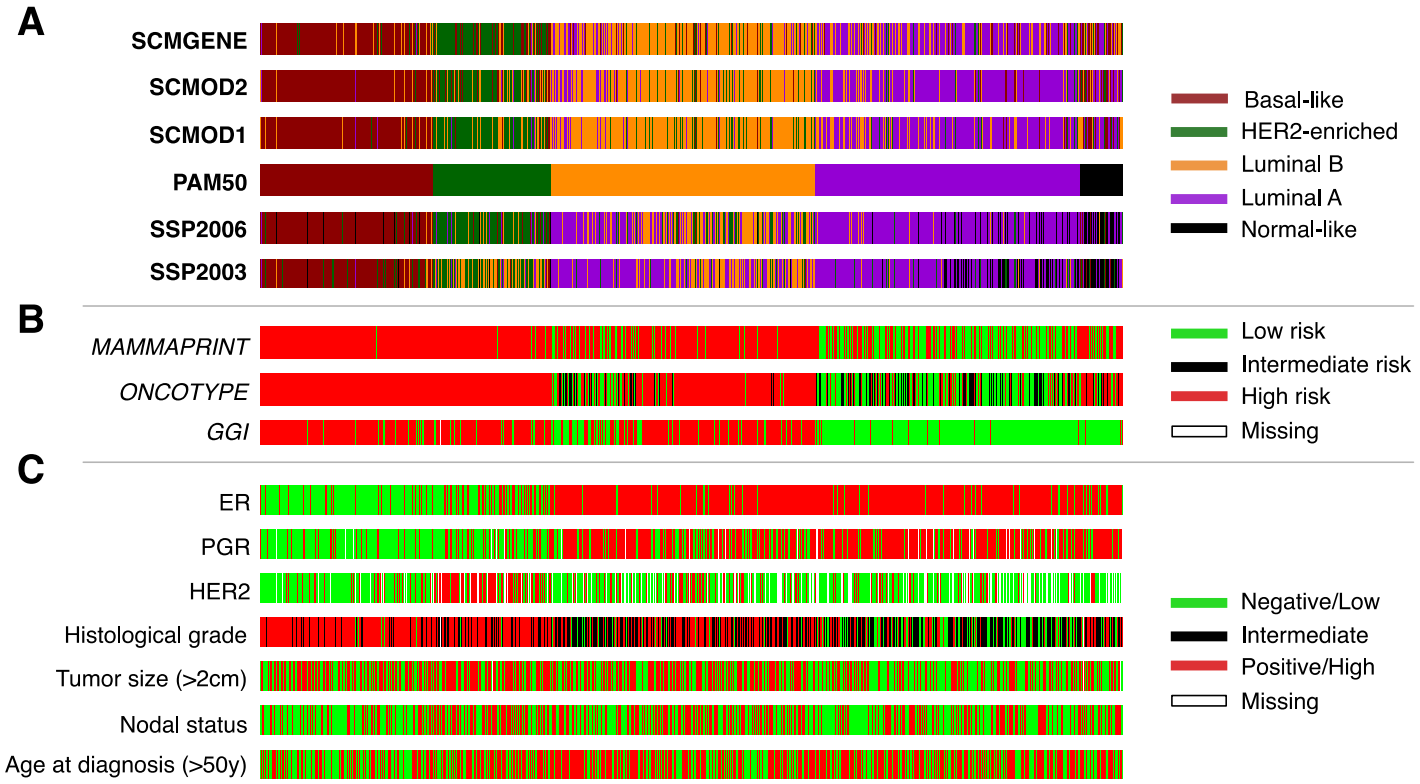
Supplementary Figure 1. Concordance of classifiers for breast cancer molecular subtyping. Colored bars illustrate the molecular subtypes as computed by each of the six classifiers applied to our compendium of 5715 breast tumors (**A**); SCMGENE: three-gene subtype classification model; SCMOD2: subtype classification model published by Wirapati et al. (8); SCMOD1: subtype classification model published by Desmedt et al. (1); PAM50: single sample predictor published by Parker et al. (3); SSP2006: single sample predictor published by Hu et al. (2). SSP2003: single sample predictor published Sorlie et al. (6). The most recent single sample predictor (PAM50) was taken as reference, i.e., the patients (tumors) were unambiguously ordered using the correlation to each centroid as estimated by PAM50. The corresponding risk predicted by the prognostic gene signatures [MAMMAPRINT: prognostic gene signature published by van't Veer et al. (14); ONCOTYPE: prognostic gene signature published by Paik et al. (15); GGI: prognostic gene signature published by Sotiriou et al. (16)] are represented in the middle of the figure (**B**). The clinical parameters are illustrated at the bottom (**C**) with estrogen receptor (ER) and progesterone receptor (PGR) status being defined by immunohistochemistry (IHC), human epidermal growth factor 2 (HER2) status by IHC or fluorescent in situ hybridization (FISH), histological grade being assessed separately in each dataset, and age at diagnosis (> 50 years) and tumor size (> 2 cm) being considered as binary variables.


A
SCMGENE
SCMOD2
SCMOD1
PAM50
SSP2006
SSP2003
Basal-like
HER2-enriched
Luminal B
Luminal A
Normal-like
B
MAMMAPRINT
ONCOTYPE
GGI
Low risk
Intermediate risk
High risk
Missing
C
ER
PGR
HER2
Histological grade
Tumor size (>2cm)
Nodal status
Age at diagnosis (>50y)
Negative/Low
Intermediate
Positive/High
Missing