Table S1. Summary of mutations used as clinical factors in this study

	Mutations		
	KRAS	**BRAF**	**PIK3CA**
PETACC3	G12, G13	V600E	Microarray
TCGA	G12, G13	V600E	R88, E542, E545, H1047
Agendia (GSE42284)	G12, G13, Q61	V600E	E542, E545, Q546, D549, H1047
Agendia (ICO)	G12, G13, Q61	V600E	E542, E545, Q546, D549, H1047
Agendia (VHB)	G12, G13, Q61	V600E	E542, E545, Q546, D549, H1047
Marisa *et al*. (GSE39582)	G12, G13	V600E	No info
Schlicker *et al*. (GSE35896)	NA	NA	NA
GSE75315	G12, G13, Q61	V600E	E542, E545, Q546, H1047
GSE75316	G12, G13, Q61	V600E	E542, E545, Q546, H1047

NA: Data not available
No info: No information available concerning mutation origin

Legend

This table shows which amino acid change were considered as mutations for the KRAS, BRAF and PIK3CA genes according to the cohort.

Table S2. Non-zero coefficients genes obtained by training a penalized generalized linear model to classify BM subtypes

Positive Genes	Coefficient	Negative Genes	Coefficient
BASP1	0.342485896	ARSE	-0.01401327
CMTM3	0.031282542	CASP6	-0.09803343
DPT	0.005840771	CDCA7	-0.090138226
EGR2	0.133714218	CLDN7	-0.232541451
F13A1	0.188165758	ETFA	-0.180873579
FBXL7	0.144898935	FAHD1	-0.116693331
GAS7	0.042158651	IL17RE	-0.094899931
GGT5	0.031779867	KIF11	-0.116408835
GPX7	0.229157793	LTA4H	-0.187723635
HEYL	0.071816609	METAP1	-0.047317976
ID4	0.118511862	MGST1	-0.005672195
KLHDC7B	0.052945152	PIK3R4	-0.11800317
LRRC33	0.230244064	POLR3B	-0.058432825
LTBP2	0.092759565	RAD51	-0.476755217
MAN1C1	0.000244544	SYAP1	-0.274105589
NGFR	0.195062095	TOM1L1	-0.236090048
PIK3CD	0.089712298	TPD52	-0.283762662
PPP1R16B	0.121076115	YARS2	-0.016994029
PRICKLE1	0.478722755	ZWINT	-0.059964679
PTGIS	0.17287554		
RGL1	0.330410084		
RUNX1T1	0.127930532		
SLFN11	0.091951297		
TM6SF1	0.08743986		
TSHZ3	0.22659478		

Legend

Table displaying the genes with non-zeroe coefficients after having trained a generalized logistic (binomial method) model via lasso penalty on the BRAF mutant dataset and using the BM1 and BM2 calls as response variables. These genes can be used to classify any sample with a gene expression profile into BM1 or BM2.