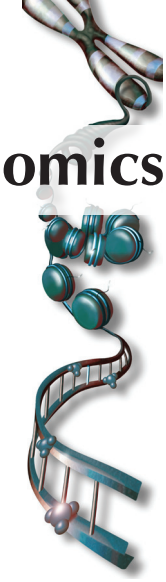
Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



H3K4 acetylation, H3K9 acetylation and H3K27 methylation in breast tumor molecular subtypes

Aim: Here, we investigated how the St Gallen breast molecular subtypes displayed distinct histone H3 profiles. **Patients & methods:** 192 breast tumors divided into five St Gallen molecular subtypes (luminal A, luminal B HER2-, luminal B HER2+, HER2+ and basal-like) were evaluated for their histone H3 modifications on gene promoters. **Results:** ANOVA analysis allowed to identify specific H3 signatures according to three groups of genes: hormonal receptor genes (*ERS1*, *ERS2*, *PGR*), genes modifying histones (*EZH2*, *P300*, *SRC3*) and tumor suppressor gene (*BRCA1*). A similar profile inside high-risk cancers (luminal B [HER2+], HER2+ and basal-like) compared with low-risk cancers including luminal A and luminal B (HER2-) were demonstrated. **Conclusion:** The H3 modifications might contribute to clarify the differences between breast cancer subtypes.

First draft submitted: 11 February 2016; Accepted for publication: 4 May 2016; Published online: 18 July 2016

Keywords: breast cancer • H3 modifications • St Gallen molecular subtypes

Gaëlle Judes[1,2], Aslihan Dagdemir[1,2], Seher Karsli-Ceppioglu[1,2,3], André Lebert[4], Maureen Echegut[2], Marjolaine Ngollo[1,2], Yves-Jean Bignon[1,2], Frédérique Penault-Llorca[2,5] & Dominique Bernard-Gallon*[,1,2]

[1]Department of Oncogenetics, Centre Jean Perrin, CBRV, 28 Place Henri Dunant, 63001 Clermont-Ferrand, France

[2]EA 4677 'ERTICA', University of Auvergne, 63011 Clermont-Ferrand, France

[3]Department of Toxicology, Faculty of Pharmacy, Marmara University, 34668 Istanbul, Turkey

[4]University Blaise Pascal, Pascal Institute UMR 6602 CNRS/UBP, 63177 Aubière, France

[5]Department of Biopathology, Centre Jean Perrin, 63011 Clermont-Ferrand, France

*Author for correspondence: dominique.gallon-bernard@clermont.unicancer.fr

Breast cancer is the main cause of women mortality in the world [1] and comprise different tumor subtypes associated with histological criteria. Prognosis for the patient has been based on the analysis of biological markers in the tumor including ER, PR, HER2 and Ki67. Previously, transcriptomic microarrays identified several molecular subtypes of breast cancer [2,3] and showed that ER+ and ER- tumors have different molecular profiles and resulted in three distinguishable groups based on mRNA expression of specific luminal, HER2 and basal genes [4–7]. Luminal breast cancer is a heterogeneous intrinsic group that was separated into two subgroups: luminal A (ER+, PR+, CK8/18+, HER2- and low rate of Ki67) and luminal B characterized by proliferation markers (KI67 >14%) [5,8,9] and PR expression <20% or lost. However, HER2 was used to define two types of luminal B, one type is HER2- with a high Ki67 index and one type is HER2+ [10]. HER2+ tumors are ER-/PR- and showed amplification and overexpression of the *ERBB2* gene. On the other hand, basal-like group has been distinguished by the expression of basal cytokeratin CK5/6, ER- and PR-. The 13th St Gallen International Breast Cancer Conference established a classification of breast cancer subtypes [11].

While breast cancer was linked to genetic mutations, several evidences showed that abnormal epigenetic modifications existed [12]. The histone modifications have been highlighted in different tumor grades of breast cancer. In particular, low levels of H3K9ac, H3K18ac and H3K4me2 have been recovered in basal-like and HER2+ tumors [13]. Other studies have shown increases in H3K27me3 or H3K9me3 and a decrease of H4K16ac in breast cancer cell lines: MDA-MB231 (mesenchymal-like subtype) and MCF7 (luminal subtype) [14]. Epigenetic modifications seemed to be modulated in different subtypes of breast cancers. Trimethylation of H3K27 is an epigenetic

Future Medicine part of fsg

mark mediating gene silencing and can have a role in recruitment of DNA methyltransferase [15–17]. H3K27 di- and tri-methylation are characteristic of Polycomb group target genes and are associated with transcriptional repression. PRC2 is associated with chromatin condensation and is composed of several proteins, including EED, SUZ12, RbAP46/48 and either EZH1 or EZH2. EZH2 has a SET domain that functions as a histone-lysine *N*-methyltransferases [18–20]. Specific marks in lysine acetylation (H3K9ac, H3K18ac and H4K12ac) are associated with transcriptionally active gene promoters [21]. Different histone acetyltransferases are involved in cancer mechanism such as GCN5 [22], P300 [23], SRC3 [24] and Tip60 [25]. A high presence of acetylated H3K9 (H3K9ac) was found at region surrounding transcriptional start sites, this mark is considered to active chromatin. In different cancers, a decrease in H3K9ac levels was reported and the H3K9ac expression was correlated with a poor prognosis for patients [26,27]. In contrast, the function of H3K4 acetylation (H3K4ac) has been poorly studied and particularly in cancer. In *Schizosaccharomyces pombe*, Xhemalce and Kouzarides [28] demonstrated that H3K4ac plays a role in the formation of repressive heterochromatin. Moreover, in *Schizosaccharomyces cerevisiae*, H3K4ac was enriched at promoters of actively transcribed genes and located just upstream of H3K4 trimethylation. Dimethylation and trimethylation of H3K4 regulated global levels of H3K4ac at gene promoters [29].

Here, we studied by ChIP-qPCR, the H3K27me3, H3K4ac and H3K9ac modifications on *BRCA1*, *EZH2*, *P300*, *SCR3*, *ERS1*, *ERS2* and *PGR* genes which are implicated in tumor initiation and progression, between sporadic breast tumors and matched normal breast tissues. A comparative analysis of epigenetic alterations was realized between the five St Gallen molecular subtypes [11] and revealed distinct patterns of epimodifications.

Patients & methods

Patients

This study included 192 breast tumors with 192 matched normal tissues. With a consent, patients accepted to give a sample for research. Surgical tissues were stored in a Biological Resource Center (AC-2013–1882). Patients with neoadjuvant chemotherapy, hormonotherapy, radiotherapy and with family background were excluded.

Molecular subtype classification

The classification of breast tumors was realized according to the St Gallen into five groups: luminal A (ER+, PGR+, KI67 <14%, histological grade 1 or 2 and HER2-), luminal B HER2- (ER+, PGR <10% or negative, KI67 >14%, high grade and HER2-), luminal B HER2+ (ER+, PGR- and HER2+), HER2+ no luminal (ER-, PGR- and HER2+) and basal-like (ER-, PGR- and HER2-) [11].

ChIP-qPCR

ChIP was performed with an SX-8G IP-Star® Compact Automated System (Diagenode) according to [30,31]. All antibodies were purchased to Diagenode: anti-H3K27me3 (pAb-069-050), anti-H3K9ac (pAb-004-050), anti-H3K4ac (pAb-165–050) and non-immune rabbit IgG (Kch-504–250).

DNA were then analyzed by q-PCR (ABI PRISM 7900HT, Applied Biosystems) with primers and probes for *ERS1*, *ERS2*, *EZH2*, *BRCA1*, *P300, SRC3* described in [31] and *PGR* [14]. ChIP efficiency of a locus was calculated by qPCR:%(ChIP/Total Input) = 2^[(Ct(×% input) – log(×%)/log2) – Ct(ChIP)] × 100%. Ct (ChIP) and Ct (×% input) are threshold values obtained from exponential phase of qPCR for immunoprecipitate DNA sample and input sample respectively. Two is the amplification efficiency and the log(×%)/log2 compensatory factor that is used to take into account the dilution 1:× of the input. The recovery is the %(ChIP/Total Input) (Diagenode).

RT-qPCR

RNAs were extracted from the same tissues shared before ChIP-qPCR using RNA-B™ (BioProbe Systems). 1 µg of RNA was used as a template for cDNA synthesis using the First-Strand DNA Synthesis Kit (GE Healthcare Life Science). Duplex q-PCR was performed in 96-well plates with 30 ng of cDNA and 12.5 µl Gene expression Master Mix (dATP, dCTP, dGTP and dUTP, MgCl2, AmpliTaqGold, Amperase uracil-*N*-glycosylase), 200 nM assay-on-demand of each gene (*EZH2* (Hs01016789_m1), *BRCA1* (Hs01556193_m1), *ERS1* (Hs00174860_m1), *ERS2* (Hs01003531_m1), *SRC3* (Hs01105251_m1), *P300* (Hs00914223_m1), *PGR* (Hs01556702_m1); applied Biosystems. Data were collected using an ABI PRISM 7900HT. Relative gene expression was determined using the comparative C_T method. The normalization was performed with 18s RNA. The level of each mRNA gene of each tumor was normalized to mRNA of normal tissues. Fold changes were determined using the ΔΔCT method [32].

Western blotting

Frozen tissue lysates were prepared in RIPA Buffer with phosphatase inhibitor and protease inhibitor cocktail (Sigma). 25–50 µg of proteins were studied on SDS-PAGE gel, electrotransferred onto nitrocellulose

membrane and respectively probed with primary Abs: anti-EZH2 (1:1000, rabbit PoAb, pAb-039–050, Diagenode), anti-ER-α (1:1000, mouse MoAb, sc-8002, Santa Cruz), anti-PGR (1:500, rabbit MoAb, PAB12723, Abnova) or anti-β Actin (1:4000, CP01, Millipore), followed by Alkalyne Phosphatase conjugated secondary Abs (anti-rabbit IgG, 1:2000, S3738 or anti-mouse, 1:2000, S372B, Promega). Bands were detected by Western Blue® Alkalyne Phosphatase (Promega).

Statistical analysis

In order to verify the normality of the distribution of all variables, four tests were performed, including Kolmogorov–Smirnov, Cramer–von Mises, Lilliefors and Anderson–Darling. When the distribution was abnormal, a one-parameter Box-Cox transformation was used. After normalization, log-transformed ratio data distributions were gathered according to molecular subtypes. Student t-test was used to compare % (ChIP/total input) and the gene expression levels between breast cancer and normal tissues. A two-way duplicated measures analysis of variance (ANOVA) was performed with a level of 0.05. At any time when the results were significant, an HSD Tukey test was used to compare ranges between different subtypes.

To preferentially highlight the gene/mark combinations connected to a certain level of risk, additional statistical processing was performed in order to predict to which subtypes, the patients belong to in respect with the three variables (genes, marks and subtypes) and the result of the ChIP-qPCR. At least, two approaches can be used depending on the relationship between the membership of subtype, the three variables and the result of the ChIP-qPCR. For instance, if a linear relation was assumed, a linear discriminant analysis (LDA) could be used. Likewise, if the shape of the relation was unknown, decision trees would allow to use some logical connections performed on the values of the variables above to distribute the tumors according to homogeneous groups. LDA and decision trees would help to emphasize the significant variables that determine a membership in one group or another subtype. Some important variables were selected by calculating the Wilks' Lambda followed by a test of Fischer for LDA or a Gini Index greater than 3.5 for decision tree. All statistical assays were performed using the R statistical software package including: version 3.20 (R Development Core Team), the Nortest package for the normality tests [33], the package car for the Box–Cox transformation [34] and the package agricolae for the HSD Tukey test [35]. The decision tree construction was performed with the rattle packages [36] and rpart [37]. LDA was made by using MASS [38] and package rrcov [39].

Results

Global H3 modifications of genes & descriptive analysis of breast tumors

H3 marks including acetylation (H3K4ac, H3K9ac) and methylation (H3K27me3) within the promoters of *BRCA1*, *SRC3*, *P300*, *EZH2*, *ERS1*, *ERS2* and *PGR* genes were studied by ChIP-qPCR in 192 breast tumors compared with 192 matched normal tissues. For all the studied genes, Figure 1 showed a significant decrease in the tumors by comparison with the normal tissues for the H3K27me3 and H3K4ac marks. For H3k9ac marks, no significant difference was established between tumors and normal tissues for *BRCA1*, *SRC3*, *P300* and *ERS1*. In addition, H3K9ac marks of *ERS2* and *PGR* genes were significantly decreased in the tumors. At the opposite, the H3K9ac marks of *EZH2* exhibited a trend to increase in the tumors. On the whole, the global assessment of H3K4ac, H3K9ac and H3K27me3 marks for the studied genes always showed a decrease in the tumors. Therefore, further studies of H3 modifications were investigated in breast tumors classified according to St Gallen.

All subtypes presented an unsignificant variation of patient frequency according to age ($p = 0.49$), histological tumor subtype ($p = 0.139$), tumor size ($p = 0.494$) and lymph node status ($p = 0.108$ and $p = 0.186$) (Table 1). Thus, the homogeneity found in these baseline patient characteristics would prevent interferences in the subsequent epigenetic analysis. Luminal A and B (HER2-) subtypes included a more significant frequency of tumors with a low SBR grade and a status of Ki67 <14%. At the opposite, luminal B (HER2+), HER2+ and basal-like subtypes exhibited an SBR grade between 2 and 3 ($p < 0.001$) and a Ki67 status ≥20% ($p < 0.001$). Concerning receptor status, a clear distinction of subtypes is found according receptor expression with St Gallen classification.

In conclusion, the three parameters (SBR grade, Ki67 and receptor status) were all found significantly different with Pearson's Chi-square.

Specific H3 modifications in tumor subtypes

In luminal A, luminal B HER2-, luminal B HER2+, HER2+ and basal-like subtypes, different histone alterations were found. In Figure 2, just *BRCA1*, *PGR* and *P300* genes are shown because *PGR* had the same epigenetic profile as *ERS1* and *ERS2*; *P300* had the same epigenetic profile as *EZH2* and *SRC3*. In Figure 2A, in HER2+ subtype and luminal B HER2+ tumors, for *BRCA1*, a significant increase was obtained in the percentage of the H3K4ac and H3K9ac acetylated marks compared with the percentage of the H3K27me3 methylated marks for HER2+ and luminal B HER2+ subtypes. At the opposite, in HER2- subtypes (luminal

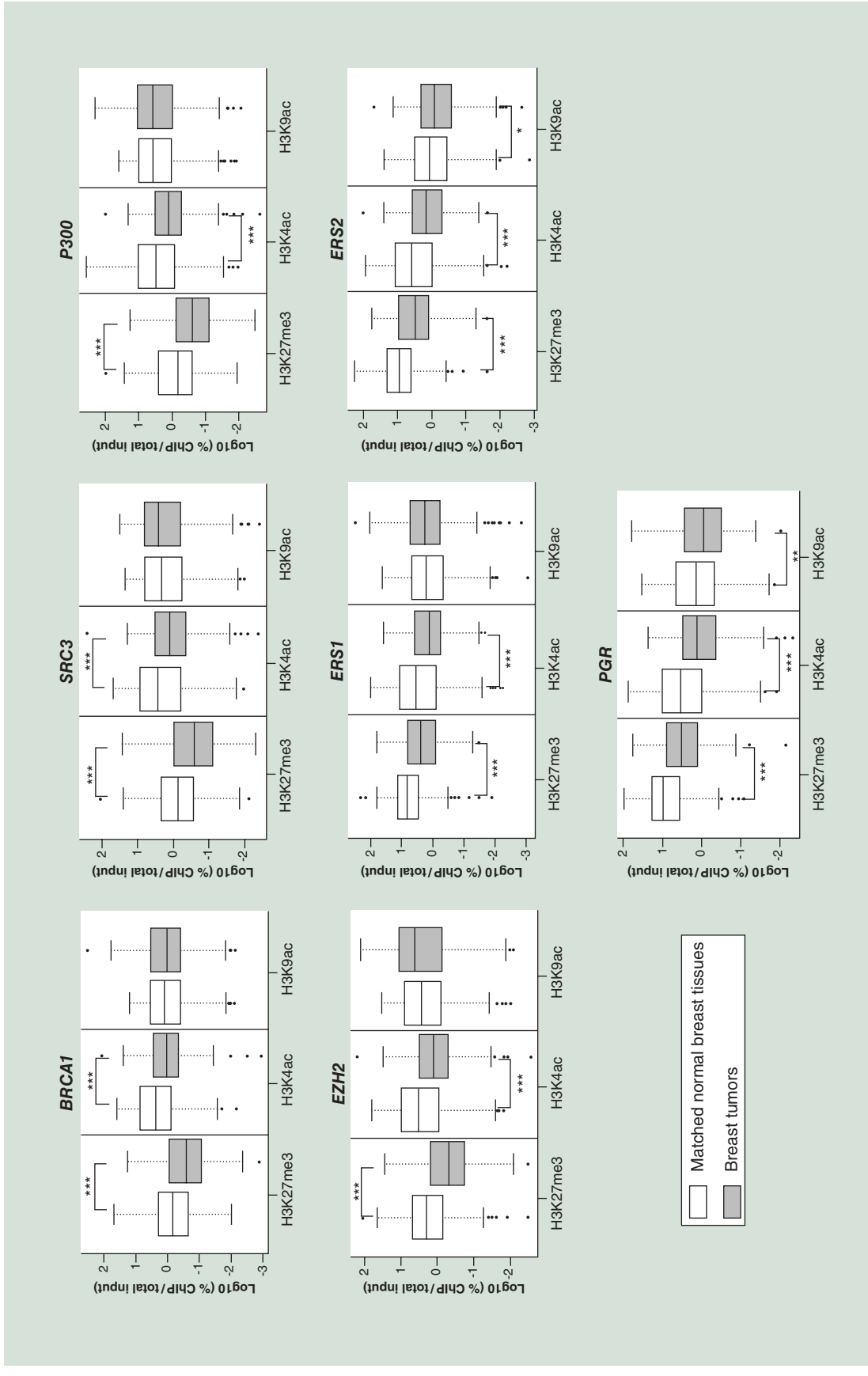


Figure 1. Levels of H3K27me3, H3K9ac and H3K4ac marks on target genes (*BRCA1, SRC3, P300, EZH2, ERS1, ERS2, PGR*) in 192 breast tumors and 192 matched normal tissues examined by ChIP-qPCR. Box plots corresponded to the three marks grouped by gene. The horizontal lines represented the median; the bottom and top of the boxes represented the 25th and 75th percentiles, respectively and the vertical bars represented the range of data. The y-axis showed the ratio of the H3 methylation or acetylation immunoprecipitation to input. Data were analyzed by paired = true Student's t-test.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 1. Baseline characteristics according to St Gallen classification of tumor subtypes.

Baseline characteristic	Luminal A, n (%)	Luminal B HER2-, n (%)	Luminal B HER2+, n (%)	HER2+, n (%)	Basal-like, n (%)	p-value†
Patients (n = 192):						
– Subtype (n)	38	38	39	39	38	
Mean (SD) age at diagnosis (years):	62.25	66.23	63.07	62.82	67.5	0.49
– <50 years	6 (15.78)	5 (13.16)	7 (17.95)	4 (10.26)	5 (13.16)	
– 50–69 years	20 (52.63)	15 (39.40)	17 (43.59)	23 (58.97)	14 (36.84)	
– ≥70 years	12 (31.58)	18 (47.36)	14 (35.90)	12 (30.77)	19 (50.00)	
Histological tumor subtypes:						
– Ductal carcinoma	28 (73.7)	29 (76.32)	33 (84.62)	37 (94.87)	31 (84.21)	0.139
– Lobular and other carcinoma	10 (26.3)	9 (23.68)	5 (12.82)	2 (5.13)	7 (18.42)	
Size of tumor (cm):						
– T1 (<2)	18 (43.37)	14 (36.84)	12 (30.77)	11 (28.21)	11 (28.95)	0.494
– T2 (≥2 to ≤5)	17 (44.74)	22 (57.89)	24 (61.54)	25 (64.10)	25 (65.79)	
– T3 (>5)	2 (5.26)	1 (2.63)	–	2 (5.13)	2 (5.26)	
– Unknown	1 (2.63)	1 (2.63)	3 (7.69)	1 (2.56)	–	
SBR grade:						
– 1	11 (28.95)	10 (26.32)	2 (5.13)	2 (5.13)	2 (5.26)	<0.001
– 2	25 (65.79)	20 (52.63)	19 (48.72)	13 (33.33)	13 (34.21)	
– 3	1 (2.63)	8 (21.05)	16 (41.03)	24 (61.54)	21 (57.89)	
– Unknown	1 (2.63)	–	1 (2.63)	–	2 (5.26)	
Lymphe node status:						
– Negative	28 (73.68)	22 (57.89)	23 (58.97)	18 (46.15)	26 (68.42)	0.108
– Positive	10 (26.32)	16 (42.11)	16 (41.03)	21 (53.85)	12 (31.58)	
– 1–3 nodes +	5 (50.0)	8 (50.0)	11 (68.75)	16 (76.19)	7 (58.33)	0.186
– ≥4 nodes +	5 (50.0)	8 (50.0)	5 (31.25)	5 (23.81)	5 (41.67)	
Ki67 status (%):						
– <14	24 (63.16)	9 (23.68)	5 (12.8)	8 (20.51)	5 (13.16)	<0.001
– 14–19	14 (38.84)	8 (21.05)	7 (17.9)	4 (10.26)	3 (7.89)	
– ≥20	–	20 (52.63)	25 (64.1)	25 (76.19)	28 (73.68)	
– Unknown	–	1 (2.63)	2 (5.1)	2 (5.13)	2 (5.26)	
Receptor status						
ER status:						
– Positive	38 (100)	38 (100)	39 (100)	–	–	
– Negative	–	–	–	39 (100)	38 (100)	
PGR status:						
– Positive	38 (100)	23 (60.53)	24 (61.54)	–	–	
– Negative	–	15 (39.47)	13 (33.33)	39 (100)	38 (100)	
HER2 status:						
– Positive	–	–	39 (100)	39 (100)	–	
– Negative	38 (100)	38 (100)	–	–	38 (100)	

Data taken from Scarff Bloom Richardson.
†Pearson's Chi-square test; n = patient number.
ER: Estrogen receptor; HER2: Human epidermal growth factor receptor-2; PGR: Progesterone receptor; SD: Standard deviation.

A, luminal B HER2- and basal-like), the H3 modifications of *BRCA1* showed no significant difference between H3K4ac, H3K9ac and H3K27me3 marks. In addition, basal-like tumors just exhibited a significant difference between the H3K4ac and H3K27me3.

In tumors, concerning hormonal receptors *ERS1*, *PGR* (Figure 2B) and *ERS2*, a significant decrease with acetylated marks and a significant increase with H3K27me3 marks were found in less aggressive breast cancer subtypes (luminal A and B HER2-) compared with more aggressive cancer (luminal B HER2+, HER2+ and basal-like) where the differences were found insignificant.

The status of histone modifications of *EZH2* which codes for an HMT and *P300*, *SRC3* which code for an HAT were investigated in tumors. The

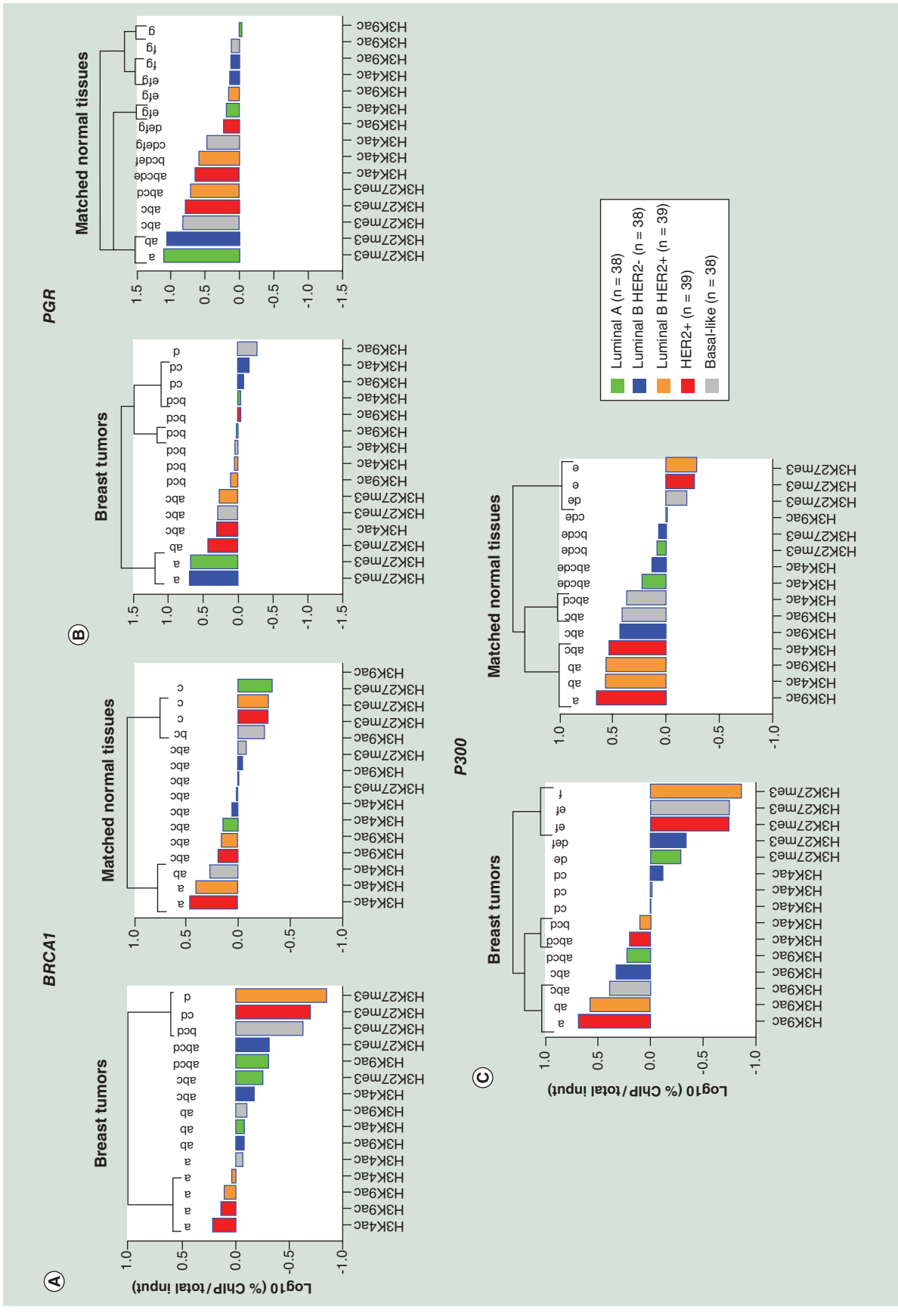


Figure 2. Epigenetic signatures of breast cancer subtypes. H3K27me3, H3K9ac and H3K4ac modifications on target genes (*BRCA1* **[A]**, *PGR* **[B]** and *P300* **[C]**) in five St Gallen molecular subtypes compared with the matched normal tissues. ChIP and Q-PCR were used to study histone H3 patterns of the studied genes. y-axis corresponded to log10% (ChIP/total input) in gene promoters of studied epigenetic marks. An analysis of variance, followed by a Tukey multiple comparison test were used to compared different breast cancer subtypes. The statistical significant between subtypes were indicated by letters 'a', 'b', 'c', 'd', 'e', 'f' and 'g'.

level of acetylated marks was found increased and the H3K27me3 marks decreased for *EZH2*, *P300* (Figure 2C), *SRC3* genes in luminal B HER2+, HER2+ and basal-like, these alterations might contribute to change of gene expression. These observations demonstrated clearly differences in epigenetic patterns in the promoters of genes implicated in the different molecular subtypes.

In this study, histone epimodifications were also investigated by ANOVA analysis in the matched normal tissues corresponding to the tumors classified according to St Gallen. Especially, in Figure 2A, for *BRCA1* gene, luminal B HER2+ and HER2+ were distinguished to other groups by the significant difference corresponding to the increase of H3K4ac and the decrease of H3K27me3.

PGR in Figure 2B, acted like other hormonal receptors (*ERS1* and *ERS2*), a trend toward increase of H3K4ac and H3K27me3 was shown in normal tissues corresponding to luminal B HER2+, HER2+ and basal-like tumors. For luminal A and luminal B HER2-, *ERS1*, *ERS2* and *PGR* showed a similar profile in the normal tissue and the tumor, that is, an increase of H3K27me3 marks and a decrease of H3K9ac marks, and the difference were significant.

P300 (Figure 2C) and *SRC3* demonstrated in luminal B HER2+, HER2+ and basal-like subtypes, a significant difference between the increase of H3K4ac and the decrease of H3K27me3 marks in the normal tissues and the tumors.

Transcriptional changes between tumors & normal tissues

To underline the histone alterations on the gene expression, RT-qPCR was used to follow the mRNA of the panel of genes in breast tumors and matched tissues in each molecular subtypes (Figure 3). Log10 (fold change) was used with reference set to 0, if the value is positive, the gene is overexpressed, and if the value is negative, the gene is underexpressed in the tumors.

Luminal A, luminal B (HER2-) and luminal B (HER2+) tumors showed significant overexpression of hormonal receptor *ERS1* compared with normal tissues ($p = 0.014$, $p = 0.012$ and $p = 0.018$). Moreover, for *PGR* expression, a significant increase ($p = 0.03$) for luminal A subtype compared with normal tissues was observed and just a trend toward an increase in luminal B (HER2-) and luminal B (HER2+). HER2+ and basal-like tumor subtypes showed that *PGR* (HER2+: $p = 0.003$; basal-like: $p = 0.007$), *ERS2* (HER2+: $p = 0.0001$; basal-like: $p = 0.04$) and *ERS1* (HER2+: $p = 0.002$; basal-like: $p = 0.0002$), were significantly underexpressed compared with normal tissues. Most of the tumors in the different subtypes presented a significant overexpression of *EZH2* mRNA in luminal A ($p = 0.03$), luminal B (HER2-; $p = 0.002$), luminal B

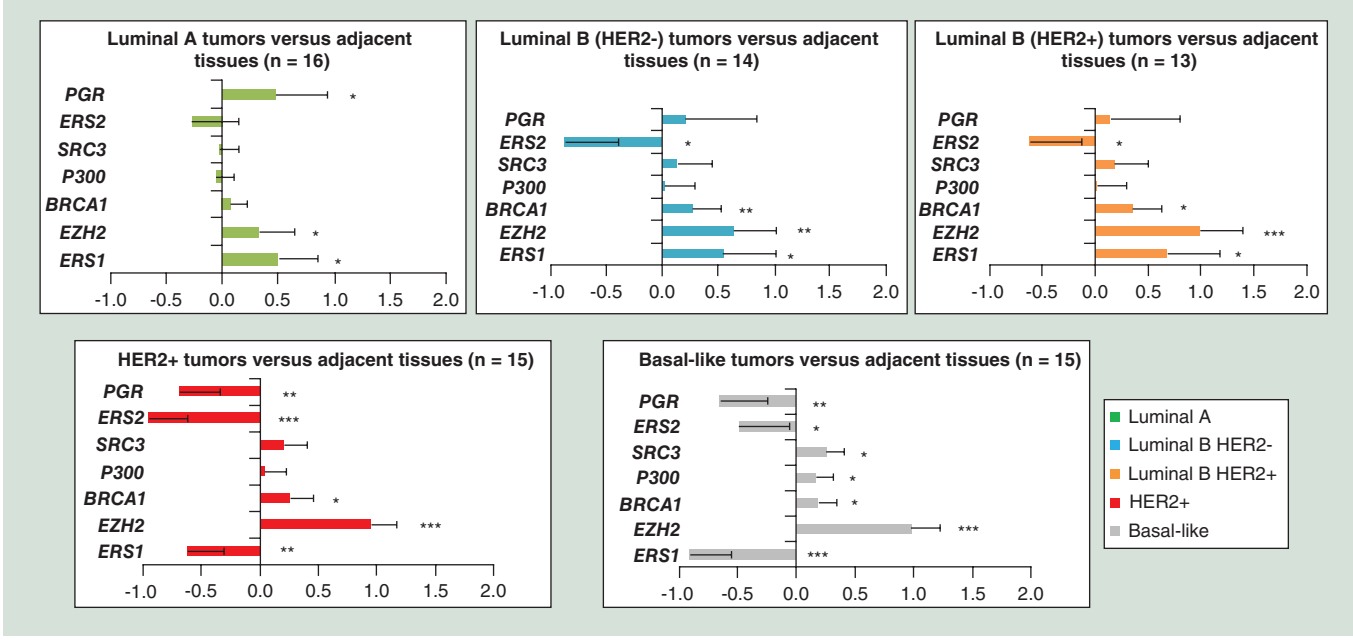


Figure 3. Relative mRNA expressions of *PGR*, *ERS2*, *SRC3*, *P300*, *BRCA1*, *EZH2* and *ERS1* in luminal A, luminal B (HER2-), luminal B (HER2+), HER2+, basal-like tumors compared with normal tissue. Graphs showed fold changes in tumor gene expressions normalized to normal tissue gene expression. x-axis represented the relative fold changes in log10 scale. Log10 (fold change) was used with reference set to 0. Data were expressed as mean for each group and error bars indicate the 95% CI of the mean value; data were analyzed by paired = true student's t-test.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

(HER2+) ($p = 0.0007$), HER2+ ($p = 3.36 \times 10^{-6}$), basal-like ($p = 1.73 \times 10^{-6}$) that seemed to increase with the aggressiveness.

In luminal A, luminal B (HER2-), luminal B (HER2+) and HER2+ subtypes, the relative *SRC3* and *P300* mRNA expressions were found closed to 0, exhibiting no difference between the tumors and the normal tissues. Only, basal-like subtype showed a significant relative mRNA overexpression of *SRC3* ($p = 0.007$) and *P300* ($p = 0.04$).

Concerning *BRCA1*, a significant mRNA overexpression was found in luminal B (HER2-; $p = 0.006$), luminal B (HER2+; $p = 0.02$), HER2+ ($p = 0.01$) and basal-like ($p = 0.04$) tumors compared with normal tissues.

Furthermore, ANOVA test was only used in tumors to identify differences in mRNA expressions of *BRCA1*, *ERS2*, *PGR*, *ERS1*, *EZH2*, *P300* and *SRC3* between the molecular subtypes. In Figure 4A, the statistical analyses allowed to identify three major genes *EZH2*, *ERS1* and *PGR* inside the molecular subtypes. *EZH2* mRNA significantly increased in basal-like (a) compared with luminal A (b) subtypes. Besides, HER2+ (ab), luminal B (HER2+) (ab), luminal B (HER2-) (ab) subtypes were identical. Hormonal receptor (*ERS1* and *PGR*) expressions significantly decreased in negative ERS1 and PGR subtypes (HER2+) (b; bc) and basal-like (b; c) compared with luminal A, B (HER2-) and B (HER2+) subtypes.

The changes found in mRNAs for *EZH2*, *ERS1* and *PGR* were checked by WB (Figure 4B). Protein levels of EZH2 were higher in tumors compared with normal tissues in the five breast cancer subtypes. HER2+ and basal-like subtypes were associated to high expression of EZH2 in tumors. Then, ER-α and PGR proteins were highly expressed in luminal A, luminal B (HER2-) and luminal B (HER2+) tumors while HER2+ and basal-like tumors expressed lower ER-α and PGR proteins compared with normal breast tissues.

Thus, protein and mRNA levels followed a similar way for each gene previously selected by ANOVA analysis.

Classification of five molecular subtypes & epigenetic model assessment

The histone H3 signatures on seven gene promoters were found different in five molecular subtypes; however, these data were not sufficient to stratify the different tumors based on epigenetic data. It is crucial to

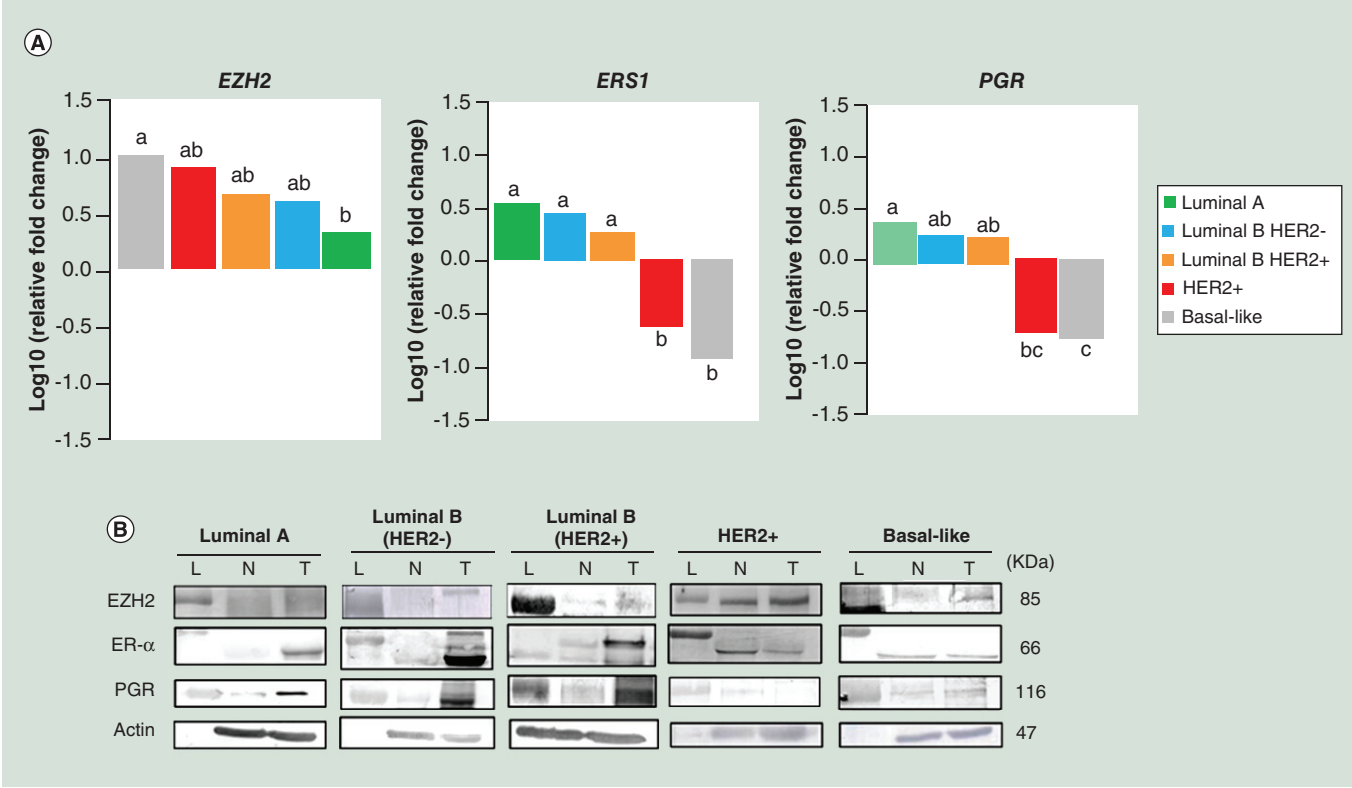


Figure 4. Expression of *EZH2*, *ERS1* and *PGR* according breast cancer subtypes. (A) The difference of relative mRNA levels of *EZH2*, *ERS1* and *PGR* in breast cancer tumors normalized to normal tissue gene expression inside the five molecular subtypes. The analysis was assessed by the two-way ANOVA test. The statistical significant between the five molecular subtypes was indicated by letters 'a', 'b', 'c' and 'd'. **(B)** Western-blot demonstrating protein expression of EZH2, ER-α and PGR in the molecular subtypes.
L: Ladder; N: Normal tissue; T: Tumor.

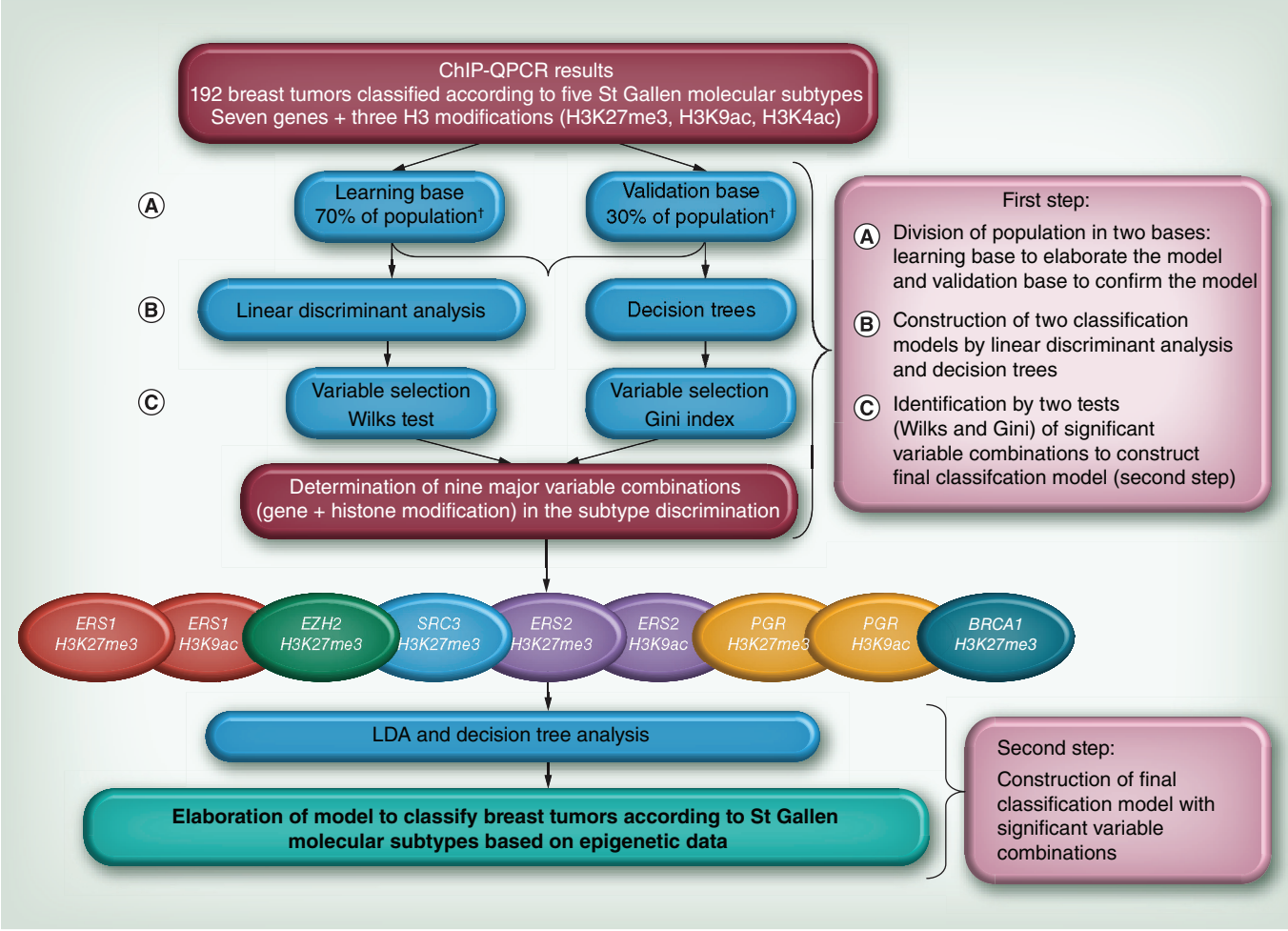


Figure 5. Statistical approaches to elaborate an epigenetic model of breast cancer subtypes.
Luminal A subtype was represented by 19.79% of total tumors, luminal B (HER2-) 19.79%, luminal B (HER2+) 20.31%, HER2+ 20.31% and basal-like 19.79%. The first step was divided into 3 parts **(A)**, **(B)** and **(C)** to elaborate the final classification.
†Inside each subtype, the ratio 70/30 was respected.
LDA: Linear discriminant analysis.

determine which variables among the epigenetic marks and genes might be the most efficient ways to differentiate the subtypes. For that reason, two supervised techniques (decision tree and LDA) were compared with help choose the best classification model. Decision tree model that evaluated mark/gene combinations among the seven genes and three histone modifications allowed a better differentiation of the molecular subtypes. In parallel, LDA was used to indicate whether there was a linear relationship between these subtypes and different existing studied variables.

In the first time, 20 random draws among 192 tumors were performed to determine both a learning basis represented by 70% of the tumor population and a validation basis represented by 30% of the tumor population (Figure 5). These random draws revealed that LDA and decision tree analysis conferred a similar building capacity and an equivalent validation capacity for elaborated models. These models classified tumors in different subtypes with a shifting effectiveness. Basically, four efficacy categories were distinguished, including good prediction (meaning 100% of tumors correctly classified), tumors classified with ±1 subtype, pejorative prediction (low-risk tumors classified by mistake in breast cancer with a high risk, e.g., luminal A tumors classified into HER2+) and risky prediction (high-risk tumors classified by mistake in breast cancer with a low risk, e.g., basal-like tumor classified into luminal A).

Two statistical tests were performed among 20 random draws to determine mark/gene combinations which act in the breast subtypes. The Gini index was used to forest analysis and Wilks' lambda followed by an F test was calculated for LDA ($p < 0.05$; Gini index ≥3.5). These two tests highlighted nine gene/mark combinations: *ERS1*/H3K27me3, *ERS1*/H3K9ac,

EZH2/H3K27me3,*SRC3*/H3K27me3,*ERS2*/H3K27me3, *ERS2*/H3K9ac, *PGR*/H3K27me3, *PGR*/H3K9ac and *BRCA1*/H3K27me. Then, these variable combinations were used to construct a multivariable decision tree and LDA. Minimal loss information was noticed following the selection of nine gene/mark couples. Figure 6 corresponded to the capacities of learning and of validation in the analysis of the LDA and the decision tree, developed according to nine gene/histone mark combinations. In the model of elaboration, the LDA analysis accurately predicted 52% (70 of 135 tumors) and 83% (with an acceptable margin of error). Regarding the decision tree analysis, we obtained similar results with the learning basis, which nicely classified 62% of tumors in five subtypes (84 of 135) and 84% with a margin of error. Results from the decision tree and LDA analysis were identical with an existence of a linear relationship between tumors subtypes and mark/gene combination. Figure 7 displays a 3D LDA representation of final model. It was observed fundamental discrepancies of H3K27me3 and H3K9ac modifications on *ERS1*, *ERS2*, *PGR*, *EZH2* and *BRCA1* promoters between luminal A, luminal B (HER2-) and luminal B (HER2+), HER2+ and basal-like tumor subtypes. The final decision tree had fourteen tiers of branching using seven categorical variable couples that resulted in sixteen terminal nodes (Figure 8).

Discussion

We have characterized epigenetic alterations of H3 (H3K27me3, H3K9ac and H3K4ac) associated with molecular subtypes defined by St Gallen classification: luminal A, luminal B (HER2-), luminal B (HER2+), HER2+, basal-like. We found that gene promoters (*BRCA1*, *ERS2*, *PGR*, *ERS1*, *EZH2*, *P300* and *SRC3*) presented a particular epigenetic profile according to molecular subtypes. Indeed, among five molecular subtypes, we characterized two major epigenetic profiles; the first profile brought luminal A and luminal B (HER2-) subtypes together, corresponding to subtypes

Learning base

Subtypes	Predicted: Lum A	Lum B HER2-	Lum B HER2+	HER2+	Basal-like	Actual total
Lum A	15	7	2	1	2	27
Lum B HER2-	6	15	3	2	1	27
Actual Lum B HER2+	2	3	16	3	3	27
HER2+	2	0	6	11	8	27
Basal-like	1	4	3	6	13	27
Predicted total	26	29	30	23	27	135

Categories	n	%		
100%	70	52%	83%	91%
± 1 subtype	42	31%		
Pejorative prediction	11	8%		
Unsafe prediction	12	9%		

Learning base

Subtypes	Predicted: Lum A	Lum B HER2-	Lum B HER2+	HER2+	Basal-like	Actual total
Lum A	18	4	0	1	4	27
Lum B HER2-	6	14	2	3	2	27
Actual Lum B HER2+	3	2	19	2	1	27
HER2+	1	2	4	18	2	27
Basal-like	2	1	1	8	15	27
Predicted total	30	23	26	32	24	135

Categories	n	%		
100%	84	62%	84%	93%
± 1 subtype	30	22%		
Pejorative prediction	11	8%		
Unsafe prediction	10	7%		

Validation base

Subtypes	Predicted: Lum A	Lum B HER2-	Lum B HER2+	HER2+	Basal-like	Actual total
Lum A	6	2	1	1	1	11
Lum B HER2-	4	3	2	1	1	11
Actual Lum B HER2+	1	4	4	0	3	12
HER2+	2	2	0	6	2	12
Basal-like	2	0	0	3	6	11
Predicted total	15	11	7	11	13	57

Categories	n	%		
100%	25	44%	74%	88%
± 1 subtype	17	30%		
Pejorative prediction	8	14%		
Unsafe prediction	7	12%		

Validation base

Subtypes	Predicted: Lum A	Lum B HER2-	Lum B HER2+	HER2+	Basal-like	Actual total
Lum A	6	0	2	3	0	11
Lum B HER2-	6	1	2	1	1	11
Actual Lum B HER2+	2	4	2	3	1	12
HER2+	0	2	0	8	2	12
Basal-like	2	1	0	6	2	11
Predicted total	16	8	6	21	6	57

Categories	n	%		
100%	19	33%	74%	88%
± 1 subtype	23	40%		
Pejorative prediction	8	14%		
Unsafe prediction	7	12%		

Figure 6. Assessment of linear discriminant analysis and decision tree models. The classification efficiency was evaluated for learning and validation bases in two analyses. Number of tumors (n) was classified into four categories (100%, ±1 subtype, pejorative prediction and unsafe prediction) and the proportion of tumors for each category was evaluated.

with a 'low risk' and the second profile gathered luminal B (HER2+), HER2+ and basal-like tumors corresponding to subtypes with 'high risk'. Luminal A and B (HER2-) tend to have a good prognosis which can be explain by gene expression and clinical feature predictors like ER positive, low% of Ki67, negative lymph node and low grade [40]. However, most aggressive behaviors, including invasion, are typical characteristics of HER2+ and basal-like tumors [41].

Moreover, among the seven genes, three groups of genes were identified and showed a similar epigenetic signature; hormonal receptor genes (*ERS1*, *ERS2* and *PGR*), tumor suppressor gene (*BRCA1*) and histone-modifying enzymes including histone methyltransferase and histone acetyltransferase (*EZH2*, *P300* and *SRC3*). Hormonal receptor gene promoters (*ERS1*, *ERS2* and *PGR*) displayed on the whole low level of acetylated marks (H3K9ac and H3K4ac) but high proportion of H3K27me3 particularly in 'low-risk' subtypes compared with 'high-risk' subtypes. In the 'high-risk' subtypes, the loss of *ERS1* and *PGR* expression is caused by a hypermethylation of the CpG islands within promoter [42,43]. Kondo *et al.* established that histone H3K27me3 silenced genes independently of promoter DNA methylation which is characterized by H3K4 and H3K9 methylation [44]. These observations could explain the downregulation of H3K27me3 on promoters of *ERS1*, *ERS2* and *PGR* which are silenced in HER2+ and basal-like subtypes. Besides, in MCF7 cells, which represented luminal breast cancer subtype, *ERS1* promoter showed an increase of H3K4me3, a mark activating transcription and a decrease of H3K9me2 a transcription repressive mark [45]. These epigenetic alterations participate in an overexpression of ER in luminal breast cancer. The acetylated marks, in this study, do not seem to contribute to the overexpression of hormone receptor in luminal A and B (HER2-) subtypes.

Another parameter, BRCA1, is important for ER expression and function. *BRCA1* is a tumor suppressor; a dysregulation and reduced expression of *BRCA1* can play a role in sporadic breast cancer development [46,47]. In ER positive breast cancer, estrogen is a known activator of *BRCA1* [48,49]. The increased BRCA1 transcription is associated with loss of HDAC1 from the BRCA1 promoter and an increase in histone H3 and H4 acetylations [50]. We found that *BRCA1* promoter acquired a specific modulation of H3 modifications in luminal B (HER2+), HER2+ and basal-like subtypes. These subtypes were characterized by a global reduction of H3K27me3 compared with acetylated marks. Moreover, Matros *et al.* observed that the majority of sporadic basal-like tumors were not methylated on *BRCA1* promoter and had high levels of BRCA1 protein [51]. Methylation of *BRCA1* was negatively associated to HER2 expression in breast cancer [52]. The histone modification profile established in this study and the variation of DNA methylation in *BRCA1* promoter could contribute to the BRCA1 expression in luminal B (HER2+), HER2+ and basal-like breast cancers.

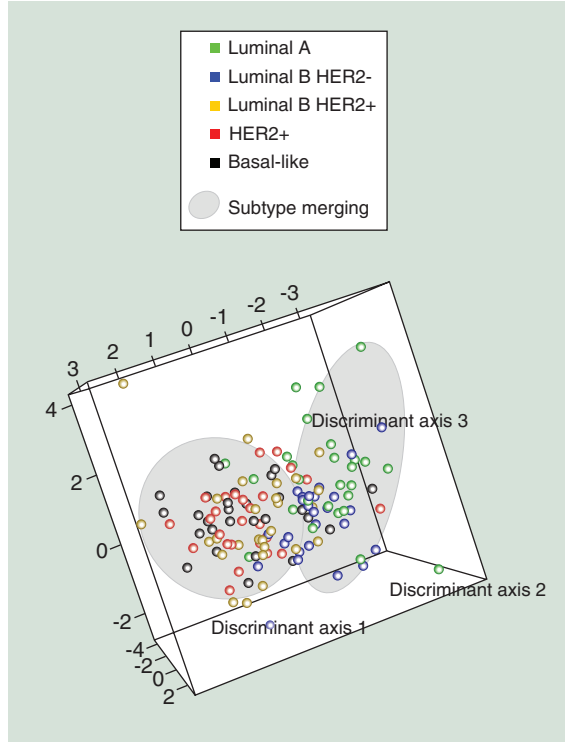


Figure 7. Linear discriminant analysis graphically depicting of luminal A, luminal B (HER2-), B (HER2+), HER2+ and basal-like subtypes according to nine gene/mark combinations (*ERS1*/H3K27me3, *ERS1*/H3K9ac, *EZH2*/H3K27me3, *SRC3*/H3K27me3, *ERS2*/H3K27me3, *ERS2*/H3K9ac, *PGR*/H3K27me3, *PGR*/H3K9ac, *BRCA1*/H3K27me3).

Another group of genes was identified by H3 profile and included *EZH2*, *P300* and *SRC3* genes. EZH2 is a histone methyltransferase; P300 and SRC3 are histone acetyltransferases. The promoter of these three genes displayed a high level of H3K4ac and H3K9ac compared with H3K27me3 in luminal B (HER2+), HER2+ and basal-like. EZH2 expression was strongly correlated with increased tumor cell proliferation and it was also associated with multiple features of aggressive breast cancer, such as HER2 expression, and markers of the basal-like subtype [53]. In fact, high abundance of EZH2 in basal-like, TNBC and HER2-enriched tumors was reported and associated with poor distant disease-free survival [54]. However, a high expression of H3K27me3 was found in luminal A tumors and was associated with better survival [54,55]. So, high expression of EZH2 does not seem correlated with increase of H3K27me3. A study showed that p300 was expressed

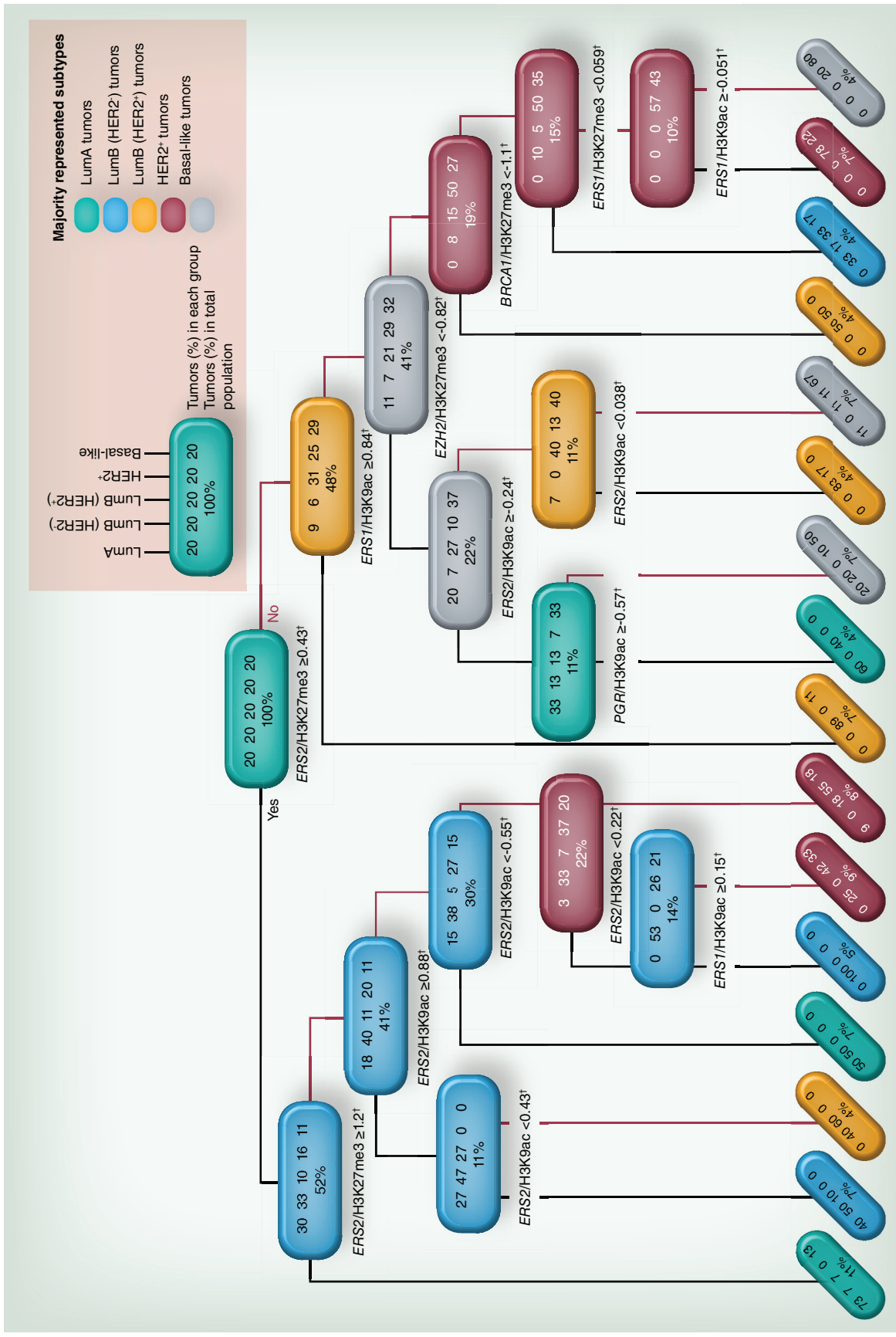


Figure 8. Decision tree classification of the five St Gallen molecular subtypes of breast cancers.
†The boxes show the decision nodes with the% (ChIP/total input) for each gene/mark combinations, and the repartition of the number tumors for each subtypes (%). These data form the splitting rules.

in triple negative cells and seems to play a critical role in invasive growth and inhibition by acetyltransferase inhibitor such as L002 that is a potential anticancer therapeutic [56]. SRC-3 expression was significantly correlated with p300/CBP [57] and it existed a combined function as well as a functional link for these proteins [58]. SRC3 was overexpressed in high grade tumor and in HER2+ tumors [59]. Therefore, *EZH2*, *P300* and *SRC3* were implicated in the development of HER2+ and basal-like breast cancer and their particular H3 signature could be implicated in these subtypes.

Moreover, we have found in normal tissues a specific profile of H3 modifications following the repartition in the different molecular subtypes of breast tumors. We have found in normal tissues a specific profile of H3 modifications following the repartition in the molecular breast subtypes. However, Yan *et al.* studied DNA methylation in adjacent breast tissues located at 1, 2, 3 and 4 cm away from the visible tumor boundary and showed that DNA methylation density progressively radiating from the tumor epicenter to the surrounding tissues. They also demonstrated that the methylation of adjacent tissues is significantly higher than in normal breast samples belonging to healthy individuals [60].

In this study, we have shown that epigenomic and transcriptional data in breast cancer reveal that H3 modifications affects gene expression and are specific to breast cancer molecular subtypes. A close relationship exists between histone modifications and gene transcription. The transcriptional regulation is a complex and dynamic process for which histone modifications are a key component. However, the link between these two regulation processes is very complex. Some combinations of histone marks are indicative of transcriptional states. Bernstein *et al.* showed 'bivalent domain' of histone modifications such as harboring both the repressive mark H3K27me3 and the active mark H3K4me3 near genes with poised transcription [61]. When H3K27me3 becomes weaker and H3K4me3 becomes stronger, the gene transcription will be activated. Thus, a high number of histone marks and their intercombinations could influence gene transcription [62] and the combinations of histone marks are indicative of transcriptional states.

Different treatments which act on histone mark modulation exist and have a potential impact in breast cancer treatment. For example, 3-deazaneplanocin A (DZNep) was reported to selectively inhibit trimethylation of lysine 27 on histone H3 (H3K27me3) and lysine 20 on histone H4 (H4K20me3) as well as reactivate silenced genes in cancer cells [63]. Suberoylanilide hydroxamic acid inhibits the activity of histone deacetylases (HDACs), including human class I and class II HDACs. HDACs have many protein targets (histones and nonhistone proteins) and alter their structures and functions by acetylation [64]. The study of histone modifications and their stratification between breast cancer subtypes and the existence of treatments could be promising tools for breast cancer therapy.

The goal of our study was to identify H3 patterns of breast cancer subtypes. Two models, LDA and decision tree, were elaborated. Nine gene/histone modification combinations were identified. Using these combinations, we can predict and determine the subtype for each tumor based on two models with 83% (LDA) and 84% (decision tree) of success with a low margin of error (±1 subtype).

Conclusion

In this study, three H3 modifications were analyzed on promoters of a gene panel in the five St Gallen molecular subtypes. Specific H3 signatures were highlighted according to genes and the aggressiveness of breast cancer subtypes and were used to elaborate two statistical models to classify breast cancer subtypes. Our results added novel information to understand the heterogeneity of the breast cancer molecular subtypes.

Future perspective

There is a high degree of diversity between and within breast tumors including a histological and molecular heterogeneity, and a variation of clinical outcome for a long time. Over the past few years, molecular classifications were elaborated by a comprehensive gene expression profiling and it allowed to develop individualized therapies in breast cancer treatment. Considering the importance of this heterogeneity, it is necessary to define new biomarkers to improve the breast cancer classification. A variety of epigenetic changes has been implicated in the development of breast cancer and these changes can be used to improve the prognosis of breast cancer. Now, the challenge is to extend the study of these changes to all epigenome by targeting other epigenetic modifications. New high-throughput "Omics" technologies such as genomics and epigenomics allow to study tumors with further analysis. In the future, the goal would be to establish clinical epigenetic therapy adapted to breast cancer subtype.

Acknowledgements

The authors thank N Gomez for corrections to the English language in the manuscript. The authors also thank 'la ligue contre le cancer-Comité du Puy-de-Dôme.'

Financial & competing interests disclosure

A Dagdemir was funded by Protema Saglik Hizm, Anonymous Society (Istanbul, Turkey) and S Karsli-Ceppioglu were

supported by The Scientific and Technology Research Council of Turkey (TUBITAK-2219) project grants. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No funded writing assistance was utilized in the production of this manuscript.

Executive summary

- Breast cancer is a heterogeneous disease. The variability, due to the numerous subtypes, and the biological, histological and molecular characteristics, induce different responses with therapy.
- New treatments targeting epigenetics: DNA methylation and histone deacetylation are in development for the treatment of cancers. Epigenetic changes can then be promising biomarkers for this disease.
- Molecular breast cancer subtypes presented different epigenetic profiles. Specific signatures of H3K27me3, H3K9ac and H3K4ac marks were identified according to aggressiveness in breast cancer.
- Statistical models were elaborated in the different breast cancer subtypes depending on epigenetic data.
- The study of epigenetic modifications in breast cancer helps to improve understanding of the mechanisms of carcinogenesis and the heterogeneity of breast tumors.

References

1 Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin.* 61(2), 69–90 (2011).

2 Perou CM, Sorlie T, Eisen MB *et al.* Molecular portraits of human breast tumours. *Nature* 406(6797), 747–752 (2000).

3 Lehmann BD, Bauer JA, Chen X *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* 121(7), 2750–2767 (2011).

4 Chung CH, Bernard PS, Perou CM. Molecular portraits and the family tree of cancer. *Nat. Genet.* 32(Suppl.), 533–540 (2002).

5 Sorlie T, Perou CM, Tibshirani R *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* 98(19), 10869–10874 (2001).

6 Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet* 378(9805), 1812–1823 (2011).

7 Bauer K, Parise C, Caggiano V. Use of ER/PR/HER2 subtypes in conjunction with the 2007 St Gallen Consensus Statement for early breast cancer. *BMC Cancer* 10, 228 (2010).

8 Ades F, Zardavas D, Bozovic-Spasojevic I *et al.* Luminal B breast cancer: molecular characterization, clinical management, and future perspectives. *J. Clin. Oncol.* 32(25), 2794–2803 (2014).

9 Cheang MC, Chia SK, Voduc D *et al.* Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J. Natl Cancer Inst.* 101(10), 736–750 (2009).

10 Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ. Strategies for subtypes – dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann. Oncol.* 22(8), 1736–1747 (2011).

11 Goldhirsch A, Winer EP, Coates AS *et al.* Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann. Oncol.* 24(9), 2206–2223 (2013).

12 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 144(5), 646–674 (2011).

13 Elsheikh SE, Green AR, Rakha EA *et al.* Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Res.* 69(9), 3802–3809 (2009).

14 Li Y, Li S, Chen J *et al.* Comparative epigenetic analyses reveal distinct patterns of oncogenic pathways activation in breast cancer subtypes. *Hum. Mol. Genet.* 23(20), 5378–5393 (2014).

15 Schlesinger Y, Straussman R, Keshet I *et al.* Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer. *Nat. Genet.* 39(2), 232–236 (2007).

16 Ohm JE, Mcgarvey KM, Yu X *et al.* A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat. Genet.* 39(2), 237–242 (2007).

17 Vire E, Brenner C, Deplus R *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439(7078), 871–874 (2006).

18 Schwartz YB, Pirrotta V. Polycomb silencing mechanisms and the management of genomic programmes. *Nat. Rev. Genet.* 8(1), 9–22 (2007).

19 Margueron R, Reinberg D. The Polycomb complex *PRC2* and its mark in life. *Nature* 469(7330), 343–349 (2011).

20 Cao R, Wang L, Wang H *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298(5595), 1039–1043 (2002).

21 Bernstein BE, Kamal M, Lindblad-Toh K *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120(2), 169–181 (2005).

22 Roth SY, Denu JM, Allis CD. Histone acetyltransferases. *Annu. Rev. Biochem.* 70, 81–120 (2001).

23 Iyer NG, Ozdag H, Caldas C. p300/CBP and cancer. *Oncogene* 23(24), 4225–4231 (2004).

24 Xu J, Wu RC, O'malley BW. Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat. Rev. Cancer* 9(9), 615–630 (2009).

25 Judes G, Rifai K, Ngollo M *et al.* A bivalent role of TIP60 histone acetyl transferase in human cancer. *Epigenomics* 7(8), 1351–1363 (2015).

26 Zhen L, Gui-Lan L, Ping Y, Jin H, Ya-Li W. The expression of H3K9Ac, H3K14Ac, and H4K20TriMe in epithelial ovarian tumors and the clinical significance. *Int. J. Gynecol. Cancer* 20(1), 82–86 (2010).

27 Mohamed MA, Greif PA, Diamond J *et al.* Epigenetic events, remodelling enzymes and their relationship to chromatin organization in prostatic intraepithelial neoplasia and prostatic adenocarcinoma. *BJU Int.* 99(4), 908–915 (2007).

28 Xhemalce B, Kouzarides T. A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. *Genes Dev.* 24(7), 647–652 (2010).

29 Guillemette B, Drogaris P, Lin HH *et al.* H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. *PLoS Genet.* 7(3), e1001354 (2011).

30 Ngollo M, Lebert A, Dagdemir A *et al.* The association between Histone 3 Lysine 27 Trimethylation (H3K27me3) and prostate cancer: relationship with clinicopathological parameters. *BMC Cancer* 14(1), 994 (2014).

31 Dagdemir A, Durif J, Ngollo M, Bignon YJ, Bernard-Gallon D. Histone lysine trimethylation or acetylation can be modulated by phytoestrogen, estrogen or anti-HDAC in breast cancer cell lines. *Epigenomics* 5(1), 51–63 (2013).

32 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25(4), 402–408 (2001).

33 Gross J, Ligges U. Nortest: Test for Normality. R. Package version 1.0–3. (2015). https://cran.r-project.org

34 Fox J, Weisberg S. *An R Companion to Applied Regression, (Second Edition).* SAGE publications, Inc., CA, USA (2011).

35 De Mendiburu F. Agricolae: Statistical Procedures for Agricultural Research. R package version 1.2 –1. (2014). https://cran.r-project.org

36 Williams GJ. Dtat Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery, Use R! (2011). https://mineriaddatos.wikispaces.com

37 Therneau T, Atkinson B., Ripley B. rpart: Recursive Partitioning and Regression Trees. R package version 4.1–9. (2015). https://cran.r-project.org

38 Venables WN, Ripley BD. *Modern Apllied Statistics with S. (Fourth Edition)*, Springer, NY, USA (2002).

39 Todorov V, Filzmoser P. An object-oriented framework for robust multivariate analysis. *J. Stat. Softw.* 32(3), 1–47 (2009).

40 Liu Z, Zhang XS, Zhang S. Breast tumor subgroups reveal diverse clinical prognostic power. *Sci. Rep.* 4, 4002 (2014).

41 Schnitt SJ. Molecular biology of breast tumor progression: a view from the other side. *Int. J. Surg. Pathol.* 18(Suppl. 3), S170–S173 (2010).

42 Giacinti L, Claudio PP, Lopez M, Giordano A. Epigenetic information and estrogen receptor alpha expression in breast cancer. *Oncologist* 11(1), 1–8 (2006).

43 Gaudet MM, Campan M, Figueroa JD *et al.* DNA hypermethylation of ESR1 and PGR in breast cancer: pathologic and epidemiologic associations. *Cancer Epidemiol. Biomarkers Prev.* 18(11), 3036–3043 (2009).

44 Kondo Y. Epigenetic cross-talk between DNA methylation and histone modifications in human cancers. *Yonsei Med. J.* 50(4), 455–463 (2009).

45 Perillo B, Ombra MN, Bertoni A *et al.* DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science* 319(5860), 202–206 (2008).

46 Turner NC, Reis-Filho JS, Russell AM *et al.* BRCA1 dysfunction in sporadic basal-like breast cancer. *Oncogene* 26(14), 2126–2132 (2007).

47 Hosey AM, Gorski JJ, Murray MM *et al.* Molecular basis for estrogen receptor alpha deficiency in *BRCA1*-linked breast cancer. *J. Natl Cancer Inst.* 99(22), 1683–1694 (2007).

48 Spillman MA, Bowcock AM. BRCA1 and BRCA2 mRNA levels are coordinately elevated in human breast cancer cells in response to estrogen. *Oncogene* 13(8), 1639–1645 (1996).

49 Marks JR, Huper G, Vaughn JP *et al.* BRCA1 expression is not directly responsive to estrogen. *Oncogene* 14(1), 115–121 (1997).

50 Di LJ, Fernandez AG, De Siervi A, Longo DL, Gardner K. Transcriptional regulation of BRCA1 expression by a metabolic switch. *Nat. Struct. Mol. Biol.* 17(12), 1406–1413 (2010).

51 Matros E, Wang ZC, Lodeiro G, Miron A, Iglehart JD, Richardson AL. BRCA1 promoter methylation in sporadic breast tumors: relationship to gene expression profiles. *Breast Cancer Res. Treat.* 91(2), 179–186 (2005).

52 Fiegl H, Millinger S, Goebel G *et al.* Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. *Cancer Res.* 66(1), 29–33 (2006).

53 Collett K, Eide GE, Arnes J *et al.* Expression of enhancer of zeste homologue 2 is significantly associated with increased tumor cell proliferation and is a marker of aggressive breast cancer. *Clin. Cancer Res.* 12(4), 1168–1174 (2006).

54 Healey MA, Hu R, Beck AH *et al.* Association of H3K9me3 and H3K27me3 repressive histone marks with breast cancer subtypes in the Nurses' Health Study. *Breast Cancer Res. Treat.* 147(3), 639–651 (2014).

55 Kleer CG, Cao Q, Varambally S *et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl Acad. Sci. USA* 100(20), 11606–11611 (2003).

56 Yang H, Pinello CE, Luo J *et al.* Small-molecule inhibitors of acetyltransferase p300 identified by high-throughput screening are potent anticancer agents. *Mol. Cancer Ther.* 12(5), 610–620 (2013).

57 Hudelist G, Singer CF, Manavi M, Pischinger K, Kubista E, Czerwenka K. Co-expression of ErbB-family members in human breast cancer: Her-2/neu is the preferred dimerization

candidate in nodal-positive tumors. *Breast Cancer Res. Treat.* 80(3), 353–361 (2003).

58 Gojis O, Rudraraju B, Gudi M *et al.* The role of SRC-3 in human breast cancer. *Nat. Rev. Clin. Oncol.* 7(2), 83–89 (2010).

59 Osborne CK, Bardou V, Hopp TA *et al.* Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. *J. Natl Cancer Inst.* 95(5), 353–361 (2003).

60 Yan PS, Venkataramu C, Ibrahim A *et al.* Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. *Clin. Cancer Res.* 12(22), 6626–6636 (2006).

61 Bernstein BE, Mikkelsen TS, Xie X *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2), 315–326 (2006).

62 Wang Z, Zang C, Rosenfeld JA *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* 40(7), 897–903 (2008).

63 Miranda TB, Cortez CC, Yoo CB *et al.* DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. *Mol. Cancer Ther.* 8(6), 1579–1588 (2009).

64 Marks PA. Discovery and development of SAHA as an anticancer agent. *Oncogene* 26(9), 1351–1356 (2007).