The Omics of Triple-Negative Breast Cancers

Hong Xu,1 Peter Eirew,1,2 Sarah C. Mullaly,1 and Samuel Aparicio1,2

BACKGROUND: Triple-negative breast cancers (TNBC) do not represent a single disease subgroup and are often aggressive breast cancers with poor prognoses. Unlike estrogen/progesterone receptor and HER2 (human epidermal growth factor receptor 2) breast cancers, which are responsive to targeted treatments, there is no effective targeted therapy for TNBC, although approximately 50% of patients respond to conventional chemotherapies, including taxanes, anthracyclines, cyclophosphamide, and platinum salts.

CONTENT: Genomic studies have helped clarify some of the possible disease groupings that make up TNBC. We discuss the findings, including copy number–transcriptome analysis, whole genome sequencing, and exome sequencing, in terms of the biological properties and phenotypes that make up the constellation of TNBC. The relationships between subgroups defined by transcriptome and genome analysis are discussed.

SUMMARY: TNBC is not a uniform molecular or disease entity but a constellation of variably well-defined biological properties whose relationship to each other is not understood. There is good support for the existence of a basal expression subtype, p53 mutated, high-genomic instability subtype of TNBC. This should be considered a distinct TNBC subtype. Other subtypes with variable degrees of supporting evidence exist within the nonbasal/p53wt (wild-type p53) TNBC, including a group of TNBC with PI3K (phosphoinositide 3-kinase) pathway activation that have better overall prognosis than the basal TNBC. Consistent molecular phenotyping of TNBC by whole genome sequencing, transcriptomics, and functional studies with patient-derived tumor xenograft models will be essential components in clinical and biological studies as means of resolving this heterogeneity.

© 2013 American Association for Clinical Chemistry

1 Department of Molecular Oncology, BC Cancer Agency, Vancouver, BC; 2 Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC

1L3. Fax 604-675-8218; e-mail saparicio@bccrc.ca.

Received September 25, 2013; accepted November 4, 2013.

* Address correspondence to this author at: Department of Molecular Oncology, BC Cancer Research Centre, 675 West 10th Ave., Vancouver, BC, Canada, V5Z 1L3. Fax 604-675-8218; e-mail saparicio@bccrc.ca.

Previously published online at DOI: 10.1373/clinchem.2013.207167

3 Nonstandard abbreviations: TBNC, triple-negative breast cancers; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; PDX, patient-derived xenografts; EGFR, epidermal growth factor receptor; BL, basal like; EMT, epithelial–mesenchymal transition; AR, androgen receptor; LOH, loss of heterozygosity; HR, homologous recombination; PARP, poly ADP ribose polymerase; VEGF, vascular endothelial growth factor; VEGF receptor 1 (VEGF-1).
levels of genomic instability and poor survival in the first 5 years after diagnosis. This group is enriched for basal-like TNBC, displaying stereotyped copy number changes with 5q loss and gains at 8q, 10p, and 12p. Analysis of mutation frequencies in these 2 groups also shows subtype-specific patterns. IntCluster10 has the highest rate of tumor protein p53 (TP53)\textsuperscript{4} mutations; in IntCluster4, the most frequently mutated gene is phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PIK3CA). Mutations in Usher syndrome 2A (autosomal recessive, mild) (USH2A), filaggrin (FLG), and FAT atypical cadherin 3 (FAT3) are associated with IntCluster10, and v-akt murine thymoma viral oncogene homolog 1 (AKT1) mutation is specifically associated with IntCluster4. The heterogeneity within TNBC supports the necessity of applying different drugs targeting specific driver genes or pathways perturbed within the patient.

Taken together, there is evidence for a major subgroup of a genomically unstable, basal expression pattern, p53-mutated subgroup of TNBC (Fig. 1); the remaining nonbasal, PIK3CA-enriched TNBCs probably form several distinct groupings and are of generally better prognosis. However, all proposed subgroups with intermediate levels of evidence have to be interpreted with some caution. Currently, clinical outcome associations with individual molecular markers are very prone to overfitting of data, especially in small groups of patients. Also, many genomic/transcriptomic markers can be carried together, so the observation of an outcome association is not in itself proof of a

\textsuperscript{4} Human genes: TP53, tumor protein p53; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha; USH2A, Usher syndrome 2A (autosomal recessive, mild); FLG, filaggrin; FAT3, FAT atypical cadherin 3; AKT1, v-akt murine thymoma viral oncogene homolog 1; BRCA1, breast cancer 1, early onset; MK67, marker of proliferation Ki-67; TOP2A, topoisomerase (DNA) II alpha 170kDa; PTEN, phosphatase and tensin homolog; BRCA2, breast cancer 2, early onset; MYO3A, myosin IIIa; RB1, retinoblastoma 1; ATR, ataxia telangiectasia and Rad3 related; UBR1, ubiquitin protein ligase E3 component n-recognin 1; CD6A3, collagen, type VI, alpha 3; Braf, v-raf murine sarcoma viral oncogene homolog B; NRAS, neuroblastoma RAS viral (v-ras) oncogene homolog; ERBB2, v-erb-b2 avian erythroblast leukemia viral oncogene homolog 2; ERBB3, v-erb-b2 avian erythroblast leukemia viral oncogene homolog 3; RB1, retinoblastoma 1; MYC, v-myc avian myelocytomatosis viral oncogene homolog; RAS, ras sarcoma viral oncogene homolog; AURKB, aurora kinase B; BCL2, B-cell CLL/lymphoma 2; BUB1, BUB1 mitotic checkpoint securin/threonine kinase; CDC43, cell division cycle associated 3; CDC44, cell division cycle associated 4; CDC20, cell division cycle 20; CDC45, cell division cycle 45; CHEK1, checkpoint kinase 1; FOXM1, forkhead box M1; HDAC2, histone deacetylase 2; IGF1R, insulin-like growth factor 1 receptor; KIF2C, kinesin family member 2C; KIF5C, kinesin family member C1; MTHFD1L, methylenetetrahydrofolate dehydrogenase (NADPH-dependent) 1-like; TTK, TTK protein kinase;UBE2C, ubiquitin-conjugating enzyme E2C; PARK2, parkin RBR E3 ubiquitin protein ligase; EGF, epidermal growth factor receptor; FBXW7, FBXW7 protein kinase; ARNT, aryl hydrocarbon receptor nuclear translocator.
biological subtype with a specific marker. Functional studies in cell lines are also problematic, because it is clear that the available TNBC cell lines do not represent the spectrum of mutations and other disease-associated events seen in TNBC. Many of these cell lines are derived from patients with multiply pretreated metastatic disease and also carry the adaptations necessary for survival in vitro. This makes the translation of function in a model system to function in situ in patients very uncertain. Studies of patient-derived xenografts (PDX) at different stages of disease can overcome some of these limitations, but also suffer from clonal adaptation during engraftment. Representative TNBC PDX are still in their infancy and much larger numbers of characterized xenografts are required. A full discussion of this is beyond the scope of this review. With all of these caveats in mind, probably the single most important current observation concerning TNBC is that this group of patients is quite clearly not a uniform clinical or molecular entity. Progress will be difficult to come by until clinical studies and disease mechanism studies also subtype patient tumors and collect the materials on which to rigorously review and reevaluate “subtype,” as new information accrues. Uniform and consistent assays that measure mutations, expression, copy number, and methylation on FFPE materials would greatly advance the field by allowing cost-effective and robust genotyping of patients in clinical studies.

The Cell Types of Origin for TNBC

The phenotype of cancers can be understood as the product of transforming events (genetic and epigenetic), which take place on the background of a cell of origin. However, considerable controversy and uncer-
tainty exists as to the cell types of origin of TNBC. Mammary epithelial lineages (luminal, myoepithelial) are generated throughout life by a tissue hierarchy that comprises long-lived, self-renewing, bilineage stem cells, as well as bilineage, luminal-restricted, and myoepithelial-restricted progenitor cells that occupy an intermediate position in the hierarchy (20–26). The cells with bilineage potential (stem cells and bilineage progenitors), as well as myoepithelial-restricted progenitors, share many phenotypic markers with basally located myoepithelial cells (e.g., α6-integrinβ1-integrinEpCAM+flow), whereas progenitor cells committed to the luminal lineage (both ER+ and ER−) express markers that overlap with differentiated luminal cells (e.g., EpCAM+CD24+) (20–22, 25, 27). The persistence in breast and other cancers of dysregulated developmental hierarchies is supported by studies showing differential repopulating capacity of sorted cell fractions when implanted into immunodeficient mice (28–32).

Key unresolved questions are: how the developmental potential of mammary cells in which transforming mutations arise modulates the impact of those mutations, and conversely, how specific mutations perturb the regulation of the hierarchical process itself. Although such information is mostly lacking in the majority of human breast cancer subtypes, some progress has been made with studies of BRCA1-deficient tumors.

Lim and colleagues demonstrated that histologically normal preneoplastic breast tissue in human breast cancer 1, early onset (BRCA1)-mutation carriers (i.e., germline-heterozygous) contains increased proportions of α6-integrinEpCAM+ cells (a phenotype that enriches for luminal progenitors) and fewer cells in the α6-integrinEpCAM− fraction (mostly mature myoepithelial cells, but also stem cells, bipotent cells, and myoepithelial progenitors) (27). Furthermore, the BRCA1-mutated luminal progenitors are functionally perturbed, with enhanced growth factor–independent clonogenic properties in vitro. The gene expression signature of the luminal progenitor-enriched normal cell fraction correlated more with human basal-type breast cancers than with the other molecular subtypes (including luminal A and B subtypes, which showed the lowest correlation). Although suggesting retention of aspects of the luminal differentiation program within basal cancers, it is likely that the correlations are partly influenced by the proliferative nature of the normal luminal-progenitor-enriched fraction, a property shared by highly proliferative breast cancer subtypes, such as BRCA1-deficient/other TNBCs.

The interaction of developmental potential with oncogenic mutations, including BRCA1 loss, has also been investigated by targeting mutations experimentally to specific cell subsets. Luminal EpCAM+ cells from normal or BRCA1-mutation–carrying human mammary tissue were more readily transformed than basal CD10+ cells by a lentiviral oncogene cocktail of mutated p53, cyclinD1, activated PI3K, and oncogenic K-ras (33). Conditional BRCA1 deletion in p53+/− mice resulted in the development of tumors using either Blg-Cre or K14-Cre drivers (the former targeting mainly ER luminal cells, including luminal progenitors, the latter targeting basal cells, including stem cells). The histopathology of mouse tumors generated under the luminal-targeting Blg-Cre more closely resembled human BRCA1−/− tumors. Differences were less evident at the overall transcriptome level, with both sets of tumors correlating with the human basal subtype (34). The Blg-Cre BCRA1−/− tumors also closely resembled tumors previously reported in a K14-Cre BCRA1−/−p53−/− model (35), highlighting the sensitivity of such studies to factors such as genetic background and transgene strategy.

Deconstructing the functional interactions between complex mutational patterns and cellular developmental properties in an evolving tissue hierarchy remains challenging. The BRCA studies may be interpreted as supporting either luminal progenitor or stem cell hypotheses of origin, and analogous studies have not yet been carried out on non-BRCA-deficient tumors (the majority of TNBCs). Key to making progress may be the development of serial mutation approaches to model the genetic profile of different TNBC subtypes on different cell types of origin backgrounds, as well as research to deepen our understanding of the molecular basis of the properties that define stem and progenitor function.

Oomics-Based Molecular Subgroups of TNBC

Attempts to understand the disease groupings that make up TNBC have received some impetus in the last 10 years from genome-scale analysis of the transcriptome, genome, and epigenome. Within the last year, genome sequencing has added to our understanding. Also, functional screens in cell lines imply some molecular features not evident from the large scale omics. This is an evolving area in which many of the disease relationships remain tenuous; we review each, with a focus on recent genome studies.

TRANSCRIPTOMICS AND THE “BASAL” TNBC SUBTYPE

This aspect of TNBC has been extensively reviewed elsewhere (36). Early microarray studies defined a molecular subtype known as basal (from similarity to the transcriptome observed in basal myoepithelial cells of the breast) (15, 16), which has a strong overlap with TNBC status. These early studies have received support
from multiple independent studies (3, 9, 37–41) confirming basal-like expression characteristics as a reproducible grouping of patients, with discrete patient survival characteristics. When gene expression data are compiled from different sources, basal-like tumors in TNBC consistently cluster within a group that express CK5, 6, 14, 17, and caveolin1/2, as well as c-kit and P-cadherin (15, 42). Approximately 50% of transcriptome-defined basal-like tumors express EGFR. Genes associated with cell proliferation, such as marker of proliferation Ki-67 (MKI67) (also known as Ki67) and topoisomerase (DNA) II alpha 170kDa (TOP2A), are also highly expressed in basal-like tumors. In parallel with the “core” basal expression grouping, evidence from the distribution of somatic p53 and PI3K mutations, copy number analysis, as well as genome sequencing of TNBC (discussed below), shows that the basal expression cancers are the most prevalent/molecularly homogenous subgroup, with a distinct patient survival characteristic. Thus, basal expression subtype, p53-mutated breast cancers, representing approximately 60%–70% of TNBC, have the strongest evidence in favor of being a TNBC subgroup.

**TRANSCRIPTOMICS OF NONBASAL TNBC SUBTYPES.**

Transcriptomic studies of TNBC have suggested that other expression-based groupings may exist. The identification of tumors with expression patterns similar to mesenchymal cells has suggested the possible existence of such a subgroup, originally named “claudin-low” (low expression of claudin genes) (43), or EMT signature-enriched TNBC (44, 45). The mesenchymal expression signature has received support from independent studies. A recent metaanalysis of breast cancer transcriptome studies combined data from 21 publicly available datasets and 587 TNBC cases of pretreatment and posttreatment tumors (discovery set 386 cases, validation set 201 cases), and suggested 6 consistent TNBC subgroups: 2 basal-like (BL1, BL2), an immune signature group, a mesenchymal group, a mesenchymal stem-like group, and a luminal androgen receptor subtype (46). The exact molecular definition of “TNBC with a mesenchymal pattern” remains uncertain; different gene sets have been proposed in different studies. It is unclear whether these tumors have distinct biology in patients, although the expression patterns have been interpreted as support for epithelial–mesenchymal transition (EMT) differentiation in TNBC. For example, recently it has been shown in mouse-knockout models (47) that the combination of p53 deletion and c-met activation results in mesenchymal expression pattern mouse tumors. Taken together, there is evidence for an EMT/mesenchymal expression pattern among TNBC; it is less clear if this is a group of cancers with distinct biology in patients, or a reflection of differentiation state, and this grouping is not well defined.

**IMMUNE GENE EXPRESSION IN TNBC**

Several studies have reported distinct immune gene expression patterns in TNBC (43, 48–53). Some of these reflect tumors with inflammatory infiltrates, although it is not clear whether these form a distinct subtype. Immune signatures have also been associated with DNA damage as an intrinsic response (54, 55) and some TNBC may have patterns of genome damage resulting in this type of immune signature. The reports of outcome associations for immune-associated signatures and infiltrates have been contradictory. It is not clear that TNBC with inflammatory gene signatures could be considered a discrete subtype, but the presence of these signatures in some tumors and not others is a recognized phenomenon which is worthy of further study to determine whether some TNBC have an associated immune response or inflammatory response. A possibility that remains to be explored is that TNBC with intrinsic immune gene expression also have specific patterns of genome damage.

**ANDROGEN RECEPTOR PATHWAY IN TNBC**

Androgen receptor (AR) expression in breast cancers was first noticed in the mid-1980s as hormone receptor assays were developed (56) and the existence of a subset of ER⁻/PR⁻ breast cancers expressing AR was noted in 2006 (57), with the identification of a AR⁺ breast cell line MDA-MB453, which has androgen responsiveness. The existence of an AR-expressing cancer subset was also noted in a recent metaanalysis of microarray studies (46); nonrecurrent somatic mutations in some AR-responsive genes (although not AR itself) were noted in the recent sequencing of TNBC (58). These observations have led to interest in the possible efficacy of AR modulators in TNBC. A recently reported phase 2 trial of bicalutamide (59) concluded that 12% of ER⁻/PR⁻ breast cancers express AR, and a clinical benefit of 19% was observed for patients in the treatment arm of this study. This suggests that there may exist a group of AR pathway-activated TNBC, with obvious potential for therapeutic intervention. Additional trials and studies are required to determine the nature of AR activation/responsiveness and to define the relationships, if any, with other molecular features of TNBC.

**SOMATIC GENOMIC LANDSCAPE OF TNBC**

The results from next-generation sequencing of breast cancers, including whole genome, or whole exome DNA sequencing, indicate that mutations in breast cancers show strong subtype-specific segregation (42, 51, 58, 59, 60, 61). In the case of TNBC, the main dis-
P53 pathway. Mutations in TP53 are highly prevalent in basal-like TNBC. The reported prevalence varies, probably due to differences in classifying tumors and in part due to ascertainment, with 60%–80% of basal TNBC reported (18, 42, 58, 62–64) as having p53 mutations. Recently it has been suggested that basal tumors are enriched for p53 nonsense and frameshift mutations (as opposed to point mutations, which are activating). A frequent occurrence is loss of heterozygosity (LOH) affecting one allele of p53, which may combine with methylation/mutation leading to inactivation of the second allele. Analysis of allele-specific expression between the tumor RNA and DNA has identified TP53 with strong transcriptional allelic imbalance, preferentially expressing only mutant TP53 alleles (65).

PI3K pathway. Mutations in the PI3K pathway represent the next most prevalent somatic mutation (42, 51, 58, 60), and these tend to occur more frequently in the nonbasal TNBC, although this is not a strict relationship. Mutations in PIK3CA are approximately 10% for TNBC, although the overall mutation frequency in all breast tumors for PIK3CA (36%) is comparable with TP53 (37%). Copy number analysis also revealed significant focal amplifications in PIK3CA. PIK3CA mutations and copy number gains lead to the activation of the PI3K pathway. The recurrent PIK3CA E5454K mutation presents most exclusively in the luminal A subtype, rather than in basal TNBC. An alternative mechanism leading to activation of the PI3K pathway is the loss of phosphatase and tensin homolog (PTEN) through mutation, with a prevalence in TNBC of approximately 7.7%. Additional PTEN silencing may occur epigenetically, and thus combined, the PI3K pathway may be activated in approximately 15%–20% of TNBC, with enrichment in the nonbasal expression-type tumors. However, trials of PI3KCA agents have not yet shown a major impact, perhaps in part because of inadequate initial patient selection. PIK3CA/PTEN-mutated TNBC tumors tend to group with better prognosis breast cancers in the revised IntClust groupings (18), rather than with the basal expression subtype. Another aspect of PI3K-Akt activity that may predominate in basal-expression type TNBC is the regulation of phosphatase expression. The phosphatase INPP4B has recently been implicated as an important regulator of PIP3 signaling, and expression of full-length INPP4B is very low in a majority of basal-expression subtype TNBC (66, 67). This may be the result of LOH at the locus, or the constitutive expression of the short isoform of INPP4B (68), which lacks a phosphatase domain in cells with a basal-like expression program.

BRCA1 AND BRCA2

BRCA1 or breast cancer 2, early onset (BRCA2) germline mutation carriers have a high probability (70%–80%) of developing breast cancer (69) and a higher risk for ovarian cancer (30%–50%). The relative risk for developing other types of cancer is also greatly increased. Whereas breast cancers from BRCA2 carriers are frequently ER+, tumors from BRCA1 carriers tend to be TNBC. However, if care is taken to eliminate previously undiagnosed BRCA1 germline carriers, somatic mutations in BRCA1 are infrequent in sporadic TNBC patients (42, 58).

TNBC patients may inactivate the BRCA1/2 pathways by methylation or downregulating BRCA function. For example, the protein ID4 is overexpressed and significantly unmethylated in TNBC, possibly accounting for repression of BRCA1 at the protein level (70). Once germline carriers are accounted for, nonmutational loss of expression may occur in perhaps 10%–15% of TNBC (42, 58). Mutational patterns suggest that recombination defects may occur even in non–BRCA mutated breast cancers (71), indicating the possible relevance of the homologous recombination (HR) pathway (72). Convincing evidence of functional loss of HR-repair capacity in a significant number of sporadic TNBC patients is lacking, and the support from therapeutic responses to drugs targeting these pathways in TNBC is mixed. A recent phase 2 trial of a poly ADP ribose polymerase (PARP) inhibitor in sporadic ovarian and TNBC (73), showed no activity in the TNBC arm, suggesting “BRCAness” may not be a frequent phenotype in TNBC. In contrast, clinical reports of responses to platinum salts (74) may suggest that some TNBC patients have a functional deficiency in HR, which could be exploited. It is possible that the nonmutational mechanisms result in hypomorphic BRCA function, sufficient to moderately impair HR, but not sufficient to confer sensitivity to agents targeting this deficiency. It is hard to estimate how many TNBC patients fall into this group, but it is probably not more than 20% of all TNBC. Larger unbiased studies of patients with TNBC that clearly distinguish between germline and somatic mutations and that account for all other possible mechanisms of inactivation recombination repair are required.

ADDITIONAL SOMATIC MUTATIONS IN TNBC.

Other significantly mutated single genes reported in TNBC include: USH2A, myosin IIIA (MYO3A), PTEN, retinoblastoma 1 (RB1), ataxia telangiectasia and Rad3 related (ATR), ubiquitin protein ligase E3 component
n-recognin 1 (UBR1), collagen, type VI, alpha 3 (COL6A3), and several well-known oncogenes [v-raf murine sarcoma viral oncogene homolog B (BRAF), neuroblastoma RAS viral (v-ras) oncogene homolog (NRAS), v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (ERBB2), and v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3 (ERBB3)] (18, 42, 58). Activating mutations of ERBB2, rare in TNBC, may present an opportunity for intervention with targeted kinase inhibitors (75). The majority of the significantly mutated genes in non-TNBC subtypes, except TP53 and PIK3CA, are only rarely observed in basal-like TNBC, indicating a subtype-specific mutational spectrum.

The distribution of allele prevalence in TNBC has a long tail; most somatic mutations are observed once, and current statistical methods are not well suited to assessing the relevance of these mutations. Prevalence-based mutation assessment in a patient population is weighted toward the discovery of early founder mutations that dominate the clonal structure of a tumor. Approaches for deciphering somatic mutations implicated in progression are less well established. One approach is to group mutations by virtue of the known functions of the genes in which they occur, using systematic network analysis. Another approach is to search for mutated genes with a disproportional effect on the transcriptional network known to associate with the gene in question (76). Using these approaches and combining copy number effects and nonmutational alterations in expression, additional genes have been implicated in TNBC, including the retinoblastoma 1 (RB1), v-myc avian myelocytomatosis viral oncogene homolog (MYC), and rat sarcoma viral oncogene homolog (RAS) gene networks. RB1 may have special significance in potential contributions to the cell cycle abnormalities and genomic instability in TNBC (77–79).

The cooccurrence of dysfunction in these networks with respect to the basal/nonbasal phenotype is less clear. Rb loss of function has been associated with both luminal and basal tumors (77, 78). Overexpression of MYC has been noted in breast cancers, and frequent amplification of the c-myc locus on chromosome 8 has been observed, although this does not always result in increased expression (18). Some studies have suggested that Ras activation is functionally important in breast cancer, however Ras is not frequently affected (<5% of all breast cancers (42, 51, 58, 60, 61)) by oncogenic mutational mechanisms in breast cancers. Better-controlled studies are needed to define the subtype-associated relevance of these pathways.

Recent genome sequencing of primary TNBC (58) revealed a distributed pattern of somatic mutations in genes implicated in cell shape and motility, and the extracellular matrix. No single gene was mutated at >5% prevalence, with the exception of MYO3A; however, the gene family and biological process was significantly overrepresented. The extracellular matrix plays a key role in the maintenance of the tumor microenvironment and contributes to tumor progression/metastasis through the promotion of angiogenesis, tumor-associated inflammation, and EMT. Actins are also involved in this process and implicated in invasion/survival processes. Whether this pattern represents functional mutations in this family of proteins, or some other process, such as transcription-dependent repair/mutation, is unknown.

COPY NUMBER AND CHROMOSOMAL LEVEL GENOME ABERRATIONS

TNBC are characterized by extensive genomic instability. Basal expression-type TNBC have significantly more genomic instability than nonbasal TNBC (18, 58, 80, 81). This may be a reflection of early p53 inactivation, or aberrant expression of c-myc by nonmutational mechanisms; however, mutation or aberrant expression of genes implicated in cell cycle checkpoints (RB1), cell division mechanisms (cyclin amplification), and genome maintenance (homologous recombination, PTEN loss), could all contribute to the landscape.

High genomic instability in basal-like tumors at the chromosome structural level has been reported (82, 83). Recurrent large-scale chromosomal changes have been noted for basal cancers, with loss of 5q and gain on 10p being somewhat characteristic of basal cancers. 5q loss may be associated with basal-specific transcription patterns, as demonstrated by in trans analysis of copy number and expression (18). This method searches for expression relationships to copy number at other than the canonical gene or region (defined as in cis). In the recent reclassification of breast cancers on the basis of copy number expression analysis, the 5q loss is associated with a strong basal-specific signature, including aurora kinase B (AURKB), B-cell CLL/lymphoma 2 (BCL2), BUB1 mitotic checkpoint serine/threonine kinase (BUB1), cell division cycle associated 3 (CDCA3), cell division cycle associated 4 (CDCA4), cell division cycle 20 (CDC20), cell division cycle 45 (CDCA45), checkpoint kinase 1 (CHEK1), forhead box M1 (FOXM1), histone deacetylase 2 (HDAC2), insulin-like growth factor 1 receptor (IGF1R), kinesin family member 2C (KIF2C), kinesin family member C1 (KIFC1), methylene tetrahydrofolate dehydrogenase (NADP+ dependent) 1-like (MTHFD1L), RAD41AP1, TTK protein kinase (TTK), and ubiquitin-conjugating enzyme E2C (UBE2C). The precise 5q driver gene(s) remain unknown, but this recurrent chromosomal loss requires further investigation to define the basal-specific signature (17). Mono-
allelic expression associated with LOH is frequent in TNBC and has been mapped to high resolution (84). Recurrent focal amplifications and deletions affecting gene expression are uncommon in TNBC, but a few have been observed. Parkin RBR E3 ubiquitin protein ligase (PARK2) was reported as a tumor suppressor in colorectal cancer and is a newly discovered tumor suppressor in TNBC. Intragenic deletion in this gene is found in 6% of the TNBC tumors. Other genes affected by copy number alteration events are RB1 (5%), epidermal growth factor receptor (EGFR) (5%), fibroblast growth factor receptor 2 (FGFR2) (3%), and phosphatase and tensin homolog (PTEN) (3%) (7, 18, 58, 81, 84, 85). For patients with these uncommon events, the tumor phenotype may be skewed by the amplicons/deletion and some (e.g., EGFR, FGFR2) may represent “actionable” events. Different mechanisms likely shape the mutation landscape and copy number variation as these are uncorrelated in extent.

Individual breast cancers of all subtypes average approximately 5 nonrecurrent fusion genes per tumor sample (86), and this is even more prevalent in TNBC. However, recurrent translocation/fusions are not a feature of breast cancer, although reports of rare recurrent fusions with Notch (86) and Akt3 (87) have been described. In TNBC, Akt3-MAGI3 fusions have been reported to upregulate the expression of AKT3, while removing the PDZ [post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1)] domain of MAGI3, required for PTEN binding. This translocation leads to a combined effect of overexpression oncogene (AKT3) and loss of tumor suppressor (60), predicting activation of the PI3K pathway.

**EPIGENETIC REGULATION**

On the basis of DNA methylation analysis in 802 breast tumors, TCGA breast tumors are classified into 5 methylation groups. Group 5-represented (basal-like tumors) has the lowest level of DNA methylation, and group 3-represented (luminal B subtype) showed hypermethylated phenotype. A study of 91 TNBCs found significant hypomethylation of stem cell markers CD44, CD133, and MSH1, confirmed by IHC (87). The functional relationship of global CpG hypomethylation to the basal TNBC phenotype remains to be explored.

**MicroRNAs**

MicroRNA expression correlates with cell differentiation and tumor developmental status. It also reflects the specific pathological features of breast cancer, such as ER and PR expression, lymph node status, vascular invasion, proliferation and tumor stage. miR-21 and miR-221, which play roles in cell growth and proliferation, are significantly higher in TNBC vs normal tissue. miR-221 also inhibits the expression of ER at the protein level. miR-210, which increases the expression of hypoxia-inducible factors, hence promoting tumor progression in hypoxia, is overexpressed in TNBC. By contrast, microRNAs that suppress cell proliferation, such as miR-145 and miR-205, are underexpressed in TNBC. miR-205 also inhibits the EMT transition and suppresses tumor expansion from the basal membrane to the stroma (88). Thus, the expression profile of microRNA associates with the highly proliferative and high EMT feature of some TNBC.

**CLONAL STRUCTURE WITHIN PRIMARY TNBC**

Next-generation sequencing methods have opened up the estimation of clonal composition in human tumors (reviewed in (89). Recently, the clonal complexity of primary TNBC was assessed for the first time using these methods (58). Using population prevalence methods of clonal estimation, the number of distinct clonal groups in primary TNBC was noted to vary widely from patient to patient at the time of diagnosis. The tumors were of a similar stage and grade, and all pT2 or smaller. These patients were treated with broadly similar regimens, yet their tumors varied in complexity from 2–3 discernable clonal populations to 7+. Basal expression-type TNBC were more clonally complex than nonbasal, in keeping with the observation that they have a higher mutational burden. Analysis of clonal prevalence (abundance of each clonal group) by gene function showed that most “founder” type mutations, such as p53, were clonally prevalent, whereas somatic mutations in genes functionally associated with progression phenotypes (such as invasion, discussed above), were of lower clonal prevalence. In some cases, it was noted that founder mutations such as p53 were not the most prevalent clone, suggesting that earlier mutations may have provoked the genomic instability in some tumors. Future analysis of clonal abundance of low-prevalence mutations in relation to relapse/progression will be needed to determine whether clonal analysis can become predictive of relapse potential in primary cancers.

**Functionally Implicated Genes and Pathways**

Beyond the genes and pathways directly implicated by mutational mechanisms or epigenetic regulation, additional pathways of potential relevance to TNBC have been reported. The relationship of these to TNBC subtypes remains mostly tenuous, as the majority of studies are either small (relative to the diversity of the patient population) or have not been rigorously tested against the omics-based subgroups. The functional
properties of TNBC have been reviewed extensively elsewhere and are only summarized here.

**RAF-MEK1/2-ERK1/2 PATHWAY**
Activated RAF-MEK1/2-ERK1/2 signaling has been noted in TNBCs, although mutations are seen in <5% of cases. Clinical trials with MEK inhibitors have failed because of drug-induced activation of alternative survival signaling pathways (90). MEK inhibitor-induced RTK stimulation overcame MEK2 inhibition, reactivating ERK and resulting in drug resistance. In cell-based assays, the combination of the MEK inhibitor AZD6244, and the RTK inhibitors sorafenib or foretinib, are synthetic-lethal.

**PTPN12 AND OTHER PHOSPHATASES.**
Tyrosine phosphatase PTPN12 is a putative tumor suppressor, and loss of PTPN12 protein expression is prevalent in TNBC (91). PTPN12 inhibits the EGFR/HER2-mitogen-activated protein kinase signaling pathway to suppress tumor transformation. PTPN12 depletion leads to increased phosphorylation of EGFR and HER2, and multiple RTKs in TNBC. Combinatorial inhibition of HER2 and PDGFR-β by lapatinib and sunitinib greatly reduces the growth of TNBC, suggesting that PTPN12-deficient TNBC may be successfully treated with a combination of TK inhibitors. It also demonstrates that in TNBC, the HER2 pathway can be activated at the phosphorylation level by PTPN12 deletion without HER2 amplification. It is noteworthy that the 5q deletion in TNBC, which drives a large in trans expression network (see above), encompasses several phosphatases.

**EGFR PATHWAY**
EGFR high-level expression is detected (27%–57%) in TNBC by IHC (92). Mutations in EGFR are associated with extreme changes in transcription of interacting genes (58). Although TK inhibitors have been successfully used to treat EGFR-mutant lung cancers, the effectiveness of EGFR-targeted agents in combination with chemotherapy in TNBC has been inconsistent and trials are ongoing (93, 94).

**NON-BRCA DNA DAMAGE REPAIR PATHWAY**
A number of genes in the DNA damage response pathway [ataxia telangiectasia mutated (ATM), ATR, checkpoint kinase 2 (CHEK2), and ubiquitin protein ligase E3 component n-recogin 5 (UBR5)] are implicated at low prevalence by mutational mechanisms, but possibly more frequently by altered expression in TNBC. Because of the importance of these genes in DNA damage repair, the primary rationale for drug design is targeting these genes in combination with the presence of genotoxic agents to increase the effectiveness of chemotherapy. Synthetic lethality is another strategy currently under investigation for tumors with loss of function mutations or hypomorphic alleles in DNA damage repair.

**ANGIOGENESIS, HYPOXIA, AND SERINE ADDICTION**
Angiogenesis is a hallmark of tumor development. TNBC express higher levels of vascular endothelial growth factor (VEGF), the predominant mediator of angiogenesis (95). Acting as a growth factor ligand, VEGF binds to the RTKs VEGF receptor 1 (VEGFR-1) and VEGFR-2. Biological options for targeting angiogenesis include: an anti-VEGF monoclonal antibody, such as bevacizumab, and RTK inhibitors, such as the sunitinib and sorafenib. Hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor) (HIF1A)/aryl hydrocarbon receptor nuclear translocator (ARNT) pathways, required for angiogenesis, are also activated in basal-like tumors, suggesting that HIF1 inhibitors or bioreductive drugs that become activated under hypoxic conditions may benefit patients with basal-like cancers.

Growth factor pathway activation and hypoxic tumor cell metabolic responses have been noted in TNBC, and predominantly in basal expression type cancers. This is reviewed elsewhere, but probably also forms an important metabolic consequence of transformation (96–98). Targeting hypoxia with CA9 inhibitors is being explored. Serine biosynthesis may be another key feature, with addiction to serine and its metabolites being a reported phenotype of cell lines with basal-expression phenotypes (99).

**Conclusions**
Integration of omics data on TNBC is beginning to shed some light on the constellation of pathologies associated with these breast cancers. Better models, which are rigorously mapped at the omic level and that relate the actual behavior and phenotypes of TNBC in patients, are needed. It is realistic to believe that more robust molecular hypotheses will be developed on the basis of this information and will guide further focused study to better understand TNBC and target these cancers.

**Author Contributions:** All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.


77. Ha G, Roth A, Dai L, Bashashati A, Ding J, Goya R, et al. Integrative analysis of genome-wide loss...


