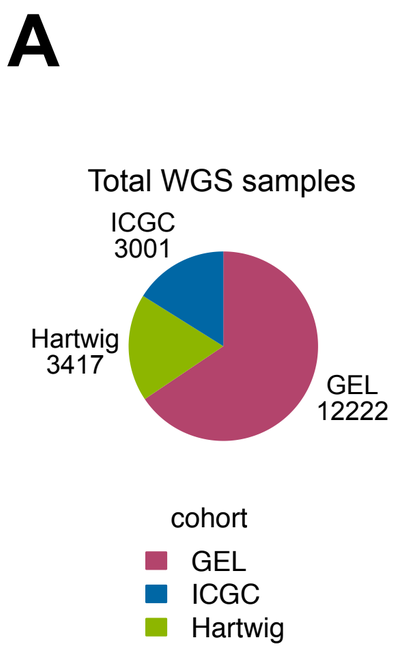
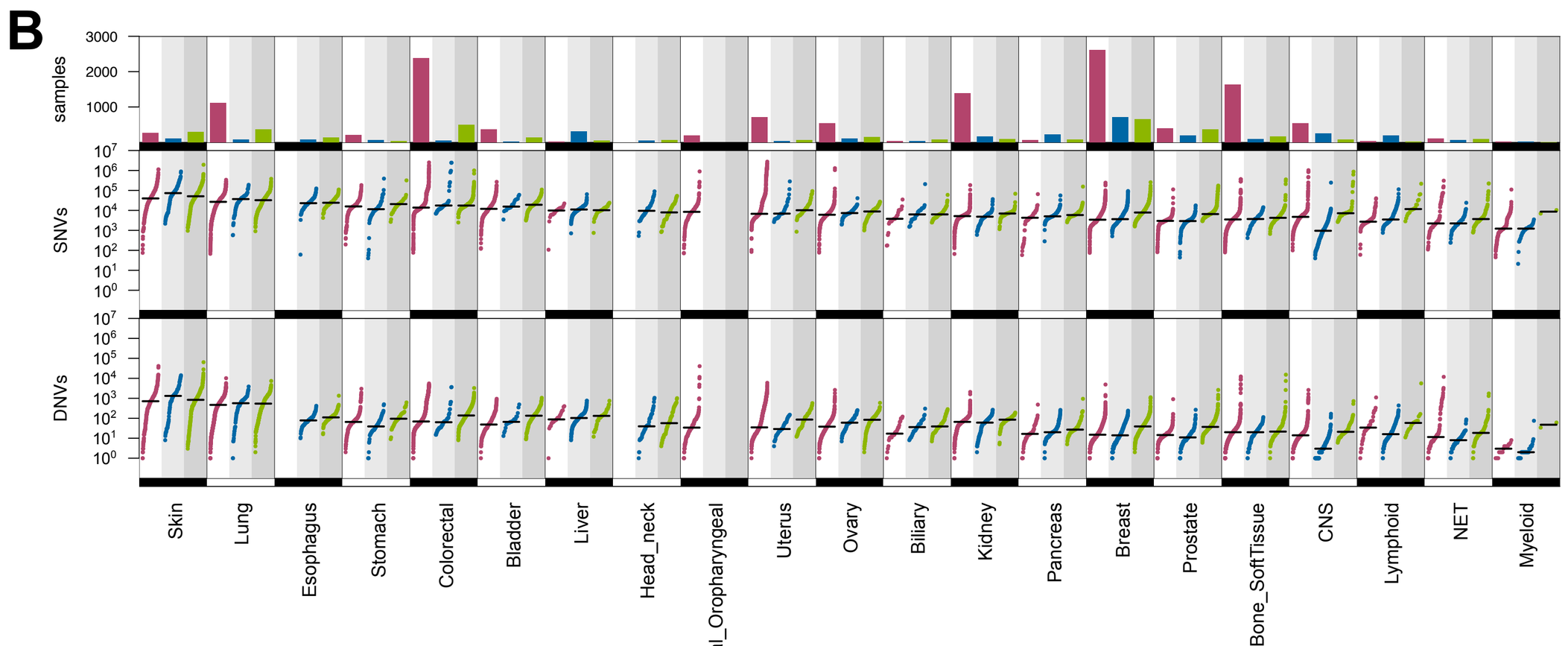
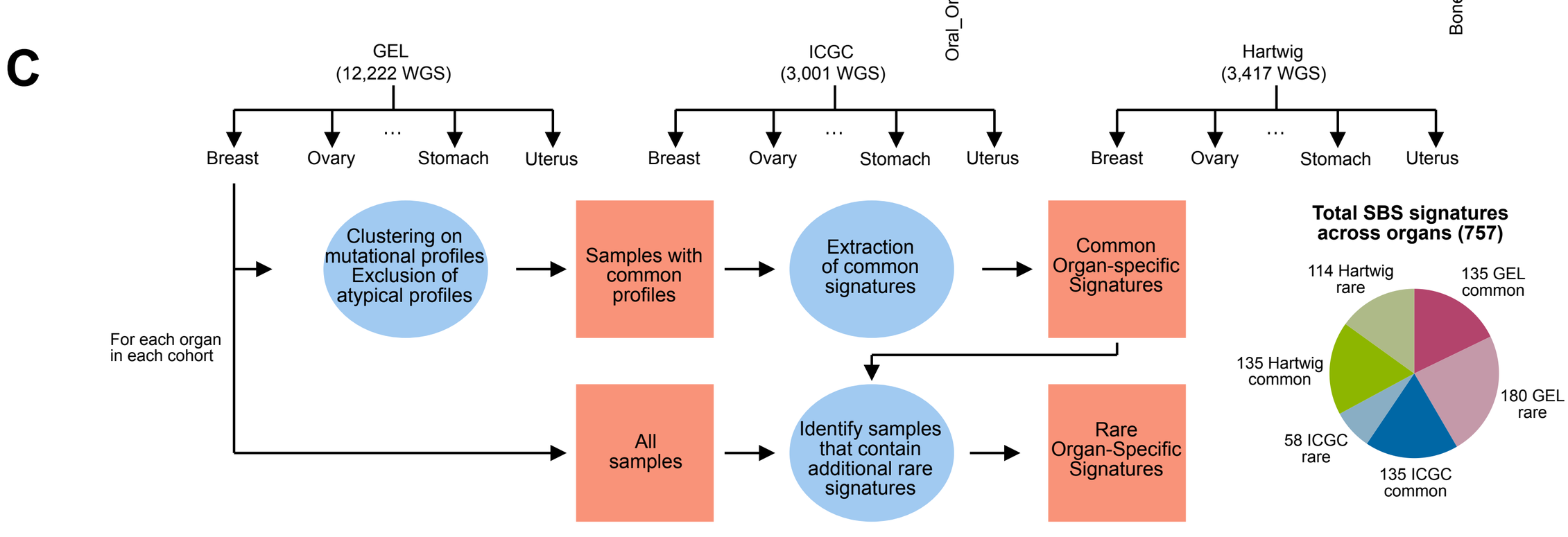
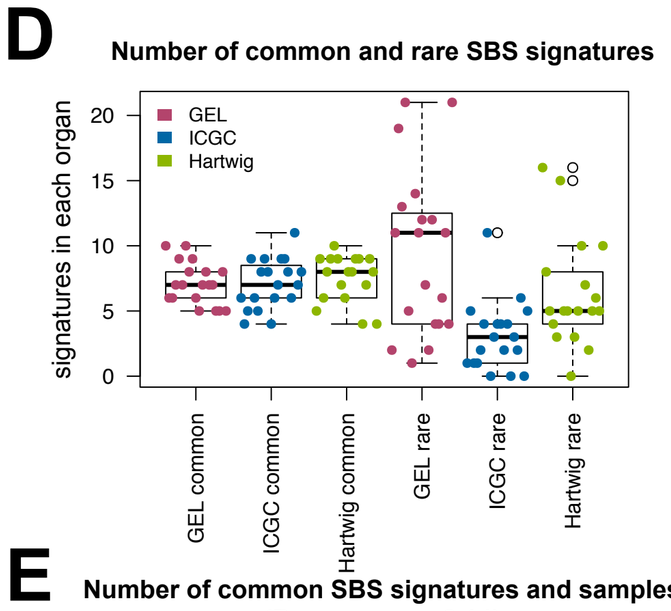
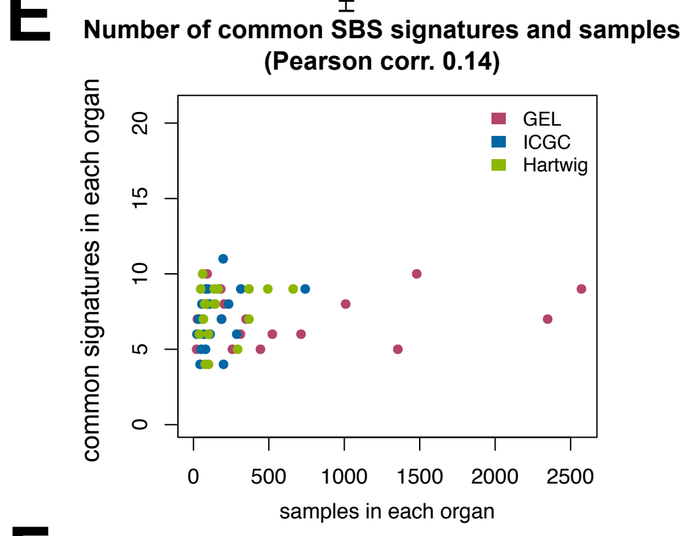
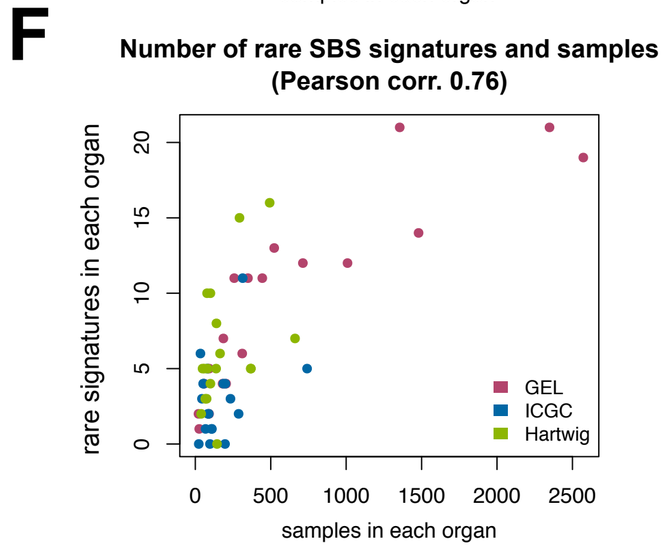

A
Total WGS samples
ICGC
3001
Hartwig
3417
GEL
12222
cohort
GEL
ICGC
Hartwig

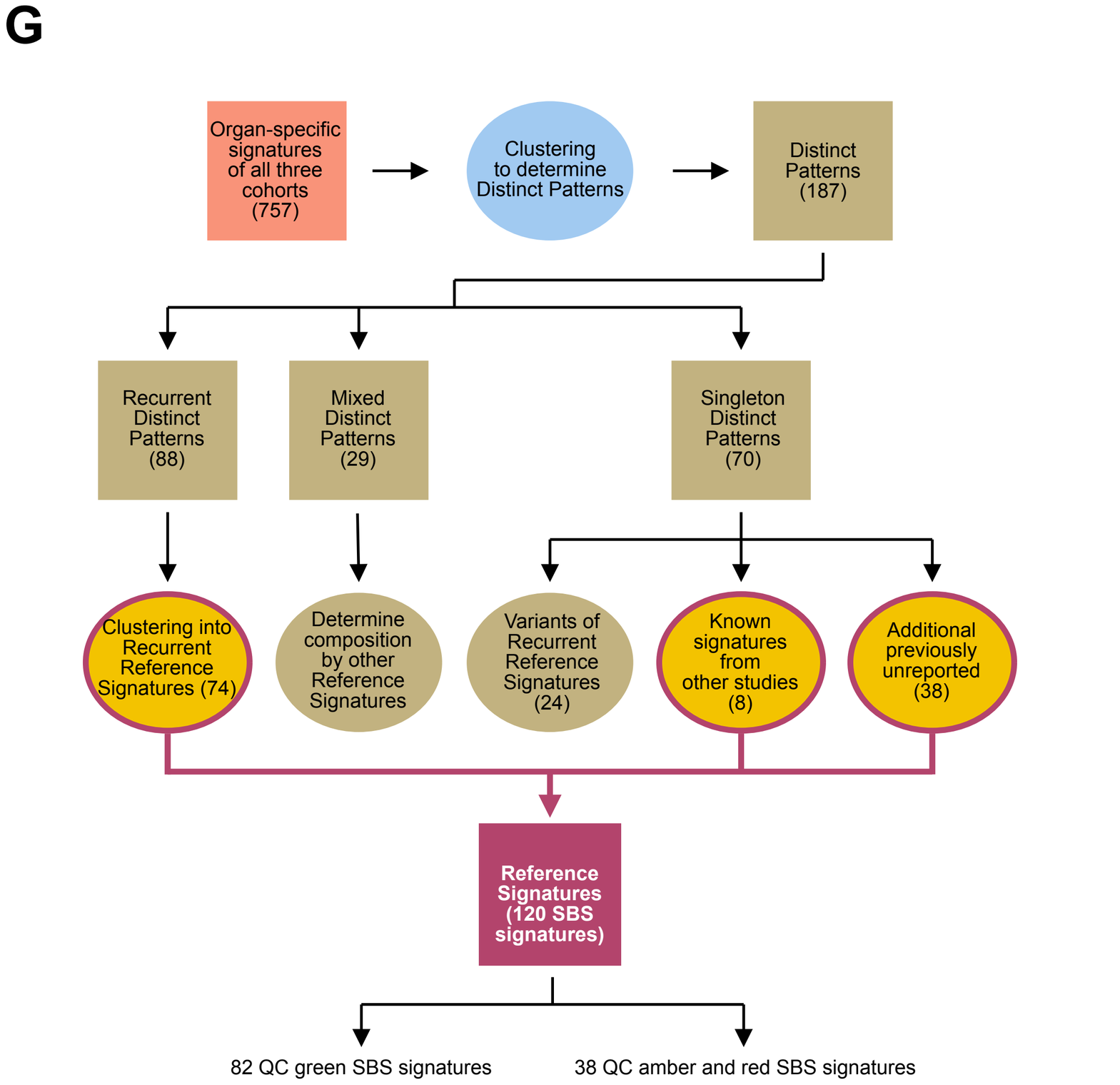

B
samples
SNVs
DNVs
Skin
Lung
Esophagus
Stomach
Colorectal
Bladder
Liver
Head_neck
Uterus
Ovary
Biliary
Kidney
Pancreas
Breast
Prostate
CNS
Lymphoid
NET
Myeloid


C
GEL
(12,222 WGS)
ICGC
(3,001 WGS)
Hartwig
(3,417 WGS)
Breast
Ovary
Stomach
Uterus
For each organ
in each cohort
Clustering on
mutational profiles
Exclusion of
atypical profiles
Samples with
common
profiles
Extraction
of common
signatures
Common
Organ-specific
Signatures
All
samples
Identify samples
that contain
additional rare
signatures
Rare
Organ-Specific
Signatures
Total SBS signatures
across organs (757)
114 Hartwig
rare
135 GEL
common
135 Hartwig
common
180 GEL
rare
58 ICGC
rare
135 ICGC
common


D
Number of common and rare SBS signatures
signatures in each organ
GEL
ICGC
Hartwig
GEL common
ICGC common
Hartwig common
GEL rare
ICGC rare
Hartwig rare


E
Number of common SBS signatures and samples
(Pearson corr. 0.14)
common signatures in each organ
samples in each organ
GEL
ICGC
Hartwig


F
Number of rare SBS signatures and samples
(Pearson corr. 0.76)
rare signatures in each organ
samples in each organ
GEL
ICGC
Hartwig


G
Organ-specific
signatures
of all three
cohorts
(757)
Clustering
to determine
Distinct Patterns
Distinct
Patterns
(187)
Recurrent
Distinct
Patterns
(88)
Mixed
Distinct
Patterns
(29)
Singleton
Distinct
Patterns
(70)
Clustering into
Recurrent
Reference
Signatures (74)
Determine
composition
by other
Reference
Signatures
Variants of
Recurrent
Reference
Signatures
(24)
Known
signatures
from
other studies
(8)
Additional
previously
unreported
(38)
Reference
Signatures
(120 SBS
signatures)
82 QC green SBS signatures
38 QC amber and red SBS signatures