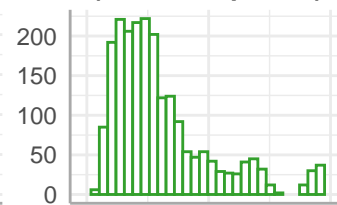
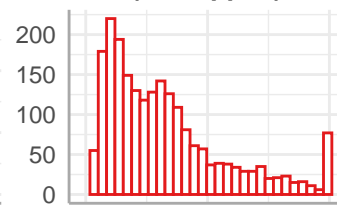
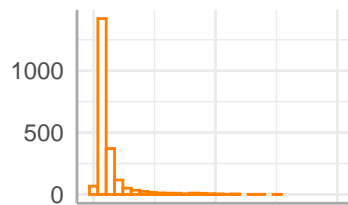


A

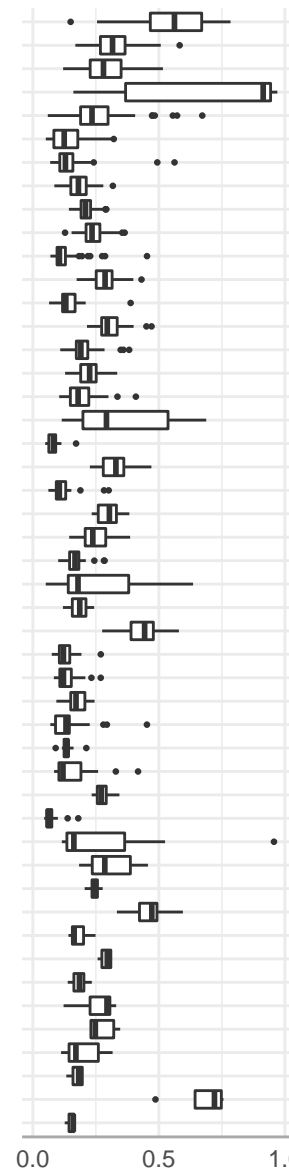
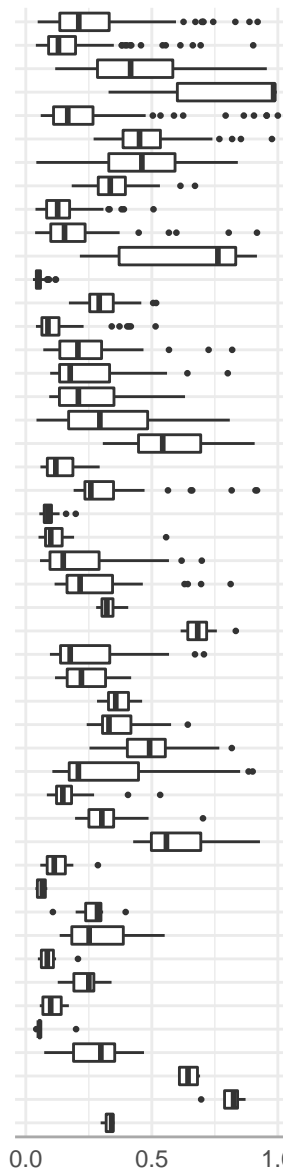
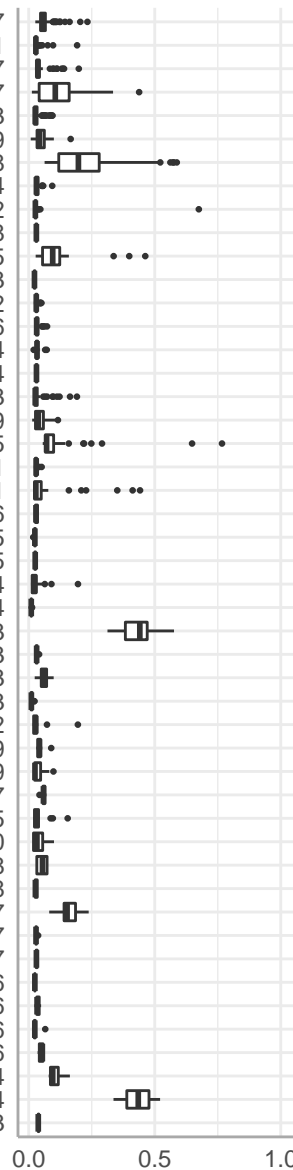
Datasets

% unmapped

% duplicate
(of mapped)% non-exonic
(of non-duplicate)**B**

Cohorts

Cohort_1, n=337
Cohort_2, n=141
Cohort_3, n=137
Cohort_4, n=127
Cohort_5, n=103
Cohort_6, n=89
Cohort_7, n=88
Cohort_8, n=84
Cohort_9, n=82
Cohort_10, n=78
Cohort_11, n=65
Cohort_12, n=63
Cohort_13, n=62
Cohort_14, n=56
Cohort_15, n=54
Cohort_16, n=44
Cohort_17, n=43
Cohort_18, n=39
Cohort_19, n=35
Cohort_21, n=31
Cohort_20, n=31
Cohort_22, n=26
Cohort_24, n=25
Cohort_23, n=25
Cohort_26, n=24
Cohort_25, n=24
Cohort_30, n=23
Cohort_29, n=23
Cohort_28, n=23
Cohort_27, n=23
Cohort_31, n=22
Cohort_33, n=19
Cohort_32, n=19
Cohort_34, n=17
Cohort_35, n=15
Cohort_36, n=10
Cohort_38, n=8
Cohort_37, n=8
Cohort_41, n=7
Cohort_40, n=7
Cohort_39, n=7
Cohort_45, n=6
Cohort_44, n=6
Cohort_43, n=6
Cohort_42, n=6
Cohort_47, n=4
Cohort_46, n=4
Cohort_48, n=3



Read type percentages in 2179 datasets

C

genes > 1 TPM

genes > 2 TPM

genes > 3 TPM

Gene counts (thousand)

 $r=-0.05$ $r=-0.04$ $r=-0.05$ $r=-0.04$ $r=-0.03$ $r=-0.03$ $r=0.22$ $r=0.21$ $r=0.2$ $r=0.23$ $r=0.26$ $r=0.28$

Total_reads

Mapped

MND

MEND

Read counts (million)