Aurora kinase B (AURKB; also known as Aik2, AIK2, AIM1, AIM-1, ARK2, ARK-2, Aurb, aurkb-sv1, aurkb-sv2, Aurora/IPL1-related kinase 2, Aurora-and IPL1-like midbody-associated protein 1, Aurora kinase B, Aurora-related kinase 2, IPL1, Serine/threonine-protein kinase 12, Serine/threonine-protein kinase aurora-B, STK-1, STK12, STK5) encodes a protein member of the aurora kinase subfamily of serine/threonine kinases that function as a regulator of the centrosome cycle and mitotic spindle assembly. The protein AURKB and interacting proteins play an important role in chromosome condensation, segregation and cytokinesis and, consequently, in ploidy maintenance during cell division. Mitotic deregulations may contribute significantly to cell division errors and development of aggressive tumour cells [1,2]. AURKB signalling was also linked to breast cancer and associated to poor prognosis [3]. This gene has been a target of inactivation via different studies and mechanisms [4-13].

<table>
<thead>
<tr>
<th>Illumina Probe_ID</th>
<th>Symbol</th>
<th>Luminal A</th>
<th>Luminal B</th>
<th>Her2 enriched</th>
<th>Normal like</th>
<th>Basal like</th>
<th>Gene and Protein Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_1684217</td>
<td>AURKB</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Cell division cycle associated 5 (CDC5; also known as Cell division cycle-associated protein 5, MGC16386, p35, Sororin, SORORN) is important for sister chromatid cohesion during mitosis, stabilizing proper chromatin association during G2 phase. The protein is also needed for efficient repair of DNA double-strand breaks and for stable presence of normal amounts of chromatin-bound cohesin population [14,15]. Carretero et al [16] reported that “the reduced accumulation of AURKB at the inner centromere in cells that lack PDSSB impairs its error correction function, promoting chromosome mis-segregation and aneuploidy”. Although systematic studies showed the up-regulation of this gene in a great majority of lung cancers, its involvement in breast cancer disease requires further investigation. The protein CDC5 confers a potential diagnostic molecule and therapeutic target for promising strategies in new drug development [17].</td>
</tr>
<tr>
<td>ILMN_1683450</td>
<td>CDCA5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Centrosomal protein 55kDa (CEP55; also known as C10orf3, Centrosomal protein of 55 kDa, Cep55, CT111, FLJ10540, Up-regulated in colon cancer 6, URCC6) encodes a mitotic phosphoprotein that acts in mitotic exit and cytokinesis. Both down- and up-regulation of CEP55 causes a cytokinesis defect [18]. The gene expression variance, nonetheless, is not affected by hormone receptors such as oestrogen and progesterone or expression patterns of ERBB2. The centrosomal protein is detected in a wide variety of tumour cell lines and is considered as a novel breast tumour-associated antigen [19]. It was also reported that CEP55 have genomic alternation in a comparison of pre-invasive ductal carcinoma in situ (DCIS) to invasive ductal carcinoma (IDC) by Colak et al [20] and it was predictive of prognosis in ER-positive patients [21].</td>
</tr>
<tr>
<td>ILMN_1747016</td>
<td>CEP55</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>The maternal embryonic leucine zipper kinase (MELK; also known as hMELK, HPK38, HPK38, KIAA0175, Maternal embryonic leucine zipper kinase, Protein kinase PK38) acts as a regulator of various processes such as cell cycle control, self-renewal of stem cells, apoptosis, and splicing regulation. The encoded protein physically interacts, phosphorylates and inhibits BCL2L14, repressing a pro-apoptotic member of the Bcl-2 family. The protein also mediates phosphorylation of CDC25B, regulating the entry into mitosis. In addition, MELK inhibits the spliceosome during mitosis by phosphorylating ZNF622 and contributes to induce other apoptosis signalling regulation. The protein kinase is a promising molecular target for the treatment of breast cancer [22] as the up-regulation is linked to poor prognosis [23-27]. Finally, it has been suggested that paclitaxel may attenuate the expression of MELK [28].</td>
</tr>
<tr>
<td>ILMN_2212909</td>
<td>MELK</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Ubiquitin-conjugating enzyme E2C (UBE2C; also known as dJ447F3.2, UbcH10, UBC1H10, Ubiquitin carrier protein C, Ubiquitin-conjugating enzyme E2 C, Ubiquitin-protein ligase C) encodes a member of the E2 ubiquitin-conjugating enzyme family. The ubiquitin modification in proteins is an important cellular mechanism of homeostasis and fate [29]. UBE2C is required for cell cycle progression and checkpoint control through targeted degradation of short-lived proteins, the mitotic cyclins. Alterations in this pathway is implicated in cancer progression and, importantly, in the pathogenesis of breast cancer [30,31]. The UBE2C up-regulation is also normally linked to high tumour grade and poor prognosis [32,33].</td>
</tr>
<tr>
<td>ILMN_1714730</td>
<td>UBE2C</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Illumina Probe_ID</th>
<th>Symbol</th>
<th>Luminal A</th>
<th>Luminal B</th>
<th>Her2 enriched</th>
<th>Normal like</th>
<th>Basal like</th>
<th>Gene and Protein Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_1796059</td>
<td>ANKR30A</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ankyrin repeat domain 30A (ANKRD30A; also known as Ankyrin repeat domain-containing protein 30A, NY-BR-1, Serologically defined breast cancer antigen NY-BR-1) is an antigen expressed in mammary glands, primary and metastatic breast carcinomas [34-36]. Interestingly, ANKR30A is almost expressed exclusively in breast epithelium; with exception of testis and sweat gland tumours. Despite the insufficient knowledge about the biology and function of the gene, the tissue specificity may be useful for the diagnosis of breast carcinomas [34,37-39]; and a potential target for treatment (immunotherapy) [40,41].</td>
</tr>
<tr>
<td>ILMN_1651329</td>
<td>LINC00993</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The long intergenic non-protein coding RNA 993 (LINC00993) matches a region in the chromosome 10 very close to ANKR30A; and contains a SNP (rs77587276) variant. The region requires further investigation as it covers relevant probes associated with breast cancer disease; markedly, up-regulated in luminal subtype.</td>
</tr>
<tr>
<td>ILMN_2310814</td>
<td>MAPT</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The microtubule-associated protein tau (MAPT; also known as DDPAC, FLJ31424, FTDP-17, MAPTL, MGC138549, Microtubule-associated protein tau, MSTD, MTBT1, MTBT2, Neurofibrillary tangle protein, Paired helical filament-tau, PHF-tau, PPND, tau, TAU) undergoes alternative splicing, originating several mRNA transcripts. The isoforms differ by having variant conserved repeat motifs in the microtubule-binding domain, and insertions in the N-terminal projection domain. Although the function of each isofrom is unknown, the protein binds to both the outer and inner surfaces of microtubules, organizing the tubulin assembly and microtubule stabilization [42]. In breast cancer, MAPT expression is high in ER-positive low grade compared to ER-negative high grade tumours [43]. Additionally, the gene up-regulation is correlated with favourable prognosis [44] and, at the same time, associated with resistance to taxanes, paclitaxel and docetaxel [42,45-49].</td>
</tr>
<tr>
<td>ILMN_1728787</td>
<td>AGR3</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
<td>The anterior gradient 3 (AGR3; also known as AG3, AG-3, Anterior gradient protein 3 homolog, BCMP11, Breast cancer membrane protein 11, HAG3, HAG-3, PDIA18, UNQ642/PRO1272) functionality has been defined in breast cancer cells as involved in hormone responsiveness, cell adhesion, migration, and metastasis. This gene encodes a membrane protein with a potential role in tumorigenesis by interacting with metastasis-associated genes [50,51].</td>
</tr>
<tr>
<td>ILMN_1688071</td>
<td>NAT1</td>
<td>+</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td>The enzyme encoded by N-acetyltransferase 1 (NAT1; also known as AAC1, Arylamide acetylase 1, Arylamine N-acetyltransferase 1, MNAT, Monomorphous arylamine N-acetyltransferase, N-acetyltransferase type 1, NAT-1, NAT1) acts metabolizing drugs and other xenobiotics, and functions in folate catabolism [52,53]. Kim et al. (2008) reported the hypomethylation of the NAT1 promoter region resulting in aberrant mRNA expression levels, with overexpression of the gene in breast carcinomas. Likewise, new insights into the associations of SNPs in the coding and control regions of NAT1 have been described and suggested as a potential susceptibility biomarker for the disease [53,54].</td>
</tr>
<tr>
<td>Illumina Probe_ID</td>
<td>Symbol</td>
<td>Luminal A</td>
<td>Luminal B</td>
<td>Her2 enriched</td>
<td>Normal like</td>
<td>Basal like</td>
<td>Gene and Protein Review</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>ILMN_1729216</td>
<td>CRYAB</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td>Crystallin, alpha B (CRYAB; also known as Alpha(B)-crystallin, Alpha-crystallin B chain, CRYA2, CTBP2, Heat shock protein beta-5, HspB5, HSPB5, Renal carcinoma antigen NY-REN-27, Rosenthal fiber component) is a member of the small heat shock protein (sHSP; also known as the HSP20) family, all of which share a common C terminal motif – the alpha crystallin domain. The protein CRYAB acts as molecular chaperones induced by ubiquitous stress and up-regulated by heat, radiation, oxidative stress and anticancer drugs. Other additional functions of alpha crystallins are the autokinase activity, participation in the intracellular architecture, and the control of large soluble protein aggregates. Additionally, CRYAB shows redundancy in interacting with various apoptosis pathways at multiple levels [55]; besides plays a critical role in vasculature homeostasis and angiogenesis [56]. The protein is expressed widely in many tissues and organs [57]. In breast cancer, CRYAB is usually high differentially expressed in invasive tumours when compared to normal breast tissue specimens [55] and might be involved in chemotheraphy response [58].</td>
</tr>
<tr>
<td>ILMN_1666845</td>
<td>KRT17</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td>The protein encoded by keratin 17 (KRT17; also known as 39.1, CK-17, Cytokeratin-17, K17, Keratin, type I cytoskeletal 17, Keratin-17, PC, PC2, PCCH1) plays a role in the formation and maintenance of various epidermal appendages, as the nail bed, hair follicle, and sebaceous glands. KRT17 regulates other protein synthesis and epithelial cell growth. In addition, the protein is a marker of basel cell differentiation as an attribute of a certain type of “stem cells”. In the context, immunohistochemical studies revealed that basel-like breast tumours present KRT17 up-regulation, with levels associated with a poor clinical outcome [59]. This gene is also up-regulated in primary BRCA1 mutant breast tumours; fact consistent with the reported connection between BRCA1 mutation and basal-like subtypes [60].</td>
</tr>
<tr>
<td>ILMN_1786720</td>
<td>PROM1*</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td>Prominin 1 (PROM1; also known as AC133, Antigen AC133, CD133, CORD12, MCDR2, MSTP061, Prominin-1, Prominin-like protein 1, PROMLI, RP41, STGD1) encodes a transmembrane glycoprotein, often expressed on adult stem cells. The protein plays an essential role in maintaining stem cell properties by suppressing differentiation. Abrant expression PROM1 is also associated with several types of cancer. In breast cancer, PROM1 overexpression is positively related to tumour size, stage, and lymph node metastasis in invasive tumours [61]. Moreover, there is an important association with p53 mutation, mammary cell dedifferentiation, and the concomitant acquisition of stemlike properties [62]. In particular, basal-like subtype shows high p53 mutation and PROM1 up-regulation, which improve tumour cells aggressiveness [63] due to activation of angiogenesis and metastasis [64]; besides chemoresistance [65].</td>
</tr>
<tr>
<td>ILMN_1753101</td>
<td>VTCN1</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td>V-set domain containing T cell activation inhibitor 1 (VTCN1; also known as B7h5, B7h4, B7h3, B7x, FLJ22418, Immune costimulatory protein B7-H4, PR01291, Protein B7s1, T-cell costimulatory molecule B7x, UNQ659/PR01291, VCTN1, V-set domain-containing T-cell activation inhibitor 1) is found on the surface of antigen-presenting cells and interacts with ligands attached to receptors on the surface of T cells. The protein negatively regulates the immune response of T cells by reducing the production of cytokines and controlling the cell cycle progression [66,67]. High levels of the VTCN1 mRNA and the related protein are associated with a number of cancers, including ovarian and breast cancers [68]. The up-regulation is also associated with tumour progression and linked to poor prognosis [69]. Although VTCN1 detection is observed in PR-negative and HER2-negative tumours, the expression is independent of grade, stage, or other clinicopathologic variables [70].</td>
</tr>
<tr>
<td>ILMN_1798108</td>
<td>C6orf211</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td>The chromosome 6 open reading frame 211 (C6orf211) is mapped in a region close to ESR1 and other genes (AKAP12 and CCDC170), suggesting further investigation of a possible connection with the ESR1 transcription and the luminal subtype in breast cancer disease.</td>
</tr>
<tr>
<td>Illumina Probe_ID</td>
<td>Symbol</td>
<td>Luminal A</td>
<td>Luminal B</td>
<td>Her2 enriched</td>
<td>Normal like</td>
<td>Basal like</td>
<td>Gene and Protein Review</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ILMN_1747911</td>
<td>CDK1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The cyclin-dependent kinase 1 (CDK1; also known as CDC2, CDC28A, Cell division control protein 2 homolog, Cell division protein kinase 1, Cyclin-dependent kinase 1, DKFZp686L20222, MGC111195, P34CDC2, p34 protein kinase) plays a key role in the control of the cell cycle by modulating the centrosome as well as mitotic onset; promotes G2-M transition, and regulates G1 progress and G1-S transition associated with multiple interphase cyclins. The kinase activity of this protein is controlled by cyclin accumulation and destruction through the cell cycle. In addition, CDK1 complexes phosphorylate several substrates that trigger centrosome separation, Golgi dynamics, nuclear envelope breakdown, chromosome condensation, and apoptosis. An abnormal phosphorylation occurs in cancer cell lines, as well as in primary breast tissues and lymphocytes. Moreover, high CDK1 activity was linked to the absence of a full DNA damage response in mitotic cells [71]. Although the impact of this gene in breast cancers remains controversial, there is a significant association with unfavourable clinicopathologic feature such as high histologic grade, large tumour size, lymph node metastases and PR-negative tumours [72]. Ultimately, CDK1 may be used as a predictive factor to identify patient’s response to neoadjuvant chemotherapy [73-75].</td>
</tr>
<tr>
<td>ILMN_1666305</td>
<td>CDKN3</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The protein encoded by cyclin-dependent kinase inhibitor 3 (CDKN3; also known as CDH1, CDK2-associated dual-specificity phosphatase, CIP2, Cyclin-dependent kinase inhibitor 3, Cyclin-dependent kinase-interacting protein 2, Cyclin-dependent kinase inhibitor 1, FLJ25787, KAP, KAP1, Kinase-associated phosphatase, MGC70625) belongs to the dual specificity protein phosphatase family, active toward substrates containing either phosphotyrosine or phosphoserine residues. CDKN3 is a cyclin-dependent kinase inhibitor, and interacts / dephosphorylates CDK2 kinase, thereby reducing its ability to phosphorylate the retinoblastoma protein (RB). Non-phosphorylated RB binds transcription factor E2F1 and prevents the G1-S transition. CDKN3 was reported to be deleted, mutated, or overexpressed in several types of cancers, including breast tumours [76].</td>
</tr>
<tr>
<td>ILMN_1678535</td>
<td>ESR1</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>Estrogen receptor 1 (ESR1; also known as DKFZp686N23123, ER, Era, ER-alpha, ESR, ESRA, Estradiol receptor, Estrogen receptor, NR3A1, Nuclear receptor subfamily 3 group A member 1) encodes a protein receptor, a ligand-activated transcription factor composed of several domains important for hormone binding, DNA binding, and activation of transcription. Estrogen and its receptors are central regulators of breast cancer disease and are associated with response to endocrine therapy. Down-regulation of ESR1, or eventual mutations, may indicate intrinsic resistance to tamoxifen, increased risk of tumour recurrence, and worse prognosis [77-79]; even though the mechanisms by which oestrogen receptor dictates tumour status are poorly understood [80].</td>
</tr>
<tr>
<td>ILMN_2149164</td>
<td>SFRP1</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>Major gene expression changes occur during progression of neoplastic cells, including down regulation of secreted frizzled-related protein 1 (SFRP1; also known as FRP, FRP1, FRP-1, FrzA, SARP2, SARP-2, Secreted apoptosis-related protein 2, Secreted frizzled-related protein 1, sFRP1, sFrp1) [81]. SFRP1 functions as a negative regulator of Wnt/β-catenin pathway, implicated in several human cancers, including breast tumours and respective cell lines [82-89]. Reduced levels of SFRP1 results in hyperplastic lesions and its loss may be a critical event in cancer initiation [90]. In breast carcinomas, SFRP1 showed significant differences in methylation patterns between ER-negative and ER-positive tumours [91]. The hypermethylation of the SFRP1 promoter and gene down-regulation has been widely reported in breast cancer [81,92-94] and associated with tumour invasion and decreased survival [95-98]. A potential combinatorial treatment - romidepsin and decitabine - has recently been administered in cell lines, promoting SFRP1 reexpression with consequently proliferation inhibition and cell death induction via apoptosis [92].</td>
</tr>
</tbody>
</table>
Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3 (SERPINA3; also known as AACT, ACT, alpha-1-antichymotrypsin, Alpha-1-antichymotrypsin, Cell growth-inhibiting gene 24/25 protein, GIG24, GIG25, MGC928254, Serpin A3) encodes a plasma protease inhibitor and member of the serine protease inhibitor class. SERPINA3 regulates the activity of neutrophil cathepsin G and is an oestrogen-induced gene. In breast cancer, the mRNA increased expression was reported as an indicator of good prognosis in oestrogen receptor positive breast cancer [99-101]. It is a maker of oestrogen regulation [102]; besides a predictor of tumour response to neoadjuvant chemotherapy [99].

**Gene and Protein Review**

<table>
<thead>
<tr>
<th>Illumina Probe_ID</th>
<th>Symbol</th>
<th>Luminal A</th>
<th>Luminal B</th>
<th>Her2 enriched</th>
<th>Normal like</th>
<th>Basal like</th>
<th><strong>Gene and Protein Review</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_1788874</td>
<td>SERPINA3</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3 (SERPINA3; also known as AACT, ACT, alpha-1-antichymotrypsin, Alpha-1-antichymotrypsin, Cell growth-inhibiting gene 24/25 protein, GIG24, GIG25, MGC928254, Serpin A3) encodes a plasma protease inhibitor and member of the serine protease inhibitor class. SERPINA3 regulates the activity of neutrophil cathepsin G and is an oestrogen-induced gene. In breast cancer, the mRNA increased expression was reported as an indicator of good prognosis in oestrogen receptor positive breast cancer [99-101]. It is a maker of oestrogen regulation [102]; besides a predictor of tumour response to neoadjuvant chemotherapy [99].</td>
</tr>
<tr>
<td>ILMN_1785570</td>
<td>SUSD3*</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>The sushi domain containing 3 (SUSD3; also known as MGC26847, Sushi domain-containing protein 3, UNQ9387/PRO14275) down-regulated in breast carcinomas is associated with a malignant phenotype (short-term overall survival, endocrine insensitivity, triple-negative status, poor tumour differentiation). SUSD3 is highly expressed in ER-positive breast tumours and the treatment with oestriadiol may increase the gene expression in cancer cells [103]. The SUSD3 abnormal mRNA levels and the protein function, however, are still unclear and require urgent investigation [100].</td>
</tr>
<tr>
<td>ILMN_1803236</td>
<td>CLCA2</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The protein encoded by chloride channel accessory 2 (CLCA2; also known as CACC, CACC3, CaCC-3, Calcium-activated chloride channel family member 2, Calcium-activated chloride channel protein 3, Calcium-activated chloride channel regulator 2, CLCRG2, FLJ97885, hCaCC-3, hCLCA2) belongs to the calcium sensitive chloride conductance protein family. The protein plays a role in modulating chloride current across the plasma membrane in a calcium-dependent mode. CLCA2 is also involved in basal cell adhesion and/or stratification of squamous epithelia. In addition, the molecule is involved in the p53 network and may act as a tumour suppressor in breast and colorectal cancer, inhibiting cancer cell migration and invasion [104-106]. The mechanisms behind the silencing of CLCA2 in luminal breast cancers and the up-regulation in HER2-enriched and basal-like subtypes, however, have not been elucidated. Ultimately, cell lines CLCA2-negative treated with demethylating agents restored the expression of the gene, suggesting an epigenetic control [107].</td>
</tr>
<tr>
<td>ILMN_2161820</td>
<td>GLYATL2*</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The enzyme encoded by glycine-N-acyltransferase-like 2 (GLYATL2; also known as Acyl-CoA:glycine N-acyltransferase-like protein 2, BXMAS2-10, GATF-B, Glycine N-acyltransferase-like protein 2, MGC244096) conjugates medium- and long-chain saturated and unsaturated acyl-CoA esters to glycine, resulting in the production of N-oleoyl glycine and also N-arachidonoyl glycine. N-Oleoyl glycine and N-arachidonoyl glycine are identified as signalling molecules that regulate the perception of pain and body temperature, and also have anti-inflammatory properties [108]. GLYATL2 is up-regulated in salivary gland and trachea, and detected also in spinal cord and skin fibroblasts. In addition, the high levels of the gene in skin and lung may indicate a role in barrier function/immune response and lipid signalling [109].</td>
</tr>
<tr>
<td>ILMN_1810978</td>
<td>MUCL1*</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The mucin-like 1 (MUCL1; also known as Mucin-like protein 1, Protein BS106, SBEM, Small breast epithelial mucin, UNQ590/PDB1160) encodes a 90 amino acids glycoprotein that exhibits characteristics of members of the mucin family. The presence of a hydrophobic signal peptide within the protein sequence suggests that MUCL1 is a secreted and subjected to proteolytic processing. The putative gene is expressed only in mammary and salivary glands and it is promising as a new biomarker with high tissue specificity [110], besides with a great potential for predicting metastasis and response to neoadjuvant chemotherapy [111]. In breast cancer, the protein is more frequently observed in ER-negative than in ER-positive cancers, and positively associated with HER2 overexpression. The evaluation of MUCL1 expression, nonetheless, needs to consider also the heterogeneity and different molecular subtypes [112]. In general, increased expression levels of MUCL1 are strongly associated with higher tumour grades, lymph node metastasis [113] and reduced patient overall survival [114,115].</td>
</tr>
<tr>
<td>Illumina Probe_ID</td>
<td>Symbol</td>
<td>Luminal A</td>
<td>Luminal B</td>
<td>Her2 enriched</td>
<td>Normal like</td>
<td>Basal like</td>
<td>Gene and Protein Review</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------------------</td>
</tr>
<tr>
<td>ILMN_1773459</td>
<td>SOX11*</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SRY (sex determining region Y)-box 11 (SOX11; also known as Transcription factor SOX-11) encodes a protein involved in the regulation of embryonic development and in the determination of the cell fate. The protein acts as a transcriptional regulator after modelling a complex with other proteins. SOX11 functions in the developing nervous system and play a role in tumorigenesis. In breast cancer patients, the gene up-regulation might contribute to a proliferative genotype which may be linked to poor prognosis and therefore worse overall survival [116]. Interestingly, SOX11 levels were higher in basal-like and HER2-enriched breast cancers compared with other subtypes [117].</td>
</tr>
<tr>
<td>ILMN_1674533</td>
<td>TRPV6</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Transient receptor potential cation channel, subfamily V, member 6 (TRPV6; also known as ABP/ZF, Calcium transport protein 1, Cat1, CAT1, CATL, CatT-L, CatT-like, ECAc2, ECAc2, Epithelial calcium channel 2, HSA27799.0, LP6728, Transient receptor potential cation channel subfamily V member 6, TrpV6, ZFAB) encodes a member of a family of multpass membrane proteins that functions as calcium channels. The calcium selective cation channel is involved in Ca2+ uptake in various tissues, including Ca2+ reabsorption in intestine. The up-regulation of TRPV6 at both the mRNA and protein levels is observed in several tumours such as breast, prostate, colon, thyroid and ovary; and in various tumour cell lines including those of colon, human leukaemia and prostate [118,119]. In particular, increased TRPV6 expression is a feature of ER-negative breast tumours (HER2-enriched and basal-like subtypes) and has been associated to patient decreased survival [120]. The exact mechanism underlying the TRPV6-mediated regulation of cancer progression and its downstream signalling, however, remain poorly understood [121]. Inhibitors of the TRPV6 channel have been investigated as potential targets for diagnosis, prognosis and/or therapeutic approaches in human cancers [122,123].</td>
</tr>
<tr>
<td>ILMN_1687235</td>
<td>HPN</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The gene hepsin (HPN; also known as Serine protease hepsin, TMPRSS1, Transmembrane protease serine 1) encodes a type II transmembrane serine protease that is involved in diverse cellular functions, including cell growth and maintenance of cell morphology. The protein is cleaved into a catalytic serine protease chain and a non-catalytic scavenger receptor. The expression of the encoded protein is associated with the progression of several types of malignancies, nevertheless little is known about its clinical and biological significance in breast cancer. HPN is up-regulated in breast tumours; besides significantly associated with tumour stage, lymph node metastasis, oestrogen receptor positivity, and progesterone receptor positivity [124].</td>
</tr>
<tr>
<td>ILMN_1655915</td>
<td>MMP11</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The matrix metallopeptidase 11 (stromelysin 3) (MMP11; also known as Matrix metalloproteinase-11, MMP-11, SL-3, ST3, STM13, Stromelysin-3) is a member of the matrix metalloproteinase (MMP) family of proteases. These proteins are constituent of the extracellular matrix and act on the epithelial/connective interface in embryogenesis, wound healing, tissue involution, and reproduction. MMP11 expressed in fibroblasts near areas of invasive carcinoma lead to the gain of metastatic potential for spread of tumour cells, with patterns of subsequent invasion and migration for different types of solid tumours and cell lines [125-129]. In breast cancer, higher expression level of MMP11 is correlated with patients having poorly differentiated tumours, increased invasiveness, node metastasis, and worse prognosis [130-134]. MMP11 gene expression analysis may also be used in clinical applications for breast cancer diagnosis, management and therapy [135,136]. Ultimately, the up-regulation of this gene is linked to other markers such as p53, ER and HER2 [137].</td>
</tr>
<tr>
<td>ILMN_1711470</td>
<td>UBE2T</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The ubiquitin-conjugating enzyme E2T (putative) (UBE2T; also known as Cell proliferation-inducing gene 50 protein, HSPC150, PIG50, Ubiquitin carrier protein T, Ubiquitin-conjugating enzyme E2 T, Ubiquitin-protein ligase T) accepts ubiquitin from the E1 complex and catalyses its covalent attachment to other proteins. The covalent conjugation of ubiquitin to proteins regulates diverse cellular pathways and proteins. Ubiquitin is transferred to a target protein through a concerted action of ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin ligase (E3). UBE2T acts as a specific E2 ubiquitin-conjugating enzyme for the Fanconi anemia complex and contribute to ubiquitination and degradation of BRCA1. The enzyme is up-regulated in different types of cancer including breast, bladder, lung, and prostate cancers; playing essential role in cell proliferation. In breast tumours, the gene up-regulation cause the decrease of the BRCA1 levels, however, major pathways involving UBE2T are still poorly understood [138].</td>
</tr>
<tr>
<td>Illumina Probe_ID</td>
<td>Symbol</td>
<td>Luminal A</td>
<td>Luminal B</td>
<td>Her2 enriched</td>
<td>Normal like</td>
<td>Basal like</td>
<td>Gene and Protein Review</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------</td>
<td>------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>ILMN_1789507</td>
<td>COL11A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The collagen, type XI, alpha 1 (COL11A1; also known as COL11A1, COLL6, Collagen alpha-1(XI) chain, STL2) encodes one of the two alpha chains of type XI collagen, a minor fibrillar collagen. Type XI collagen is a heterotrimer and play an important role in fibrillogenesis by controlling lateral growth of collagen fibrils. COL11A1 is expressed by both the epithelial and stromal compartments and its expression is deregulated in a range of cancers, such as breast and colon. In particular, molecules related to extracellular matrix remodelling (e.g. COL11A1) are differentially expressed in breast tumours in situ and invasive; enriched in metastatic tumour cells [81,139,140].</td>
</tr>
<tr>
<td>ILMN_1740609</td>
<td>CCL15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>The chemokine (C-C motif) ligand 15 (CCL15; also known as C-C motif chemokine 15, Chemokine CC-2, HCC-2, HMRP-2B, Leukotactin-1, LKN1, Lkn-1, LKN-1, Macrophage inflammatory protein 5, MIP-1d, MIP-1D, MIP-1 delta, MIP5, MIP-5, Mip-2b, MRP-2B, NCC3, NCC-3, SCYA15, SCYL3, Small-inducible cytokine A15, SY15) encodes a secreted protein characterized by two adjacent cysteines, further processed into numerous smaller functional peptides. The protein has chemotactic factor that attracts T cells and monocytes; acts through C-C chemokine receptor type 1 (CCR1) and also binds to type 3 (CCR3). In hepatocellular carcinoma, the up-regulation of CCL15 promotes cell migration and invasion [141]. High levels of CCL15 also increase the expression of matrix metalloproteinase and induce angiogenesis [142].</td>
</tr>
<tr>
<td>ILMN_1651282</td>
<td>COL17A1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>The collagen, type XVII, alpha 1 (COL17A1; also known as 180 kDa bullous pemphigoid antigen 2, BA16H23.2, BP180, BPAG2, Bullous pemphigoid antigen 2, Collagen alpha-1(XVII) chain, FLJ60881, KIAA0204, LAD-1) encodes the alpha chain of type XVII collagen, a transmembrane protein or a soluble form generated by proteolytic processing of the full length form. The protein is a structural component of hemidesmosomes, multiprotein complexes at the dermal-epidermal basement membrane zone that mediate adhesion of basal keratinocytes to the underlying membrane. Hemidesmosomal components are also implicated in signal transduction and thereby are able to influence cell growth, motility and differentiation. In neoplastic tissue, COL17A1 aberrant expression depends on the stage of the tumour, down-regulated in mild dysplasia and up-regulation as the tumour further evolves [143].</td>
</tr>
<tr>
<td>ILMN_1723684</td>
<td>DARC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>Duffy blood group, atypical chemokine receptor (DARC; also known as CCBP1, CD234, Dfy, Duffy antigen/chemokine receptor, FY, Fy glycoprotein, Glycoprotein D, GPD, GpFy, Plasmodium vivax receptor, WBCQ1) encodes a glycosylated membrane protein and a non-specific receptor for several chemokines. Polymorphisms in this gene are the basis of the Duffy blood group system. It is reported that DARC plays a negative regulatory role in human breast cancer. Overexpression of DARC protein in breast cancer cells leads to significant inhibition of tumorigenesis and metastasis [144-146]. DARC is also correlated with breast cancer incidence, auxiliary lymph node metastasis and overall survival [147].</td>
</tr>
<tr>
<td>ILMN_1809099</td>
<td>IL33*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>Interleukin 33 (IL33; also known as Citof26, DKFZp586H0523, DV227, IL1F11, IL-1F11, IL-33, Interleukin-1 family member 11, Interleukin-33, NFEHEV, NFHEV, NF-HEV, Nuclear factor from high endothelial venules, RP11-57SC20.2) is a member of the IL1 family that induces production of T helper-2 (Th2) associated cytokines. The protein acts as a chemoattractant for Th2 cells, and amplifies immune responses during tissue injury. IL33 also functions as a chromatin-associated nuclear factor with transcriptional repressor properties. In breast cancer cells, a frequent overexpression is observed, though the gene deregulation is not clearly understood in the disease [148].</td>
</tr>
<tr>
<td>Illumina Probe_ID</td>
<td>Symbol</td>
<td>Luminal A</td>
<td>Luminal B</td>
<td>Her2 enriched</td>
<td>Normal like</td>
<td>Basal like</td>
<td>Gene and Protein Review</td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------</td>
<td>------------------------</td>
</tr>
<tr>
<td>ILMN_1766650</td>
<td>FOXA1</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The forkhead box A1 (FOXA1; also known as Forkhead box protein A1, Hepatocyte nuclear factor 3-alpha, HNF3A, HNF-3A, HNF-3-alpha, MGC33105, TCF3A, TCF-3A, Transcription factor 3A) encodes a member of the forkhead class of DNA-binding proteins. The nuclear factor is a transcriptional activator involved in embryonic development, establishment of tissue-specific gene expression and regulation of gene expression in differentiated tissues. The protein is also implicated in the development of multiple organs such as liver, pancreas, thyroid, prostate and breast. Basically, it modulates the transcriptional activity of nuclear hormone receptors [149]. FOXA1 acts in both androgen receptor (AR) and oestrogen receptor (ER), directing the binding location, and therefore the transcriptional activity [150-152]. In breast cancer, FOXA1 plays a pivotal role in mammary ductal morphogenesis [153], tumour early stage, drug response, and metastatic disease [154,155]. Mutation and SNP variation located in enhancer regions may alter FOXA1 binding affinity and affect breast cancer risk [156-159]. In addition, high expression of FOXA1 is correlated with luminal A subtype and it is a significant predictor of survival in patients with ER-positive tumours [149,160-162], and a marker of good prognosis [163-166]. Ultimately, genome analysis of ER-FOXA1 interactions is required to understand the molecular mechanisms of ER activity [167-170].</td>
</tr>
<tr>
<td>ILMN_1811387</td>
<td>TFF3</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The trefoil factor 3 (intestinal) (TFF3; also known as hITF, HTF, hP1.B, Intestinal trefoil factor, ITF, P1B, Polypeptide P1.B, TFI, Trefoil factor 3, FREAC3, FREAC, Trefoil factor 3, FREAC, Polyptide P1.B, TFI, Trefoil factor 3) gene is translated in a stable secretory protein having at least one copy of the trefoil motif and a domain with three conserved disulphides. TFF3 functions as ‘luminal epithelium guardian’, involved in the maintenance and repair of the mucosa after damage. Besides, promotes the mobility of epithelial cells in healing processes. Up-regulation of TFF3 is observed in various neoplastic diseases, including breast cancer, where the gene has been target as a biomarker [171]. TFF3 is induced by hormones such as oestrogen, and is usually combined with TFF1 expression in ER-positive malignant breast tumour cells [172,173]. Moreover, the TFF3 levels are close to that of ESR1, yet reduce after tamoxifen treatment [32,174]. These genes are components of the ‘luminal epithelial’ signature defining a well-differentiated, low-grade intrinsic subtype of breast cancer: the luminal A [172], Basal-like and claudin-low breast cancer subtypes showed frequent hypermethylation of the TFF3 promoter region [175,176].</td>
</tr>
<tr>
<td>ILMN_1738401</td>
<td>FOXC1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The forkhead box C1 (FOXC1; also known as ARA, FKHL7, Forkhead box protein C1, Forkhead-related protein FKHL7, Forkhead-related transcription factor 3, FREAC3, FREAC-3, IGDA, HG1, IRID1, RIEG3) is part of the forkhead family of transcription factors, characterized by a common DNA-binding domain. All the mechanisms through FOXC1 are not yet determined; however, the gene plays important roles in cell growth, survival, differentiation, and migration. FOXC1 is identified as a functionally important biomarker of breast cancer aggressiveness, particularly associated with basal-like breast cancer subtype [177,178]. The gene up-regulation in breast tumour cells induces epithelial-mesenchymal transition, drug resistance, and increased cell proliferation and invasion [179].</td>
</tr>
<tr>
<td>ILMN_1689146</td>
<td>GABRP</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gamma-aminobutyric acid (GABA) A receptor, pi (GABRP; also known as GABA(A) receptor subunit pi, Gamma-aminobutyric acid receptor subunit pi, MGC126386, MGC126387) encodes a transmembrane protein composed by multisubunit in the chloride channel that mediates synaptic transmission in the central nervous system. The gene is also expressed in several non-neuronal tissues including the uterus, breast, and ovaries. In breast tissue, GABRP is mainly expressed in myoepithelial/basal cells, and the function is related to tissue contractility [172]. In breast cancer cells, the gene is normally up-regulated among ER-negative patients [180,181].</td>
</tr>
<tr>
<td>Illumina Probe_ID</td>
<td>Symbol</td>
<td>Luminal A</td>
<td>Luminal B</td>
<td>Her2 enriched</td>
<td>Normal like</td>
<td>Basal like</td>
<td>Gene and Protein Review</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ILMN_1807423</td>
<td>IGF2BP3*</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The protein encoded by insulin-like growth factor 2 mRNA binding protein 3 (IGF2BP3; also known as CT98, DKFZp686F1078, hKOC, IGF2 mRNA-binding protein 3, IGF-II mRNA-binding protein 3, IMP3, IMP-3, Insulin-like growth factor 2 mRNA-binding protein 3, KH domain-containing protein overexpressed in cancer, KOC1, VICKZ3, VICKZ family member 3) is primarily located in the nucleolus and belongs to a conserved family of RNA-binding proteins. The RNA-binding factor recruits target transcripts to cytoplasmic protein-RNA complexes (mRNP), modulating the rate and location of endonuclease attacks or microRNA-mediated degradation. The protein, nonetheless, is involved not only in RNA synthesis and metabolism, but in various important aspects of cell function, such as cell polarization, migration, morphology, metabolism, proliferation and differentiation. IGF2BP3 is largely absent in adult tissues but de novo synthesized or severely up-regulated in various tumours and tumour-derived cells. In breast cancer, the up-regulation enhances tumour growth, angiogenesis and metastasis, resulting in poorer survival [182-184], and chemoresistance [185]. High expression has also been associated with triple-negative breast carcinomas [186-188].</td>
</tr>
<tr>
<td>ILMN_1692938</td>
<td>PSAT1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The phosphoserine aminotransferase 1 (PSAT1; also known as EPIP, MGC1460, Phosphohydroxythreonine aminotransferase, Phosphoserine aminotransferase, PSA, PSAT) encodes a member of the class-V pyridoxal-phosphate-dependent aminotransferase family. Mutations in this gene are associated with phosphoserine aminotransferase deficiency. PSAT1 methylation and aberrant expression are strongly correlated with specific clinical and pathologic features of breast cancer. Notably, the PSAT1 hypermethylation is associated with low-grade, low-proliferation, hormone receptor ER-positive, lymph node positive breast cancer in post-menopausal women [189]; besides it is an indicator of response to tamoxifen endocrine therapy [190]. On the other hand, high expression of PSAT1 is associated with decreased relapse-free and overall survival of patients, and linked to malignant phenotypic features of breast cancer [189].</td>
</tr>
<tr>
<td>ILMN_1668766</td>
<td>ROPN1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The rhophilin associated tail protein 1 (ROPN1; also known as Cancer/testis antigen 91, CT91, DKFZp434B1222, ODF6, Rhophilin-associated protein 1A, RHPNAP1, ROPN1A, rhopporin, Ropporin-1A) is an important reproduction related gene. The protein is involved in sperm maturation, motility, capacitation, hyperactivation and acrosome reaction. Other important functions such as cAMP-dependent protein kinase regulator activity, protein binding activity, phosphorylation and signal transduction regulation were reported; even though ROPN1 requires further investigation [191]. Recently, ROPN1 was validated with diagnostic significance in basal-like breast cancer cells as one of the conserved elements of the SOX10 signature [192].</td>
</tr>
</tbody>
</table>

*Elements named according to Human HT (Illumina HT-12 v3) and matching different regions of the genome with more than one annotation in UCSC and iHOP.*

**Note:** Genes are described according to the RefSeq (Reference Sequence) and UniProt databases; and improved with related citations in PubMed.
References


