_initial	N/A	n=67 (100%)
therapy_hormonal_initial	N/A	n=67 (100%)
other_treatment	N/A	n=67 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.10 ren

	Instances	Missingness
overall_survival_status	N/A	n=65 (100%)
days_to_overall_survival	N/A	n=65 (100%)
age_at_initial_diagnosis	[49,64.25,69,74,80], NA n=7	n=7 (11%)
year_diagnosis	N/A	n=65 (100%)
gleason_grade	NA: 65 (100%)	-
gleason_major	3: 29 (45%);4: 27 (42%);5: 9 (14%)	-
gleason_minor	4: 39 (60%);3: 18 (28%);5: 8 (12%)	-

source_of_gleason	N/A	n=65 (100%)
grade_group	3+4: 23 (35%);>=8: 14 (22%);4+3: 13 (20%); ...	n=10 (15%)
T_pathological	N/A	n=65 (100%)
T_substage_pathological	N/A	n=65 (100%)
T_clinical	[2,2,2,3,4], NA n=1	n=1 (2%)
T_substage_clinical	c: 38 (58%);b: 11 (17%);a: 10 (15%); ...	n=6 (9%)
ERG_fusion_CNA	N/A	n=65 (100%)
ERG_fusion_IHC	N/A	n=65 (100%)
ERG_fusion_GEX	N/A	n=65 (100%)
disease_specific_recurrence_status	N/A	n=65 (100%)
days_to_disease_specific_recurrence	N/A	n=65 (100%)
metastasis_occurrence_status	N/A	n=65 (100%)
days_to_metastatic_occurrence	N/A	n=65 (100%)
psa	[5.28,10.73,17.3,33.4,271], NA n=2	n=2 (3%)
race	N/A	n=65 (100%)
smoking_status	N/A	n=65 (100%)
extraprostatic_extension	N/A	n=65 (100%)
perineural_invasion	N/A	n=65 (100%)
seminal_vesicle_invasion	N/A	n=65 (100%)
angiolymphatic_invasion	N/A	n=65 (100%)
androgen_ablation	N/A	n=65 (100%)
capsule	N/A	n=65 (100%)
M_stage	N/A	n=65 (100%)
M_substage	N/A	n=65 (100%)
other_patient	N/A	n=65 (100%)
sample_type	N/A	n=65 (100%)
genomic_alterations	N/A	n=65 (100%)
tumor_margins_positive	N/A	n=65 (100%)
tissue_source	N/A	n=65 (100%)
metastatic_site	Feature not found in MAE	-
microdissected	N/A	n=65 (100%)
frozen_ffpe	N/A	n=65 (100%)
other_feature	N/A	n=65 (100%)
batch	N/A	n=65 (100%)
other_sample	N/A	n=65 (100%)
tumor_purity_pathology	N/A	n=65 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=65 (100%)
zone_of_origin	N/A	n=65 (100%)
zone_of_origin_estimated	N/A	n=65 (100%)
mutational_signatures	N/A	n=65 (100%)
neoantigen_load	N/A	n=65 (100%)
AR_activity	N/A	n=65 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=65 (100%)
N_substage	N/A	n=65 (100%)
therapy_radiation_initial	N/A	n=65 (100%)

therapy_radiation_salvage	N/A	n=65 (100%)
therapy_surgery_initial	N/A	n=65 (100%)
therapy_hormonal_initial	N/A	n=65 (100%)
other_treatment	N/A	n=65 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.11 sun

	Instances	Missingness
overall_survival_status	N/A	n=79 (100%)
days_to_overall_survival	N/A	n=79 (100%)
age_at_initial_diagnosis	N/A	n=79 (100%)
year_diagnosis	N/A	n=79 (100%)
gleason_grade	N/A	n=79 (100%)
gleason_major	N/A	n=79 (100%)
gleason_minor	N/A	n=79 (100%)
source_of_gleason	N/A	n=79 (100%)
grade_group	N/A	n=79 (100%)
T_pathological	N/A	n=79 (100%)
T_substage_pathological	N/A	n=79 (100%)
T_clinical	N/A	n=79 (100%)
T_substage_clinical	N/A	n=79 (100%)
ERG_fusion_CNA	N/A	n=79 (100%)
ERG_fusion_IHC	N/A	n=79 (100%)
ERG_fusion_GEX	N/A	n=79 (100%)
disease_specific_recurrence_status	0: 40 (51%);1: 39 (49%)	-
days_to_disease_specific_recurrence	N/A	n=79 (100%)
metastasis_occurrence_status	N/A	n=79 (100%)
days_to_metastatic_occurrence	N/A	n=79 (100%)
psa	N/A	n=79 (100%)
race	N/A	n=79 (100%)
smoking_status	N/A	n=79 (100%)
extraprostatic_extension	N/A	n=79 (100%)
perineural_invasion	N/A	n=79 (100%)
seminal_vesicle_invasion	N/A	n=79 (100%)
angiolymphatic_invasion	N/A	n=79 (100%)
androgen_ablation	N/A	n=79 (100%)
capsule	N/A	n=79 (100%)
M_stage	N/A	n=79 (100%)
M_substage	N/A	n=79 (100%)
other_patient	N/A	n=79 (100%)
sample_type	primary: 79 (100%)	-
genomic_alterations	N/A	n=79 (100%)
tumor_margins_positive	N/A	n=79 (100%)
tissue_source	N/A	n=79 (100%)
metastatic_site	Feature not found in MAE	-
microdissected	N/A	n=79 (100%)
frozen_ffpe	N/A	n=79 (100%)

other_feature	N/A	n=79 (100%)
batch	N/A	n=79 (100%)
other_sample	N/A	n=79 (100%)
tumor_purity_pathology	N/A	n=79 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=79 (100%)
zone_of_origin	N/A	n=79 (100%)
zone_of_origin_estimated	N/A	n=79 (100%)
mutational_signatures	N/A	n=79 (100%)
neoantigen_load	N/A	n=79 (100%)
AR_activity	N/A	n=79 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=79 (100%)
N_substage	N/A	n=79 (100%)
therapy_radiation_initial	N/A	n=79 (100%)
therapy_radiation_salvage	N/A	n=79 (100%)
therapy_surgery_initial	N/A	n=79 (100%)
therapy_hormonal_initial	N/A	n=79 (100%)
other_treatment	N/A	n=79 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.12 taylor

	Instances	Missingness
overall_survival_status	N/A	n=218 (100%)
days_to_overall_survival	N/A	n=218 (100%)
age_at_initial_diagnosis	N/A	n=218 (100%)
year_diagnosis	N/A	n=218 (100%)
gleason_grade	[5,6,7,7,9], NA n=1	n=1 (0%)
gleason_major	[3,3,3,4,5], NA n=21	n=21 (10%)
gleason_minor	[3,3,4,4,5], NA n=21	n=21 (10%)
source_of_gleason	N/A	n=218 (100%)
grade_group	<=6: 104 (48%);>=8: 34 (16%);3+4: 34 (16%); ...	n=31 (14%)
T_pathological	[2,2,2,3,4], NA n=18	n=18 (8%)
T_substage_pathological	b: 86 (39%);a: 55 (25%);c: 39 (18%); ...	n=38 (17%)
T_clinical	[1,1,2,2,4], NA n=9	n=9 (4%)
T_substage_clinical	c: 127 (58%);a: 40 (18%);b: 34 (16%); ...	n=17 (8%)
ERG_fusion_CNA	0: 164 (75%);1: 54 (25%)	-
ERG_fusion_IHC	N/A	n=218 (100%)
ERG_fusion_GEX	0: 76 (35%);1: 74 (34%);NA: 68 (31%)	n=68 (31%)
disease_specific_recurrence_status	0: 137 (63%);1: 61 (28%);NA: 20 (9%)	n=20 (9%)
days_to_disease_specific_recurrence	[3,717.25,1386.5,1974.25,4909], NA n=20	n=20 (9%)
metastasis_occurrence_status	N/A	n=218 (100%)
days_to_metastatic_occurrence	N/A	n=218 (100%)
psa	N/A	n=218 (100%)

race	N/A	n=218 (100%)
smoking_status	N/A	n=218 (100%)
extraprostatic_extension	N/A	n=218 (100%)
perineural_invasion	N/A	n=218 (100%)
seminal_vesicle_invasion	N/A	n=218 (100%)
angiolymphatic_invasion	N/A	n=218 (100%)
androgen_ablation	N/A	n=218 (100%)
capsule	N/A	n=218 (100%)
M_stage	N/A	n=218 (100%)
M_substage	N/A	n=218 (100%)
other_patient	N/A	n=218 (100%)
sample_type	primary: 181 (83%);metastasis: 37 (17%)	-
genomic_alterations	N/A	n=218 (100%)
tumor_margins_positive	N/A	n=218 (100%)
tissue_source	N/A	n=218 (100%)
metastatic_site	N/A	n=218 (100%)
microdissected	N/A	n=218 (100%)
frozen_ffpe	N/A	n=218 (100%)
other_feature	N/A	n=218 (100%)
batch	N/A	n=218 (100%)
other_sample	N/A	n=218 (100%)
tumor_purity_pathology	N/A	n=218 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=218 (100%)
zone_of_origin	N/A	n=218 (100%)
zone_of_origin_estimated	N/A	n=218 (100%)
mutational_signatures	N/A	n=218 (100%)
neoantigen_load	N/A	n=218 (100%)
AR_activity	N/A	n=218 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=218 (100%)
N_substage	N/A	n=218 (100%)
therapy_radiation_initial	N/A	n=218 (100%)
therapy_radiation_salvage	N/A	n=218 (100%)
therapy_surgery_initial	N/A	n=218 (100%)
therapy_hormonal_initial	N/A	n=218 (100%)
other_treatment	N/A	n=218 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.13 tcga

	Instances	Missingness
overall_survival_status	0: 484 (98%);1: 10 (2%)	-
days_to_overall_survival	[23.18, 524.9, 925.68, 1466.06, 5034.03]	-
age_at_initial_diagnosis	[41, 56, 61, 66, 78]	-
year_diagnosis	[2000,2009,2011,2012,2013], NA n=31	n=31 (6%)

gleason_grade	[6, 7, 7, 9, 10]	-
gleason_major	[2, 3, 4, 4, 5]	-
gleason_minor	4: 234 (47%);3: 150 (30%);5: 110 (22%)	-
source_of_gleason	biopsy: 494 (100%)	-
grade_group	>=8: 205 (41%);3+4: 145 (29%);4+3: 99 (20%); ...	-
T_pathological	[2,2,3,3,4], NA n=7	n=7 (1%)
T_substage_pathological	a: 171 (35%);c: 164 (33%);b: 142 (29%); ...	n=17 (3%)
T_clinical	[1,1,2,2,4], NA n=91	n=91 (18%)
T_substage_clinical	c: 224 (45%);NA: 106 (21%);a: 93 (19%); ...	n=106 (21%)
ERG_fusion_CNA	N/A	n=494 (100%)
ERG_fusion_IHC	N/A	n=494 (100%)
ERG_fusion_GEX	N/A	n=494 (100%)
disease_specific_recurrence_status	0: 397 (80%);1: 91 (18%);NA: 6 (1%)	n=6 (1%)
days_to_disease_specific_recurrence	[23.18,427.38,823.04,1375.86,5034.03], NA n=6	n=6 (1%)
metastasis_occurrence_status	N/A	n=494 (100%)
days_to_metastatic_occurrence	N/A	n=494 (100%)
psa	[0,0.03,0.1,0.11,323], NA n=57	n=57 (12%)
race	NA: 338 (68%);caucasian: 147 (30%);african_american: 7 (1%); ...	-
smoking_status	N/A	n=494 (100%)
extraprostatic_extension	N/A	n=494 (100%)
perineural_invasion	N/A	n=494 (100%)
seminal_vesicle_invasion	N/A	n=494 (100%)
angiolymphatic_invasion	N/A	n=494 (100%)
androgen_ablation	N/A	n=494 (100%)
capsule	N/A	n=494 (100%)
M_stage	0: 452 (91%);NA: 39 (8%);1: 3 (1%)	n=39 (8%)
M_substage	: 491 (99%);a: 1 (0%);b: 1 (0%); ...	-
other_patient	N/A	n=494 (100%)
sample_type	Primary: 493 (100%);Metastasis: 1 (0%)	-
genomic_alterations	N/A	n=494 (100%)
tumor_margins_positive	N/A	n=494 (100%)
tissue_source	N/A	n=494 (100%)
metastatic_site	N/A	n=494 (100%)
microdissected	N/A	n=494 (100%)
frozen_ffpe	NA: 493 (100%);: 1 (0%)	-
other_feature	N/A	n=494 (100%)
batch	N/A	n=494 (100%)
other_sample	N/A	n=494 (100%)
tumor_purity_pathology	N/A	n=494 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=494 (100%)
zone_of_origin	NA: 218 (44%);peripheral: 137 (28%);mixed: 127 (26%); ...	-
zone_of_origin_estimated	N/A	n=494 (100%)
mutational_signatures	N/A	n=494 (100%)
neoantigen_load	N/A	n=494 (100%)
AR_activity	N/A	n=494 (100%)

prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	0: 343 (69%);1: 79 (16%);NA: 72 (15%)	n=72 (15%)
N_substage	N/A	n=494 (100%)
therapy_radiation_initial	0: 246 (50%);NA: 208 (42%);1: 40 (8%)	n=208 (42%)
therapy_radiation_salvage	N/A	n=494 (100%)
therapy_surgery_initial	N/A	n=494 (100%)
therapy_hormonal_initial	N/A	n=494 (100%)
other_treatment	0: 244 (49%);NA: 207 (42%);1: 43 (9%)	n=207 (42%)
psa_category	N/A	n=494 (100%)
genome_altered	[0,0.02,0.06,0.12,0.58], NA n=6	n=6 (1%)

2.14 true

	Instances	Missingness
overall_survival_status	N/A	n=29 (100%)
days_to_overall_survival	N/A	n=29 (100%)
age_at_initial_diagnosis	N/A	n=29 (100%)
year_diagnosis	N/A	n=29 (100%)
gleason_grade	[6, 7, 7, 7, 9]	-
gleason_major	3: 15 (52%);4: 12 (41%);5: 2 (7%)	-
gleason_minor	4: 14 (48%);3: 13 (45%);5: 2 (7%)	-
source_of_gleason	N/A	n=29 (100%)
grade_group	4: 11 (38%);3: 10 (34%);5: 6 (21%); ...	n=2 (7%)
T_pathological	N/A	n=29 (100%)
T_substage_pathological	N/A	n=29 (100%)
T_clinical	N/A	n=29 (100%)
T_substage_clinical	N/A	n=29 (100%)
ERG_fusion_CNA	N/A	n=29 (100%)
ERG_fusion_IHC	N/A	n=29 (100%)
ERG_fusion_GEX	N/A	n=29 (100%)
disease_specific_recurrence_status	N/A	n=29 (100%)
days_to_disease_specific_recurrence	N/A	n=29 (100%)
metastasis_occurrence_status	N/A	n=29 (100%)
days_to_metastatic_occurrence	N/A	n=29 (100%)
psa	[1,5,6.35,8.62,15.9], NA n=1	n=1 (3%)
race	N/A	n=29 (100%)
smoking_status	N/A	n=29 (100%)
extraprostatic_extension	N/A	n=29 (100%)
perineural_invasion	N/A	n=29 (100%)
seminal_vesicle_invasion	N/A	n=29 (100%)
angiolymphatic_invasion	N/A	n=29 (100%)
androgen_ablation	N/A	n=29 (100%)
capsule	N/A	n=29 (100%)
M_stage	N/A	n=29 (100%)
M_substage	N/A	n=29 (100%)
other_patient	N/A	n=29 (100%)
sample_type	N/A	n=29 (100%)

genomic_alterations	N/A	n=29 (100%)
tumor_margins_positive	0: 18 (62%);1: 11 (38%)	-
tissue_source	N/A	n=29 (100%)
metastatic_site	N/A	n=29 (100%)
microdissected	1: 29 (100%)	-
frozen_ffpe	N/A	n=29 (100%)
other_feature	Age:60-69 Volume:8 LCM_Gleason_Pattern:5: 2 (7%);Age:40-49 Volume:2.2 LCM_Gleason_Pattern:5: 1 (3%);Age:40-49 Volume:3.5 LCM_Gleason_Pattern:3: 1 (3%); ...	-
batch	N/A	n=29 (100%)
other_sample	N/A	n=29 (100%)
tumor_purity_pathology	N/A	n=29 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=29 (100%)
zone_of_origin	N/A	n=29 (100%)
zone_of_origin_estimated	N/A	n=29 (100%)
mutational_signatures	N/A	n=29 (100%)
neoantigen_load	N/A	n=29 (100%)
AR_activity	N/A	n=29 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=29 (100%)
N_substage	N/A	n=29 (100%)
therapy_radiation_initial	0: 29 (100%)	-
therapy_radiation_salvage	0: 29 (100%)	-
therapy_surgery_initial	N/A	n=29 (100%)
therapy_hormonal_initial	0: 29 (100%)	-
other_treatment	NA: 21 (72%);fish_oil: 2 (7%);selenium: 2 (7%); ...	n=21 (72%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.15 wallace

	Instances	Missingness
overall_survival_status	N/A	n=83 (100%)
days_to_overall_survival	N/A	n=83 (100%)
age_at_initial_diagnosis	N/A	n=83 (100%)
year_diagnosis	N/A	n=83 (100%)
gleason_grade	[5, 6, 7, 7, 9]	-
gleason_major	N/A	n=83 (100%)
gleason_minor	N/A	n=83 (100%)
source_of_gleason	N/A	n=83 (100%)
grade_group	N/A	n=83 (100%)
T_pathological	N/A	n=83 (100%)

T_substage_pathological	N/A	n=83 (100%)
T_clinical	N/A	n=83 (100%)
T_substage_clinical	N/A	n=83 (100%)
ERG_fusion_CNA	N/A	n=83 (100%)
ERG_fusion_IHC	N/A	n=83 (100%)
ERG_fusion_GEX	N/A	n=83 (100%)
disease_specific_recurrence_status	N/A	n=83 (100%)
days_to_disease_specific_recurrence	N/A	n=83 (100%)
metastasis_occurrence_status	N/A	n=83 (100%)
days_to_metastatic_occurrence	N/A	n=83 (100%)
psa	N/A	n=83 (100%)
race	caucasian: 45 (54%);african_american: 38 (46%)	-
smoking_status	1: 37 (45%);NA: 25 (30%);0: 21 (25%)	n=25 (30%)
extraprostatic_extension	0: 54 (65%);1: 27 (33%);NA: 2 (2%)	n=2 (2%)
perineural_invasion	1: 69 (83%);0: 12 (14%);NA: 2 (2%)	n=2 (2%)
seminal_vesicle_invasion	0: 69 (83%);1: 12 (14%);NA: 2 (2%)	n=2 (2%)
angiolymphatic_invasion	0: 76 (92%);1: 5 (6%);NA: 2 (2%)	n=2 (2%)
androgen_ablation	N/A	n=83 (100%)
capsule	N/A	n=83 (100%)
M_stage	N/A	n=83 (100%)
M_substage	N/A	n=83 (100%)
other_patient	N/A	n=83 (100%)
sample_type	primary: 69 (83%);adjacentnormal: 14 (17%)	-
genomic_alterations	N/A	n=83 (100%)
tumor_margins_positive	0: 51 (61%);1: 29 (35%);NA: 3 (4%)	n=3 (4%)
tissue_source	N/A	n=83 (100%)
metastatic_site	N/A	n=83 (100%)
microdissected	0: 83 (100%)	-
frozen_ffpe	frozen: 83 (100%)	-
other_feature	N/A	n=83 (100%)
batch	N/A	n=83 (100%)
other_sample	N/A	n=83 (100%)
tumor_purity_pathology	N/A	n=83 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=83 (100%)
zone_of_origin	N/A	n=83 (100%)
zone_of_origin_estimated	N/A	n=83 (100%)
mutational_signatures	N/A	n=83 (100%)
neoantigen_load	N/A	n=83 (100%)
AR_activity	N/A	n=83 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=83 (100%)
N_substage	N/A	n=83 (100%)
therapy_radiation_initial	0: 83 (100%)	-
therapy_radiation_salvage	0: 83 (100%)	-
therapy_surgery_initial	0: 83 (100%)	-
therapy_hormonal_initial	0: 83 (100%)	-

other_treatment	N/A	n=83 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

2.16 wang

	Instances	Missingness
overall_survival_status	N/A	n=148 (100%)
days_to_overall_survival	N/A	n=148 (100%)
age_at_initial_diagnosis	N/A	n=148 (100%)
year_diagnosis	N/A	n=148 (100%)
gleason_grade	N/A	n=148 (100%)
gleason_major	N/A	n=148 (100%)
gleason_minor	N/A	n=148 (100%)
source_of_gleason	N/A	n=148 (100%)
grade_group	N/A	n=148 (100%)
T_pathological	N/A	n=148 (100%)
T_substage_pathological	N/A	n=148 (100%)
T_clinical	N/A	n=148 (100%)
T_substage_clinical	N/A	n=148 (100%)
ERG_fusion_CNA	N/A	n=148 (100%)
ERG_fusion_IHC	N/A	n=148 (100%)
ERG_fusion_GEX	N/A	n=148 (100%)
disease_specific_recurrence_status	N/A	n=148 (100%)
days_to_disease_specific_recurrence	N/A	n=148 (100%)
metastasis_occurrence_status	N/A	n=148 (100%)
days_to_metastatic_occurrence	N/A	n=148 (100%)
psa	N/A	n=148 (100%)
race	N/A	n=148 (100%)
smoking_status	N/A	n=148 (100%)
extraprostatic_extension	N/A	n=148 (100%)
perineural_invasion	N/A	n=148 (100%)
seminal_vesicle_invasion	N/A	n=148 (100%)
angiolympathic_invasion	N/A	n=148 (100%)
androgen_ablation	N/A	n=148 (100%)
capsule	N/A	n=148 (100%)
M_stage	N/A	n=148 (100%)
M_substage	N/A	n=148 (100%)
other_patient	N/A	n=148 (100%)
sample_type	N/A	n=148 (100%)
genomic_alterations	N/A	n=148 (100%)
tumor_margins_positive	N/A	n=148 (100%)
tissue_source	N/A	n=148 (100%)
metastatic_site	N/A	n=148 (100%)
microdissected	N/A	n=148 (100%)
frozen_ffpe	N/A	n=148 (100%)
other_feature	N/A	n=148 (100%)
batch	N/A	n=148 (100%)
other_sample	N/A	n=148 (100%)

tumor_purity_pathology	N/A	n=148 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=148 (100%)
zone_of_origin	N/A	n=148 (100%)
zone_of_origin_estimated	N/A	n=148 (100%)
mutational_signatures	N/A	n=148 (100%)
neoantigen_load	N/A	n=148 (100%)
AR_activity	N/A	n=148 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=148 (100%)
N_substage	N/A	n=148 (100%)
therapy_radiation_initial	N/A	n=148 (100%)
therapy_radiation_salvage	N/A	n=148 (100%)
therapy_surgery_initial	N/A	n=148 (100%)
therapy_hormonal_initial	N/A	n=148 (100%)
other_treatment	N/A	n=148 (100%)
psa_category	N/A	n=148 (100%)
genome_altered	Feature not found in MAE	-

2.17 weiner

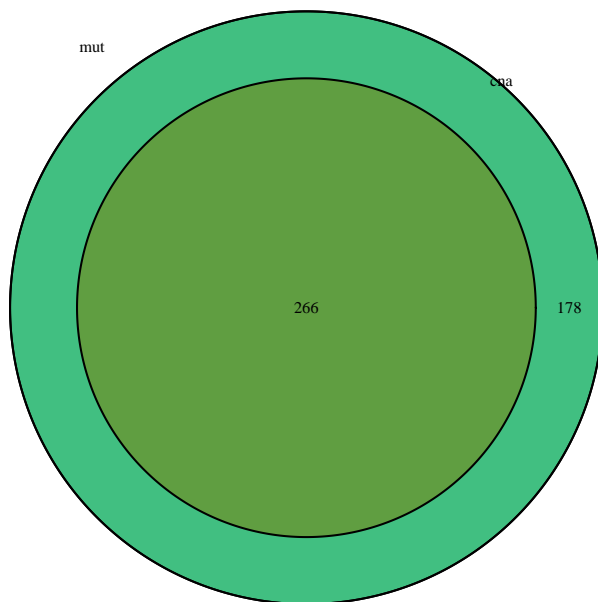
	Instances	Missingness
overall_survival_status	N/A	n=838 (100%)
days_to_overall_survival	N/A	n=838 (100%)
age_at_initial_diagnosis	[39, 57, 61, 65, 76]	-
year_diagnosis	N/A	n=838 (100%)
gleason_grade	N/A	n=838 (100%)
gleason_major	N/A	n=838 (100%)
gleason_minor	N/A	n=838 (100%)
source_of_gleason	N/A	n=838 (100%)
grade_group	3+4: 419 (50%);4+3: 183 (22%);>=8: 171 (20%); ...	-
T_pathological	N/A	n=838 (100%)
T_substage_pathological	N/A	n=838 (100%)
T_clinical	N/A	n=838 (100%)
T_substage_clinical	N/A	n=838 (100%)
ERG_fusion_CNA	N/A	n=838 (100%)
ERG_fusion_IHC	N/A	n=838 (100%)
ERG_fusion_GEX	N/A	n=838 (100%)
disease_specific_recurrence_status	N/A	n=838 (100%)
days_to_disease_specific_recurrence	N/A	n=838 (100%)
metastasis_occurrence_status	N/A	n=838 (100%)
days_to_metastatic_occurrence	N/A	n=838 (100%)
psa	[0,5.2,7.4,11.42,83.44], NA n=18	n=18 (2%)
race	NA: 538 (64%);african_american: 150 (18%);caucasian: 150 (18%)	n=538 (64%)
smoking_status	N/A	n=838 (100%)

extraprostatic_extension	N/A	n=838 (100%)
perineural_invasion	N/A	n=838 (100%)
seminal_vesicle_invasion	N/A	n=838 (100%)
angiolympathic_invasion	N/A	n=838 (100%)
androgen_ablation	N/A	n=838 (100%)
capsule	N/A	n=838 (100%)
M_stage	N/A	n=838 (100%)
M_substage	N/A	n=838 (100%)
other_patient	N/A	n=838 (100%)
sample_type	primary: 838 (100%)	-
genomic_alterations	N/A	n=838 (100%)
tumor_margins_positive	N/A	n=838 (100%)
tissue_source	prostatectomy: 838 (100%)	-
metastatic_site	N/A	n=838 (100%)
microdissected	0: 838 (100%)	-
frozen_ffpe	FFPE: 838 (100%)	-
other_feature	N/A	n=838 (100%)
batch	Durham Veterans Affairs Hospital: 538 (64%); Johns Hopkins Medical Institute: 300 (36%)	-
other_sample	N/A	n=838 (100%)
tumor_purity_pathology	N/A	n=838 (100%)
tumor_purity_demixt	Feature not found in MAE	-
tumor_purity_absolute	N/A	n=838 (100%)
zone_of_origin	N/A	n=838 (100%)
zone_of_origin_estimated	N/A	n=838 (100%)
mutational_signatures	N/A	n=838 (100%)
neoantigen_load	N/A	n=838 (100%)
AR_activity	N/A	n=838 (100%)
prolaris	Feature not found in MAE	-
decipher	Feature not found in MAE	-
oncotypedx	Feature not found in MAE	-
N_stage	N/A	n=838 (100%)
N_substage	N/A	n=838 (100%)
therapy_radiation_initial	0: 838 (100%)	-
therapy_radiation_salvage	0: 838 (100%)	-
therapy_surgery_initial	0: 838 (100%)	-
therapy_hormonal_initial	0: 838 (100%)	-
other_treatment	N/A	n=838 (100%)
psa_category	Feature not found in MAE	-
genome_altered	Feature not found in MAE	-

3 Omics sample overlap within MAE-objects

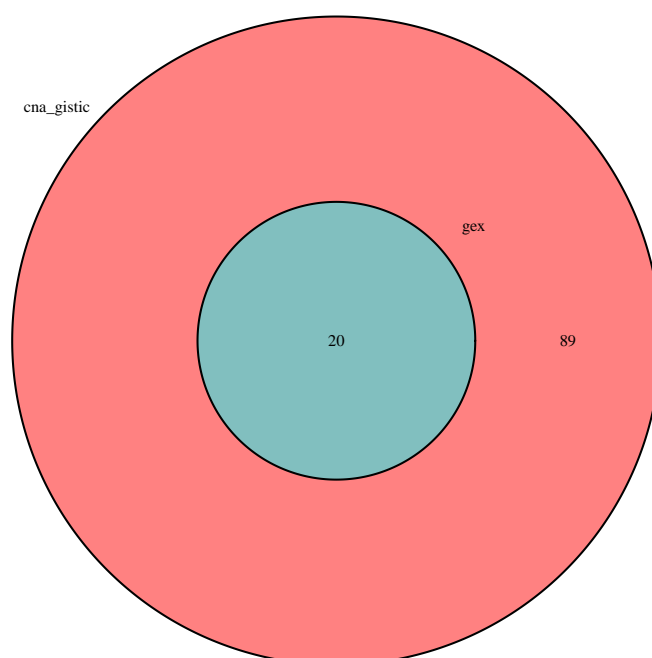
```
null device
1
```

3.1 abida

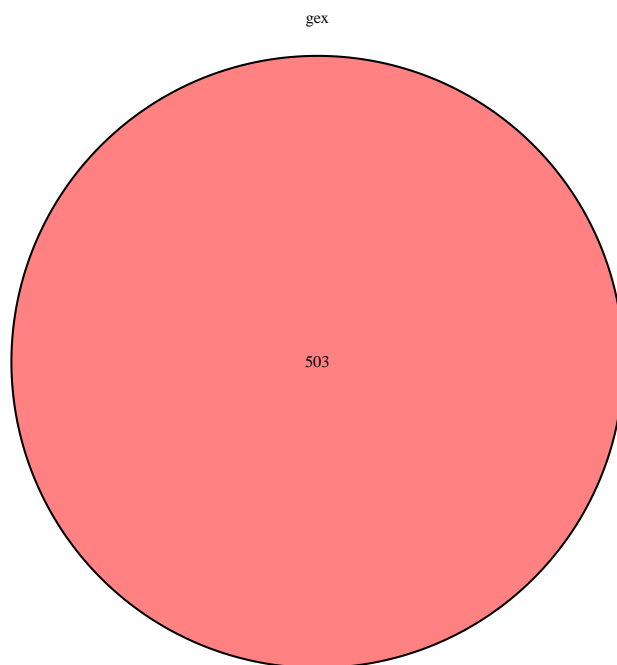


ex polvA

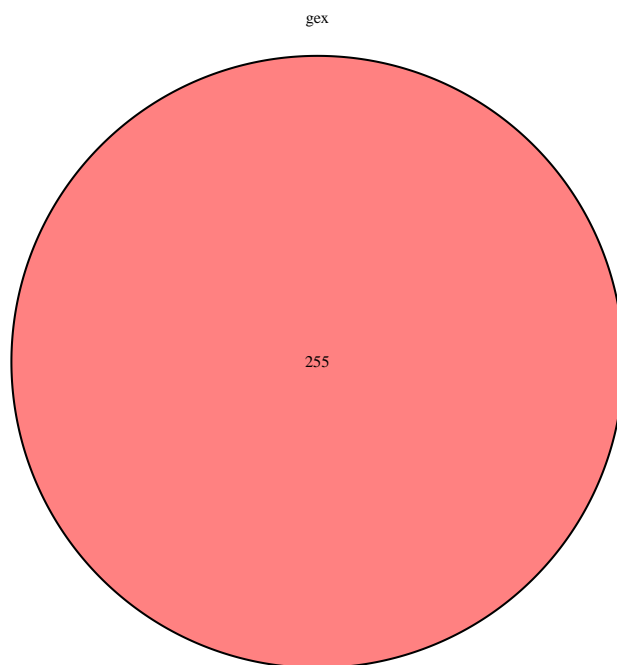
3.2 barbieri



3.3 chandran

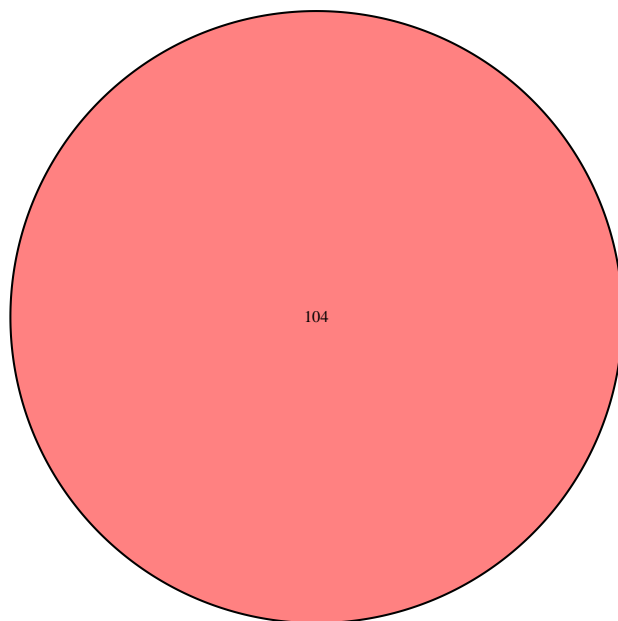


3.4 friedrich



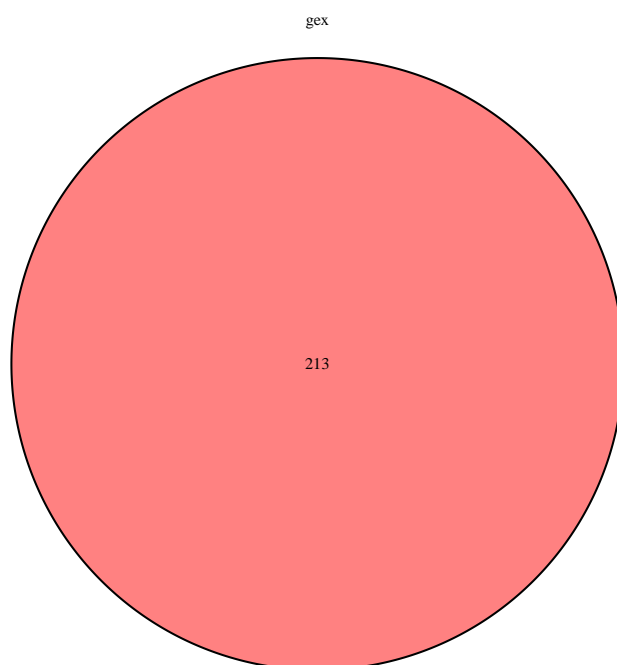
3.5 hieronymus

cna



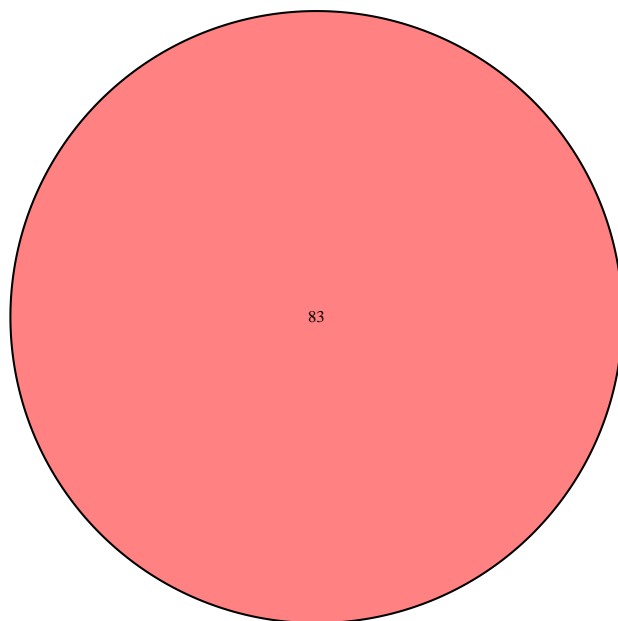
104

3.6 icgcca



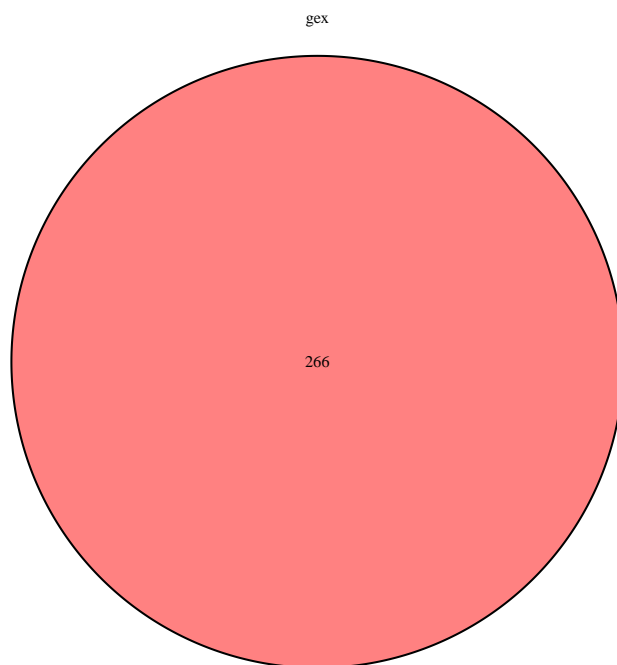
3.7 igc

gex



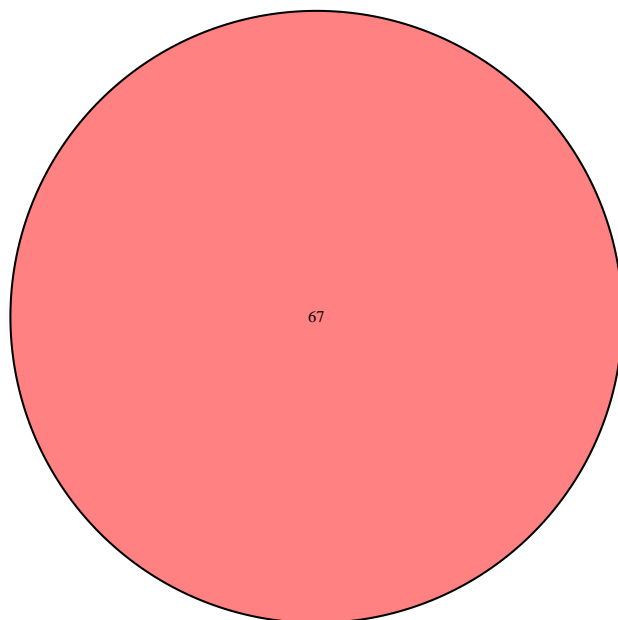
83

3.8 kim



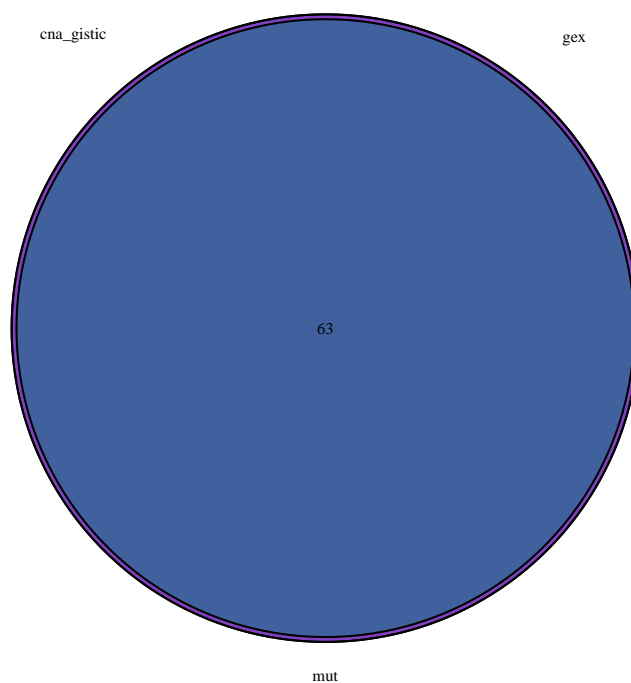
3.9 kunderfranco

gex

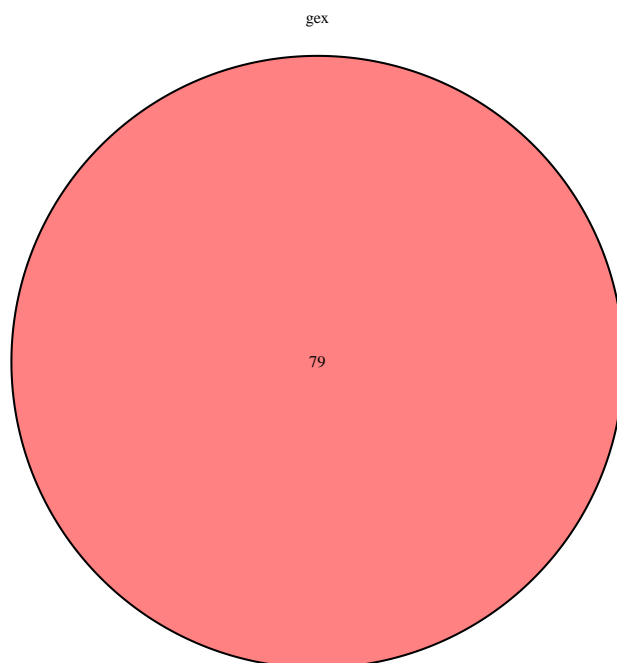


67

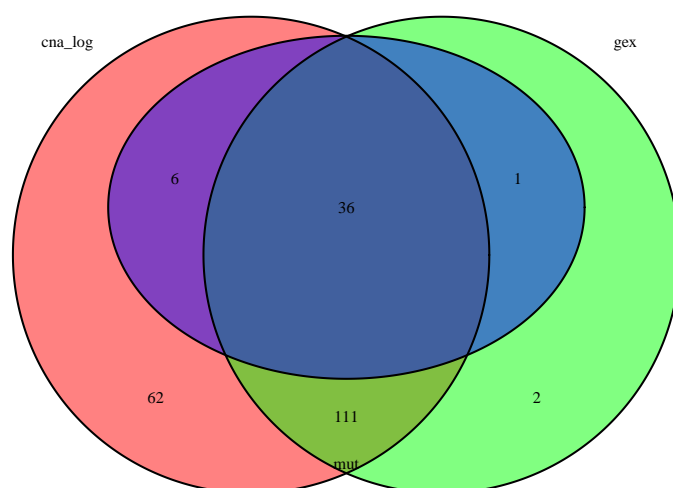
3.10 ren



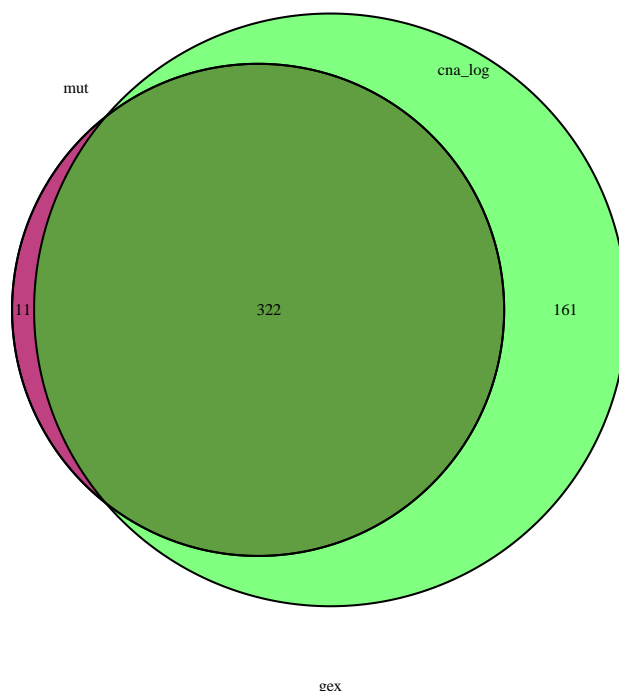
3.11 sun



3.12 taylor

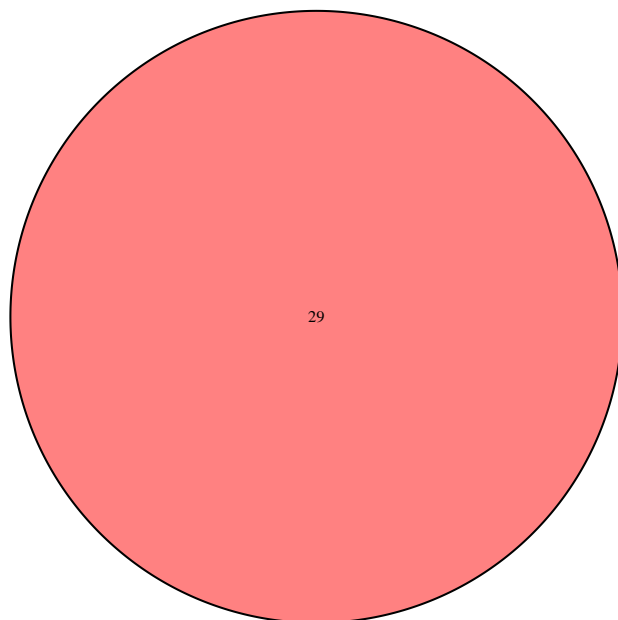


3.13 tcga



3.14 true

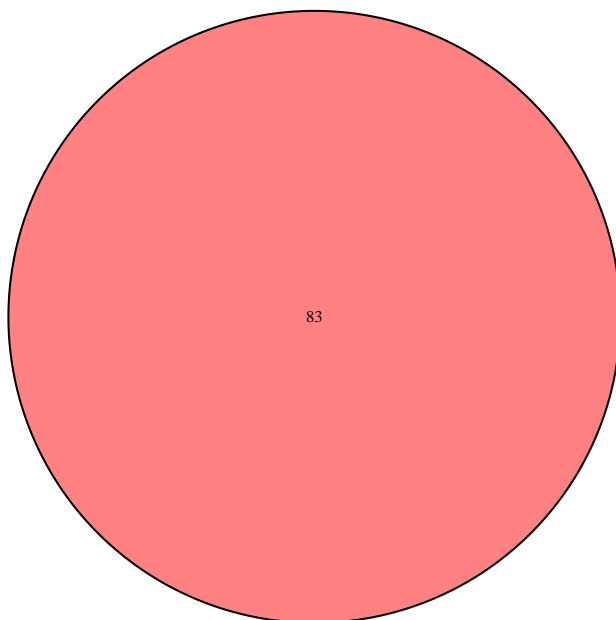
gex



29

3.15 wallace

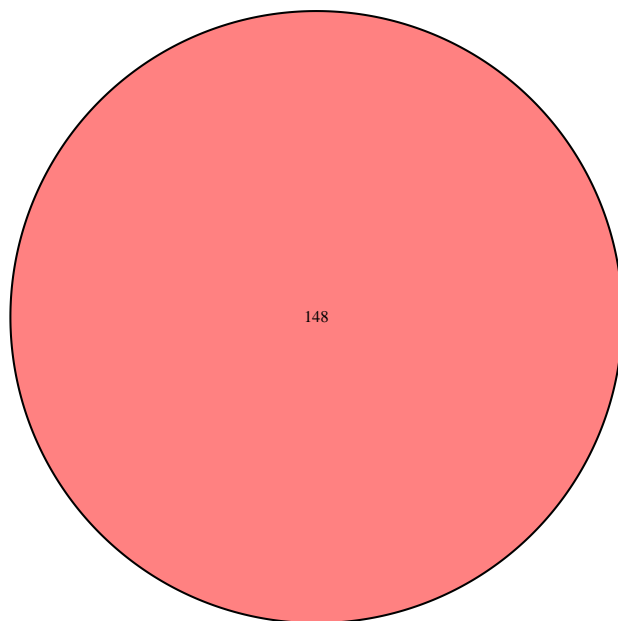
gex



83

3.16 wang

gex



148

3.17 weiner

