

Three-Way Cross-Tabulation: PurIST × DeCAF × DeSurv Risk Group

DeSurv Paper — Supplementary Analysis

Overview

This report presents the joint distribution of two established PDAC molecular classifiers — PurIST (Basal-like vs Classical) and DeCAF (proCAF vs restCAF) — stratified by the DeSurv-derived dichotomized risk groups (High vs Low). Risk groups are defined by applying the log-rank-optimized z-score cutpoint to the DeSurv linear predictor, standardized using the training cohort mean and standard deviation.

The analysis is shown for both the training cohort (TCGA+CPTAC) and the pooled external validation cohorts (Dijk, Moffitt, PACA-AU, Puleo). Per-cohort enrichment figures follow.

1. Pooled External Validation Cohorts

Table 1: PurIST × DeCAF × Risk Group — Pooled Validation (logrank cutpoint)

PurIST	DeCAF	High	Low	Total	% High
Classical	restCAF	14	222	236	5.9%
Classical	proCAF	53	182	235	22.6%
Basal-like	restCAF	7	18	25	28.0%
Basal-like	proCAF	42	32	74	56.8%
Total		116	454	570	20.4%

Cohorts: Dijk (n=90), Moffitt (n=123), PACA-AU (n=69), Puleo (n=288). Fisher's exact test: PurIST × Risk p=3.0e-13; DeCAF × Risk p=4.7e-12.

2. Training Cohort (TCGA+CPTAC)

Table 2: PurIST × DeCAF × Risk Group — Training (logrank cutpoint)

PurIST	DeCAF	High	Low	Total	% High
Classical	restCAF	0	119	119	0.0%
Classical	proCAF	6	88	94	6.4%
Basal-like	restCAF	5	13	18	27.8%
Basal-like	proCAF	25	17	42	59.5%
Total		36	237	273	13.2%

Training cohort: TCGA PAAD + CPTAC (n=273). Fisher's exact test within PurlST strata: Classical p=0.007; Basal-like p=0.047.

3. Summary

Monotonic gradient across subtype combinations. The percentage of samples classified as High risk increases monotonically across PurlST × DeCAF combinations:

Subtype Combination	% High (Training)	% High (Validation)
Classical / restCAF	0.0%	5.9%
Classical / proCAF	6.4%	22.6%
Basal-like / restCAF	27.8%	28.0%
Basal-like / proCAF	59.5%	56.8%

DeCAF adds prognostic information within PurlST strata. Among Classical tumors, proCAF status increases the High-risk rate from ~0–6% to ~6–23%. Among Basal-like tumors, proCAF increases it from ~28% to ~57–60%. Fisher's exact tests confirm this association is significant within both strata (Basal-like: p=0.020 pooled, p=0.047 training; Classical: p=1.6e-7 pooled, p=0.007 training).

Training-validation consistency. The Basal-like/proCAF combination shows 59.5% High in training vs 56.8% in validation, suggesting the cutpoint generalizes well. The training cohort has fewer Highs overall (13.2% vs 20.4%), consistent with the cutpoint being optimized on training data.

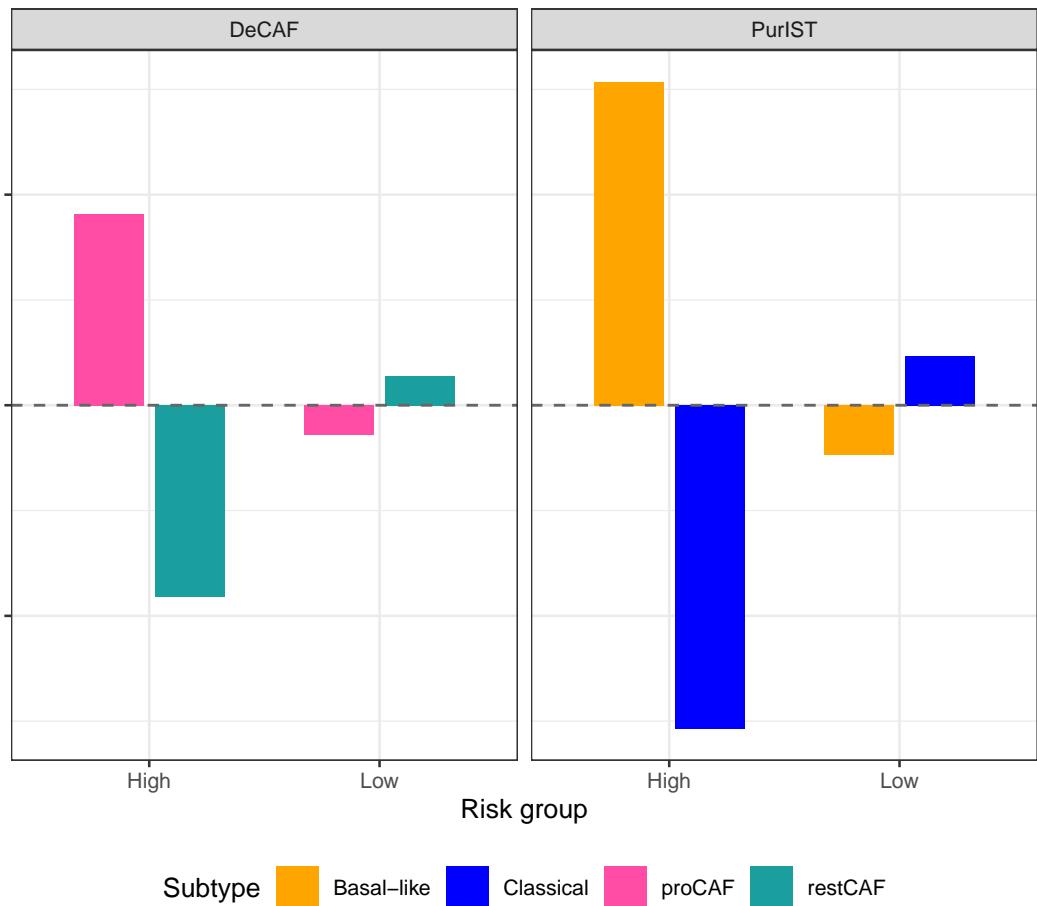
Interpretation. DeSurv's survival-driven factorization recovers a risk score that captures information from both PurlST and DeCAF simultaneously, without having access to either classifier during training. The joint enrichment pattern suggests the DeSurv linear predictor integrates both tumor-intrinsic (Basal-like) and stromal (proCAF) biology into a single prognostic axis.

4. Enrichment Figures

The following pages show the three-panel enrichment analyses generated by `plot_subtype_enrichment()`. Panel A shows enrichment relative to marginal proportions. Panel B shows DeCAF composition stratified by PuriST × Risk group. Panel C shows the raw sample counts (2×2×2 contingency table).

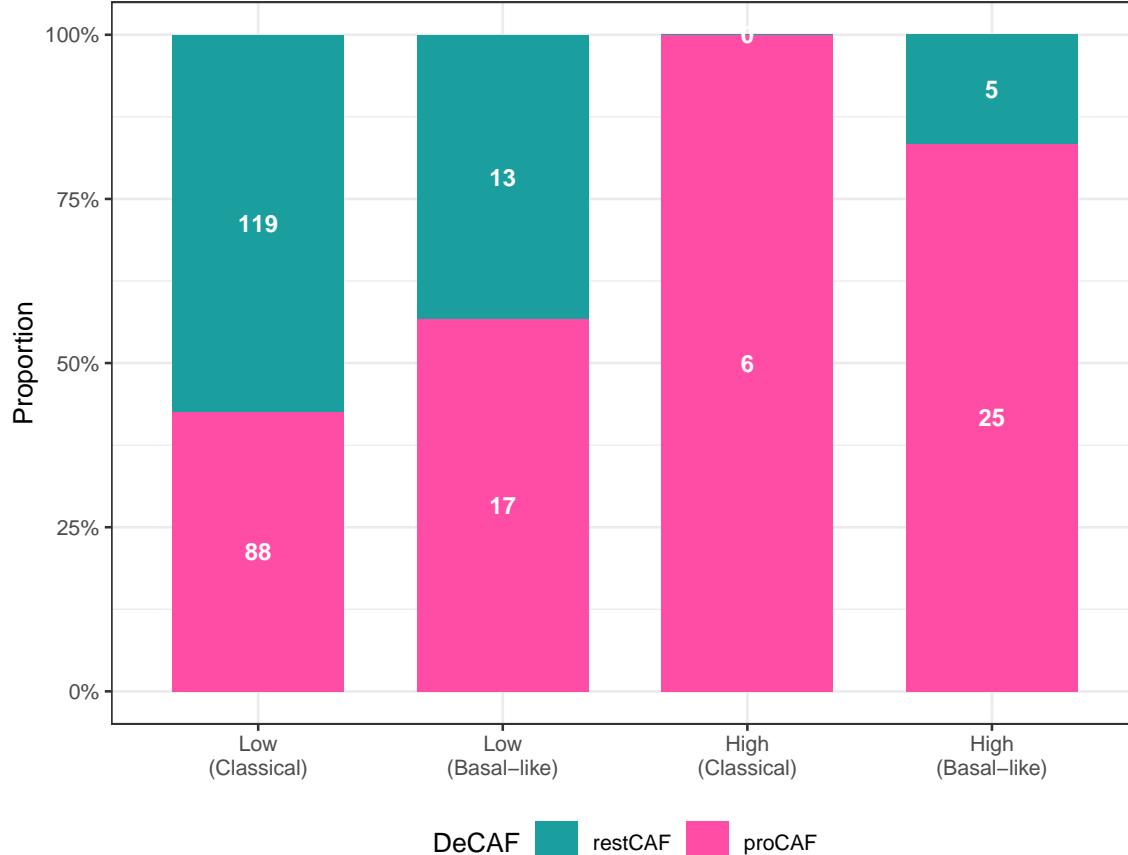
Training (logrank cutpoint)

A. Enrichment relative to marginal



B. DeCAF by PurlST × Risk group

Fisher Classical: $p = 0.007$ Fisher Basal-like: $p = 0.047$



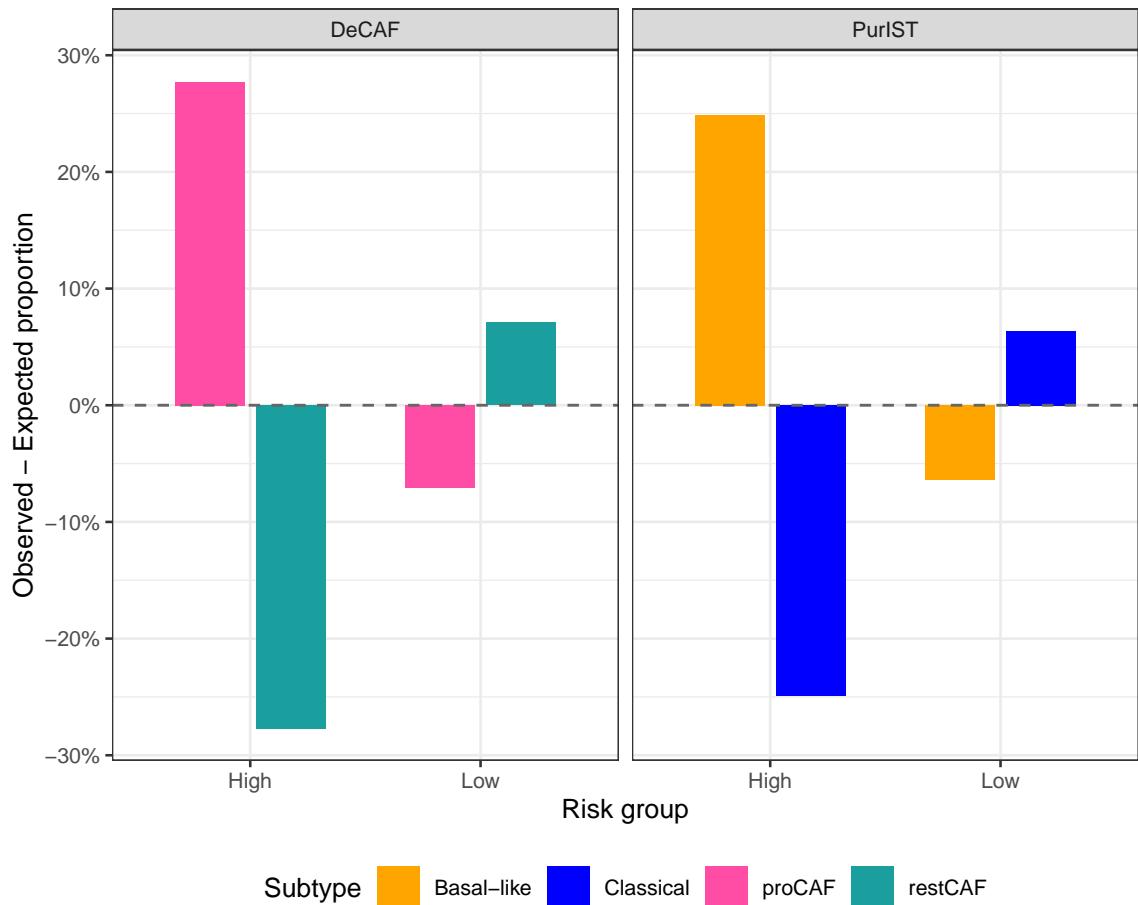
C. Sample counts

	High	Low
Classical / restCAF	0	119
Classical / proCAF	6	88
Basal-like / restCAF	5	13
Basal-like / proCAF	25	17

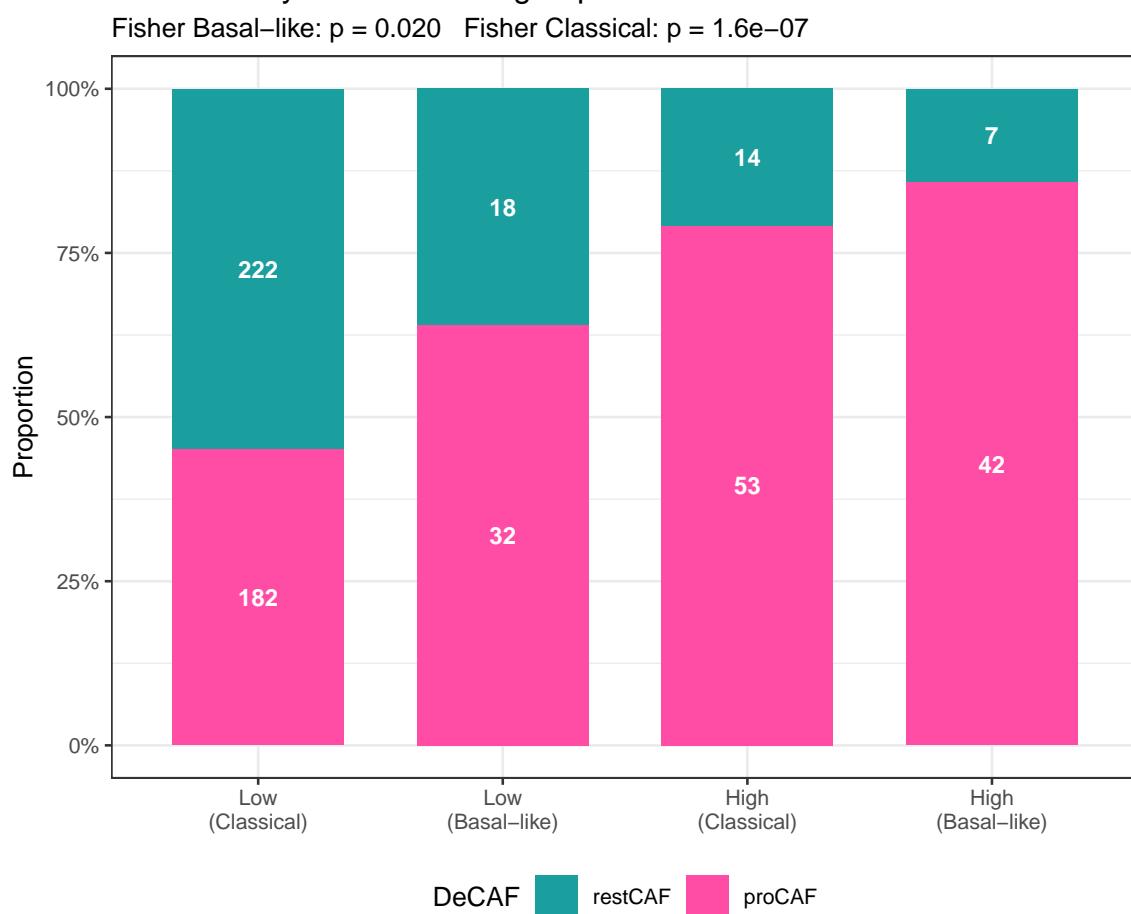
Risk group

Pooled validation (logrank cutpoint)

A. Enrichment relative to marginal



B. DeCAF by PurlST x Risk group



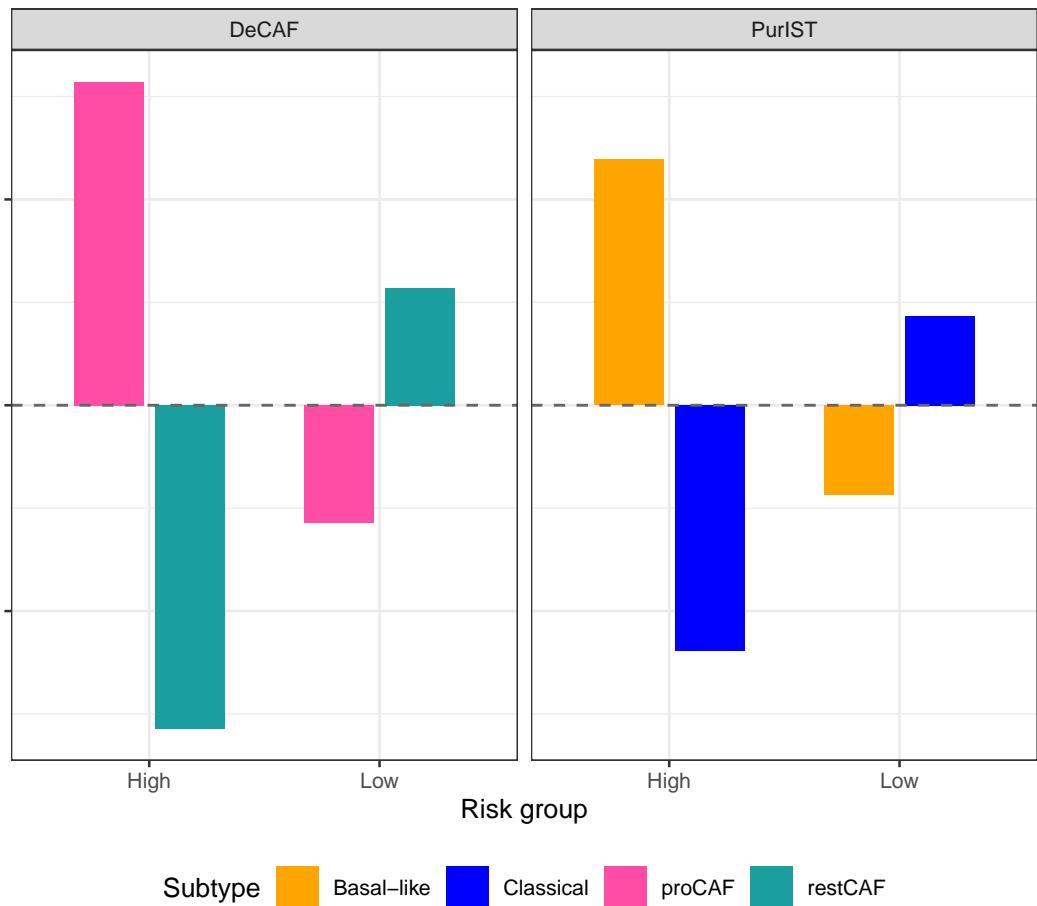
C. Sample counts

	High	Low
Classical / restCAF	14	222
Classical / proCAF	53	182
Basal-like / restCAF	7	18
Basal-like / proCAF	42	32

Risk group

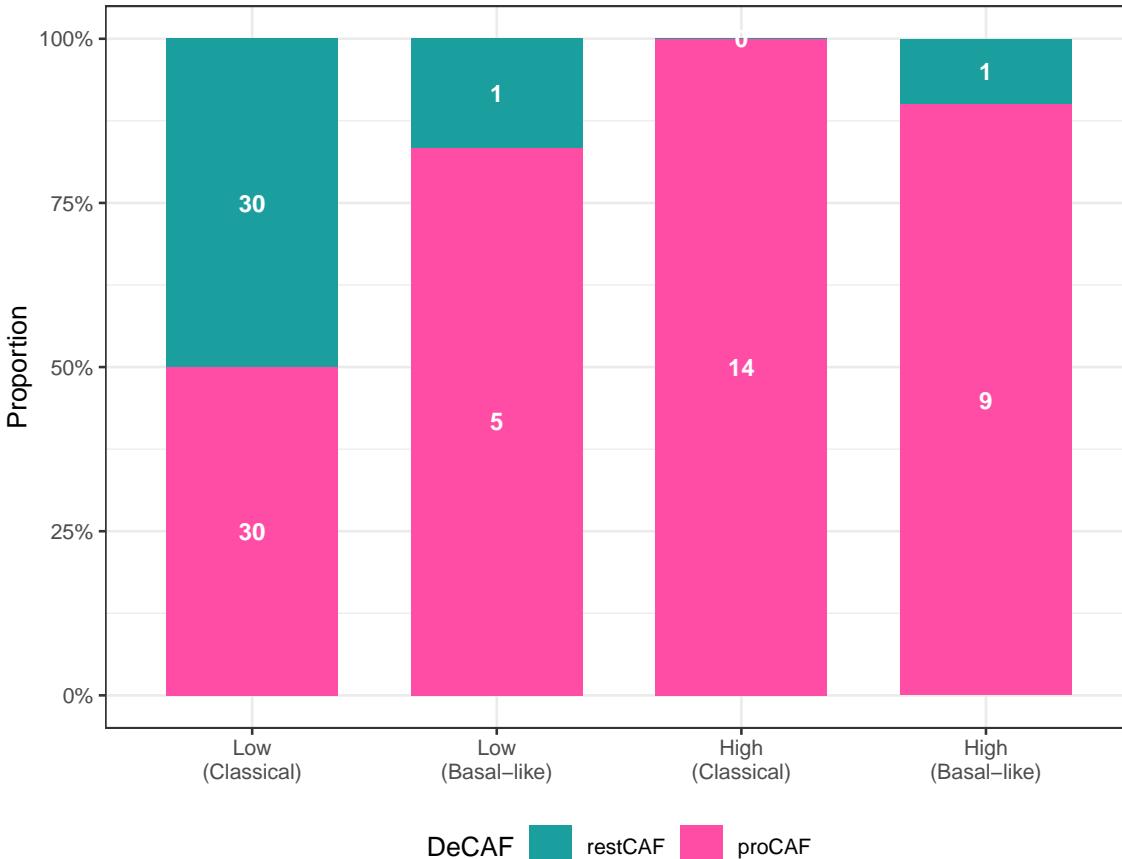
Dijk (logrank cutpoint)

A. Enrichment relative to marginal



B. DeCAF by PurlST x Risk group

Fisher Basal-like: $p = 1.000$ Fisher Classical: $p = 4.4e-04$



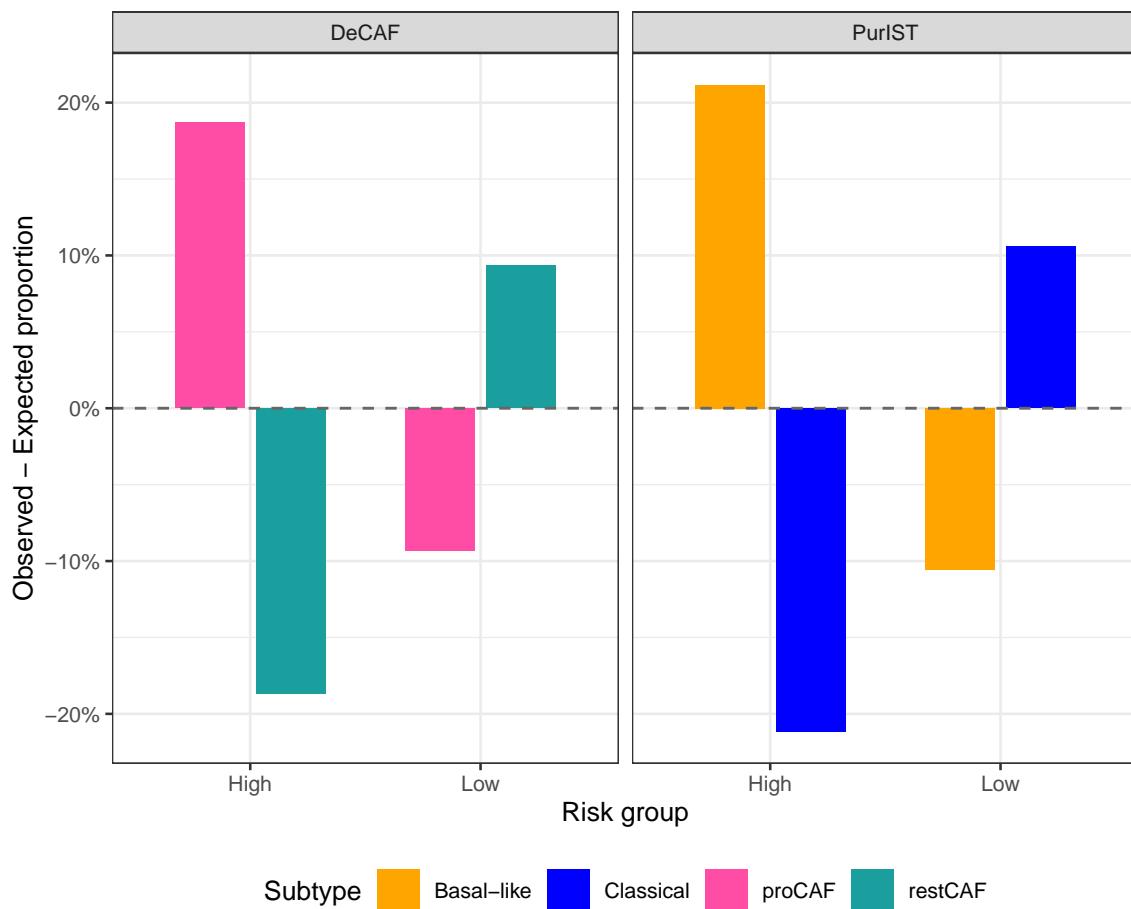
C. Sample counts

	High	Low
Classical / restCAF	0	30
Classical / proCAF	14	30
Basal-like / restCAF	1	1
Basal-like / proCAF	9	5

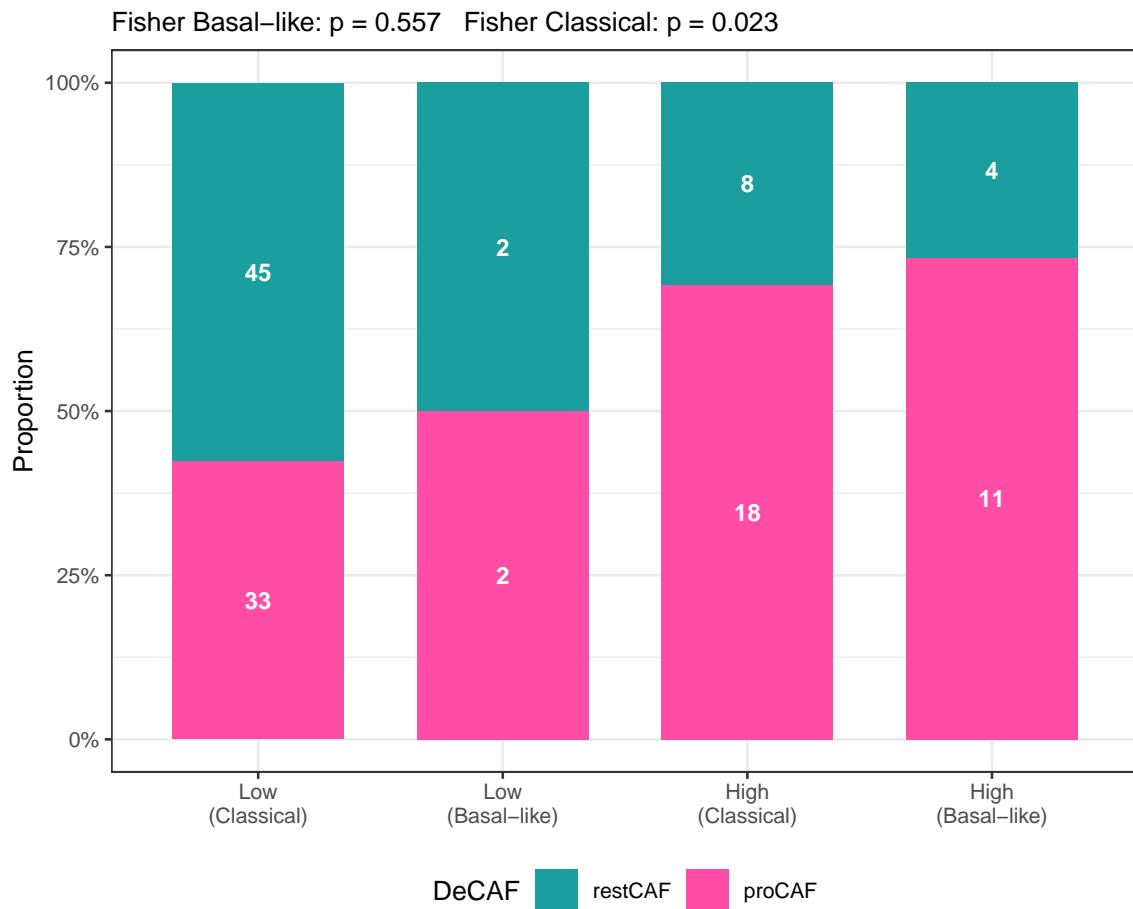
Risk group

Moffitt_GEO_array (logrank cutpoint)

A. Enrichment relative to marginal



B. DeCAF by PurlST x Risk group



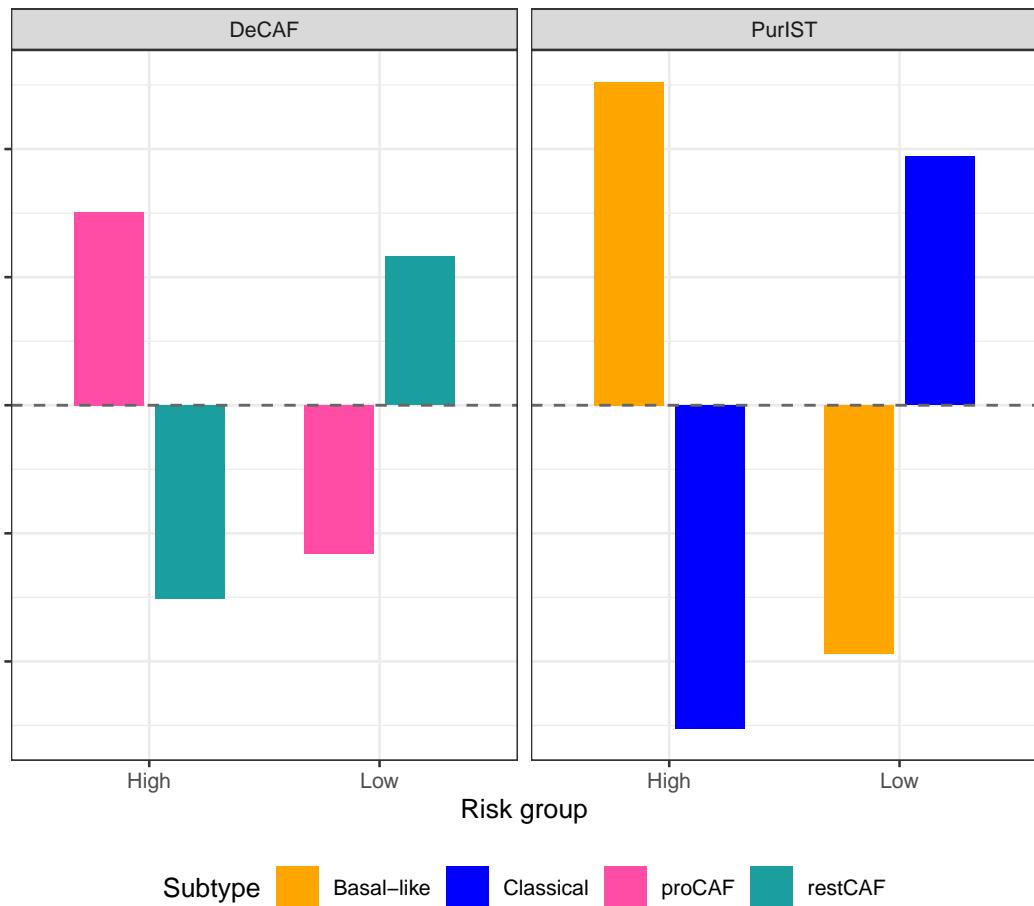
C. Sample counts

	High	Low
Classical / restCAF	8	45
Classical / proCAF	18	33
Basal-like / restCAF	4	2
Basal-like / proCAF	11	2

Risk group

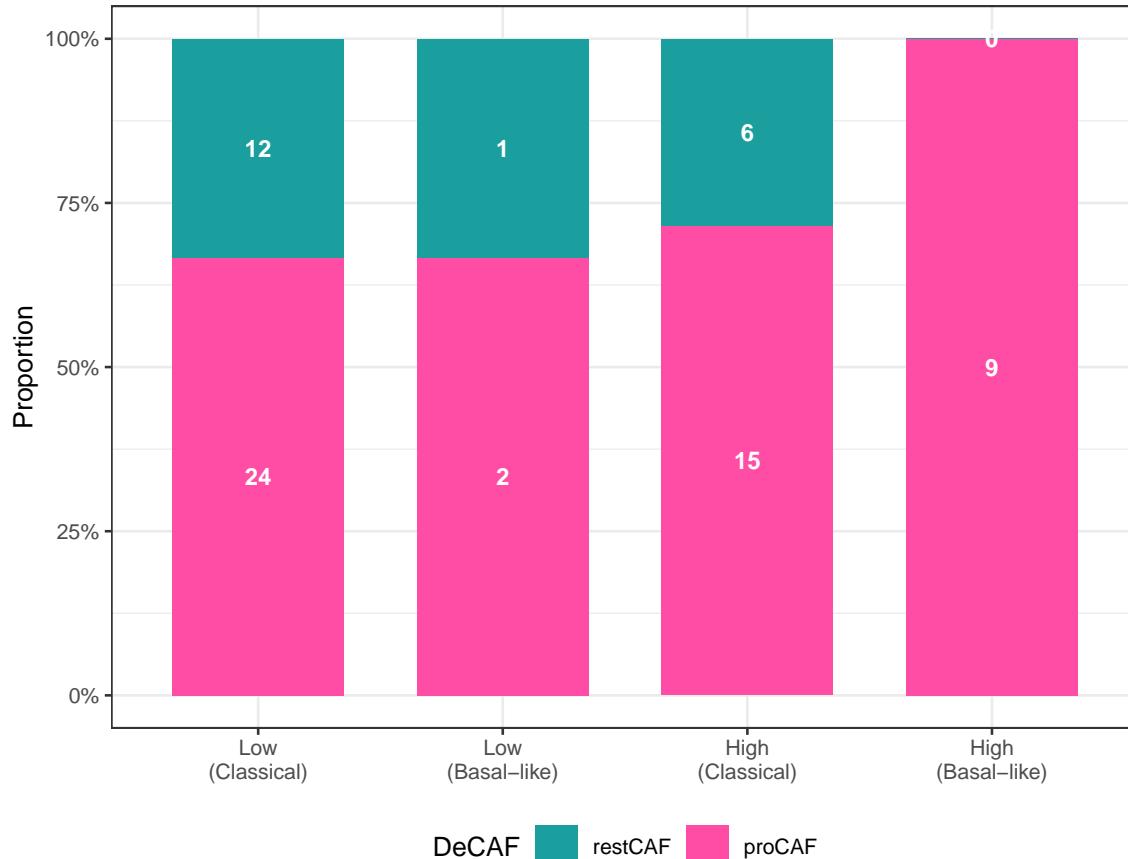
PACA_AU (logrank cutpoint)

A. Enrichment relative to marginal



B. DeCAF by PurlST x Risk group

Fisher Classical: $p = 0.775$ Fisher Basal-like: $p = 0.250$



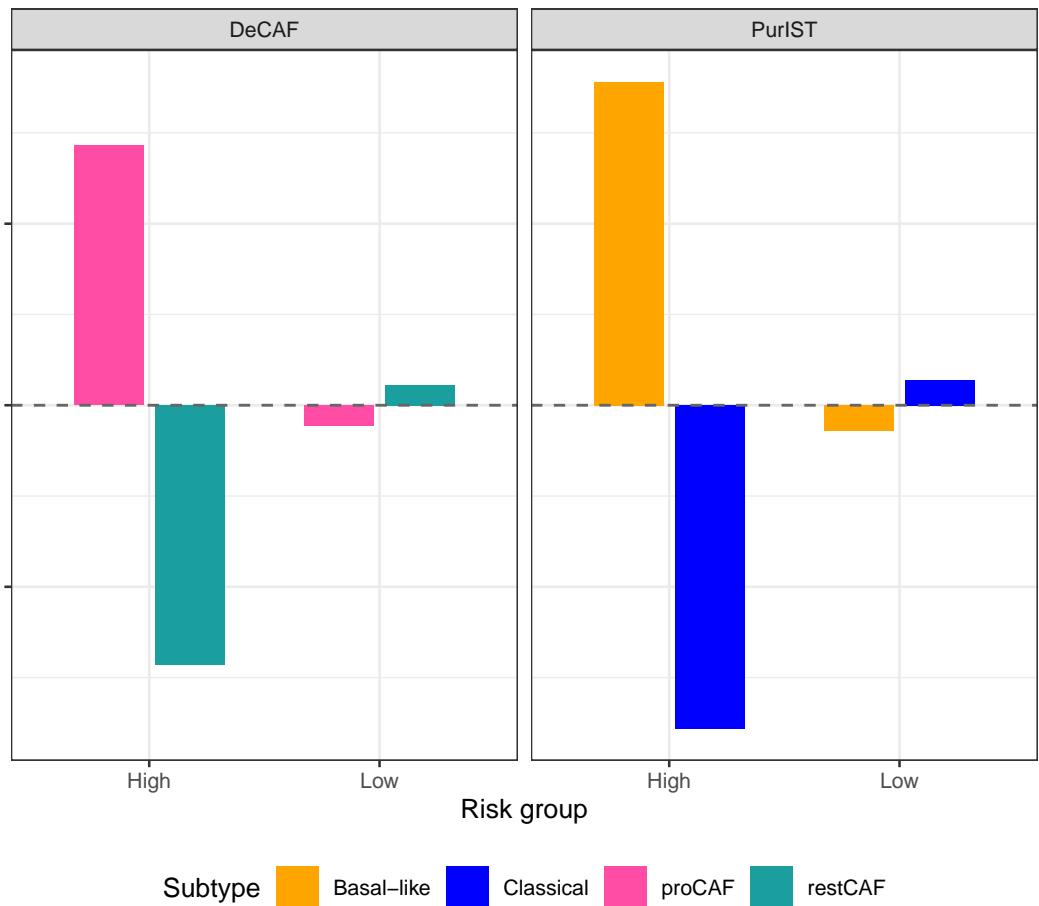
C. Sample counts

	High	Low
Classical / restCAF	6	12
Classical / proCAF	15	24
Basal-like / restCAF	0	1
Basal-like / proCAF	9	2

Risk group

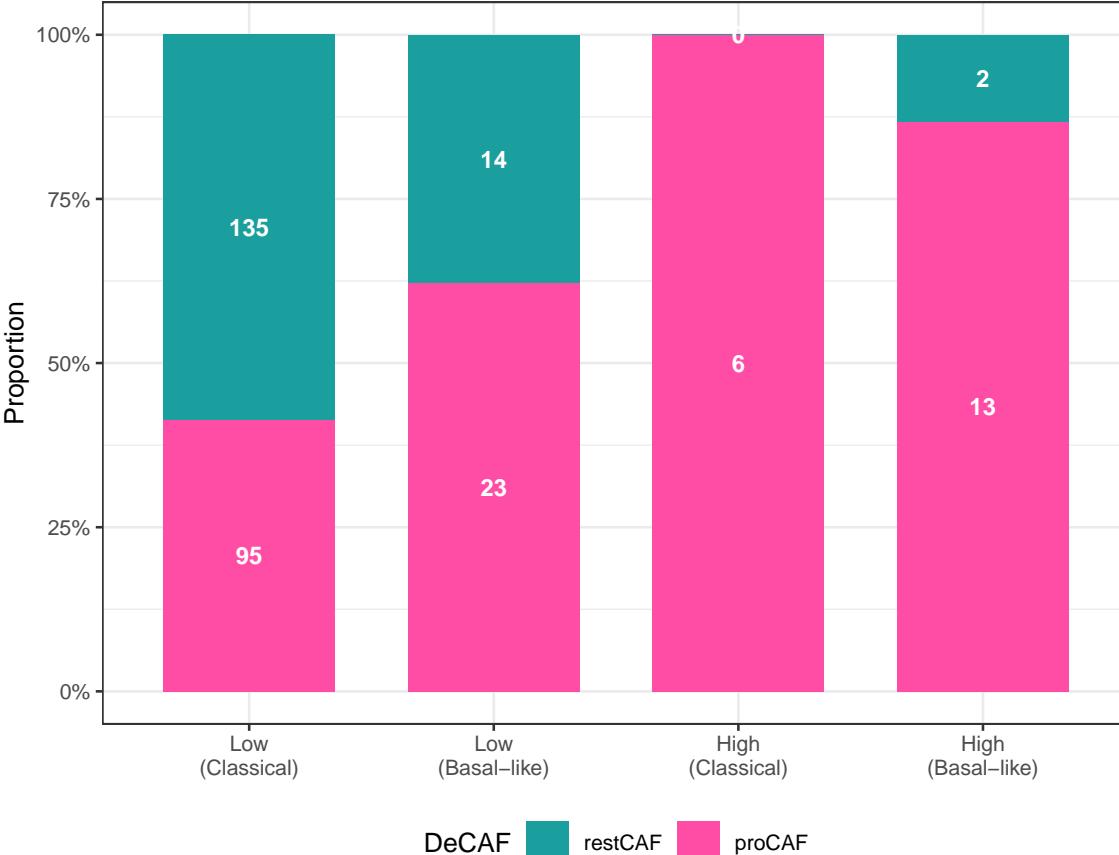
Puleo_array (logrank cutpoint)

A. Enrichment relative to marginal



B. DeCAF by PurlST x Risk group

Fisher Classical: $p = 0.006$ Fisher Basal-like: $p = 0.106$



C. Sample counts

	High	Low
Classical / restCAF	0	135
Classical / proCAF	6	95
Basal-like / restCAF	2	14
Basal-like / proCAF	13	23

Risk group