



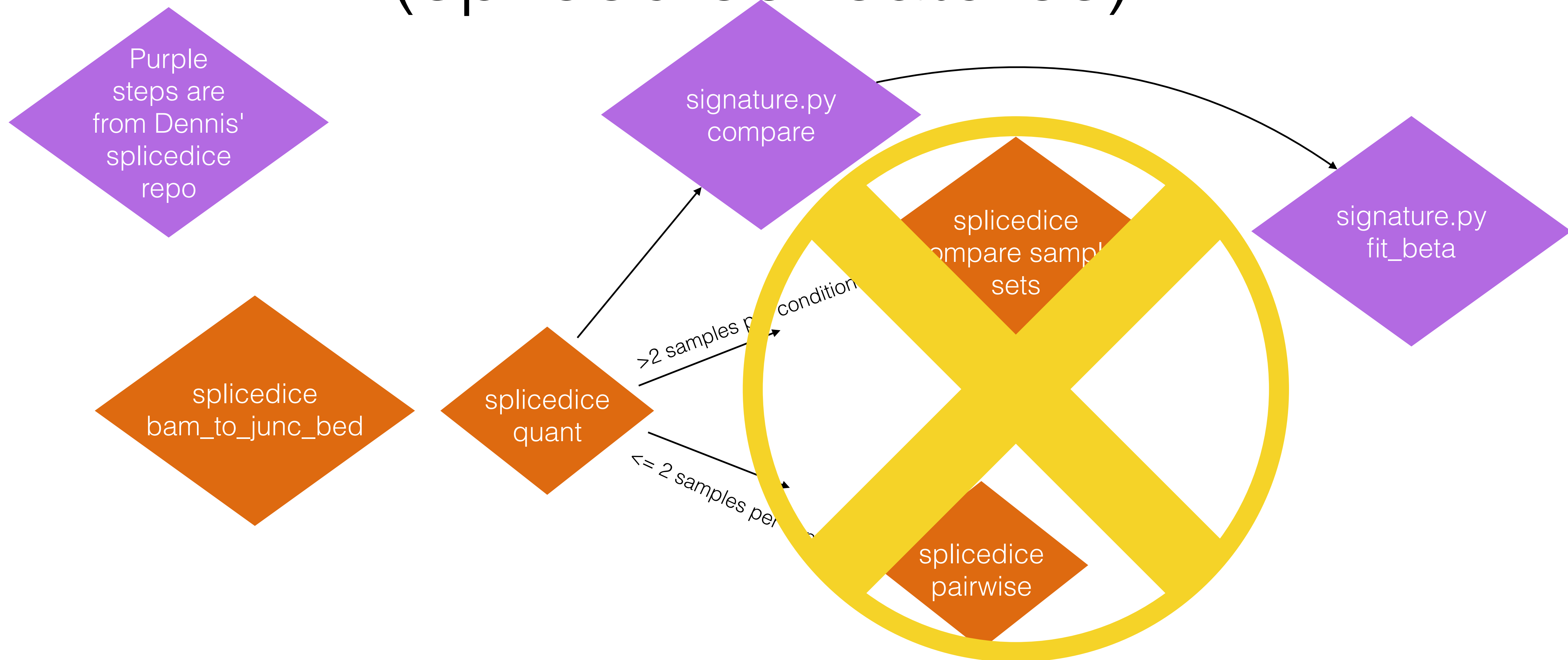
splicedice analysis of 100 TCGA-LUAD samples by U2AF1 S34F mutation status

Holly Beale
5/30/25

TCGA LUAD bam files from GDC

- attempted to download 100 (12 were u2af1-s34f per dennis's manifest)
- 93 (11) downloaded on first try; retrying the remaining 1 u2af1-s34f failed
- one u2af1-s34f was preserved in OCT (remaining "unknown", presumably frozen), so I included 3 u2af1-wt preserved in OCT
- All were 1) tumor (not normal), 2) primary (not a recurrence), and 3) aligned with STAR 2-pass Genome.
-

Splicedice workflow overview (splicedice features)



signature.py compare (splicedice)

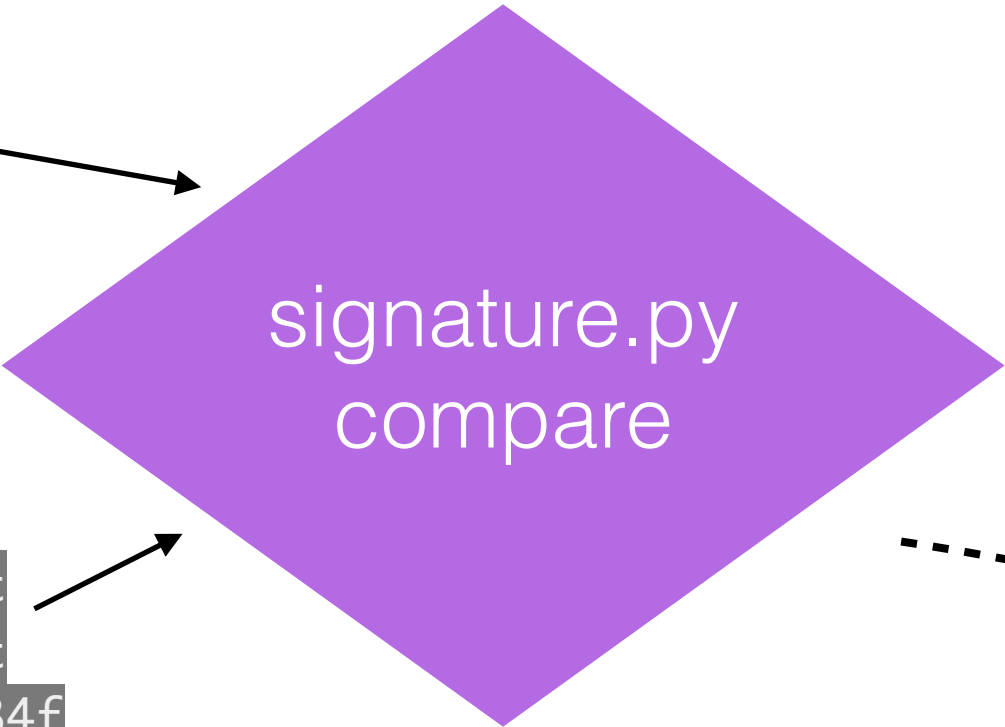
(selected contents, not in order)

allPS.tsv

cluster	TCGA-55-A4...	TCGA-78-76...
chr1:11211-12009:+	nan	nan
chr1:13052-13220:+	nan	0.500
chr1:14784-14977:-	0.000	0.000
chr1:14829-15020:+	0.857	1.000
chr1:17055-17605:-	0.060	0.007

sig_manifest.tsv

TCGA-55-A4DF-01A_4a5e9e8a-8c48-48cf-8bf0-eb564611d382	u2af1-wt
TCGA-78-7633-01A_c916f887-6e77-4fc6-a692-30375d28650f	u2af1-wt
TCGA-49-4505-01A_0ebf5cc5-f242-45ef-821a-939b51dc95a2	u2af1-s34f
TCGA-MP-A4T4-01A_9eeae6b9-2031-47fa-80db-e04d53f0bfbd	u2af1-s34f



sig.tsv

splice_interval	median_u2af1-wt	mean_u2af1-wt	delta_u2af1-wt	pval_u2af1-wt	median_u2af1-s34f	mean_u2af1-s34f	delta_u2af1-s34f	pval_u2af1-s34f
chr1:17368-17605:+	0.66	0.67	-0.01	0.50	0.78	0.78	0.11	0.02
chr1:17368-17605:-	0.79	0.79	-0.01	0.50	0.89	0.88	0.08	0.02

signature.py fit_beta (splicedice)

(selected contents, not in order)

allPS.tsv

cluster	TCGA-55-A4...	TCGA-78-76...
chr1:11211-12009:+	nan	nan
chr1:13052-13220:+	nan	0.500
chr1:14784-14977:-	0.000	0.000
chr1:14829-15020:+	0.857	1.000
chr1:17055-17605:-	0.060	0.007

sig_manifest.tsv

TCGA-55-A4DF-01A_4a5e9e8a-8c48-48cf-8bf0-eb564611d382	u2af1-wt
TCGA-78-7633-01A_c916f887-6e77-4fc6-a692-30375d28650f	u2af1-wt
TCGA-49-4505-01A_0ebf5cc5-f242-45ef-821a-939b51dc95a2	u2af1-s34f
TCGA-MP-A4T4-01A_9eeae6b9-2031-47fa-80db-e04d53f0bfbd	u2af1-s34f

sig.tsv

splice_interval	median_u2af1-wt	mean_u2af1-wt	delta_u2af1-wt	pval_u2af1-wt	median_u2af1-s34f	mean_u2af1-s34f	delta_u2af1-s34f	pval_u2af1-s34f
chr1:17368-17605:+	0.66	0.67	-0.01	0.50	0.78	0.78	0.11	0.02
chr1:17368-17605:-	0.79	0.79	-0.01	0.50	0.89	0.88	0.08	0.02

beta.tsv

splice_interval	median_u2af1-wt	alpha_u2af1-wt	beta_u2af1-wt	median_u2af1-s34f	alpha_u2af1-s34f	beta_u2af1-s34f
chr1:17368-17605:+	0.66	2.61	1.16	0.78	4.58	1.17
chr1:17368-17605:-	0.79	4.69	1.20	0.89	7.90	1.04
chr1:185350-185490:-	0.93	7.68	0.56	1.00	12.84	0.37
chr1:18366-24737:-	0.37	4.30	6.94	0.44	4.88	4.96

signature.py fit_beta

test if splicedice can find the samples it generated the query from

signature.py query (splicedice)

(selected contents, not in order)

beta.tsv

splice_interval	median_u2af1-wt	alpha_u2af1-wt	beta_u2af1-wt	median_u2af1-s34f	alpha_u2af1-s34f	beta_u2af1-s34f
chr1:17368-17605:+	0.66	2.61	1.16	0.78	4.58	1.17
chr1:17368-17605:-	0.79	4.69	1.20	0.89	7.90	1.04
chr1:185350-185490:-	0.93	7.68	0.56	1.00	12.84	0.37

allPS.tsv

cluster	TCGA-55-A4...	TCGA-78-76...
chr1:11211-12009:+	nan	nan
chr1:13052-13220:+	nan	0.500
chr1:14784-14977:-	0.000	0.000
chr1:14829-15020:+	0.857	1.000
chr1:17055-17605:-	0.060	0.007

pvals.tsv
(transposed; raw)

query	u2af1-wt_over_u2af1-s34f	u2af1-s34f_over_u2af1-wt
TCGA-55-A4...	1.2504576745225453e-299	1.0
TCGA-78-76...	1.1529596962869242e-306	1.0

pvals.tsv
(transposed; summarized)

query	u2af1-wt_over_u2af1-s34f	u2af1-s34f_over_u2af1-wt
n=82	0.00	1.00
n=11	1.00	0.00

the 11 are all u2af1-s34f; the 82 are all u2af1-wt

signature.py
query

Future options

- repeat with dividing by other mutation status or arbitrary selection of 12 samples
- QC on bams