MOLECULAR ONCOLOGY XXX (2012) 1–13

available at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/molonc

Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy

Nicole I. Park[a], Peter K. Rogan[b,c], Heather E. Tarnowski[a], Joan H.M. Knoll[a,d,*]

[a]Department of Pathology, Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON, Canada
[b]Department of Biochemistry, Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON, Canada
[c]Department of Computer Science, University of Western Ontario, London, ON, Canada
[d]Molecular Pathology, Laboratory Medicine Program, London Health Sciences Centre, ON, Canada

ARTICLE INFO

Article history:
Received 5 August 2011
Received in revised form
28 November 2011
Accepted 2 January 2012
Available online ■

Keywords:
Breast cancer
Genomic stability
Chemotherapy
Genetic selection
Array comparative genomic hybridization
Gene expression
Bioinformatics
Fluorescence in situ hybridization

ABSTRACT

Background: Cancer genomes accumulate frequent and diverse chromosomal abnormalities as well as gene mutations but must maintain the ability to survive *in vivo*. We hypothesize that genetic selection acts to maintain tumour survival by preserving copy number of specific genes and genomic regions. Genomic regions and genes that remain unaltered in copy number and expression, respectively, may be essential for maintaining tumour survival.

Methods: We analyzed copy number data of 243 previously reported breast tumours and computationally derived stable copy number regions. To identify genes in stable copy number regions with nominal changes in expression, datasets for tumour and normal samples were compared. Results were replicated by analysis of a series of independent copy number, expression and genomic sequencing studies. A subset of stable regions, including stable paralogous regions, were confirmed by quantitative PCR and fluorescence in situ hybridization (FISH) in 5 breast cancer cell lines. We deduced a comprehensive set of dually stable genes (i.e. maintaining nominal copy number and expression) which were categorized according to pathway and ontology assignments. The stability of genes encoding therapeutic drug targets was also assessed.

Results and Conclusion: Tumour genome analysis revealed 766 unstable (amplified and/or deleted) and 812 stable contiguous genomic regions. Replication analysis of an independent set of 171 breast tumours confirmed copy number stability of 1.3 Gb of the genome. We found that 5804 of these genes were dually stable. The composition of this gene set remained essentially unchanged (<2% reduction) after accounting for commonly mutated breast cancer genes found by sequencing and differential expression. The stable breast cancer genome is enriched for cellular metabolism, regulation of gene expression, DNA packaging (chromatin and nucleosome assembly), and regulation of apoptosis functions. Stable genes participating in multiple essential pathways were consistently found to be targets of chemotherapies. Preservation of stable, essential genes may be related to the effectiveness of certain chemotherapeutic agents that act on multiple gene products in this set.

* *Corresponding author*. Department of Pathology, Schulich School of Medicine and Dentistry, University of Western Ontario, Dental Sciences Building, Rm 4044, 1151 Richmond Street, London, ON, N6A 5C1 Canada. Tel.: +1 519 661 2111x86407.
E-mail address: jknoll3@uwo.ca (J.H.M. Knoll).

1574-7891/$ – see front matter © 2012 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.molonc.2012.01.001

1. Introduction

Mutation studies establish which essential gene products are critical for growth, and development of tumours. Despite extensive genomic instability, presumably, a minimal set of gene products are required for tumour cell survival. Loss-of-function mutations required for the proliferation and survival of cancer cells have been investigated using RNA interference (Ngo et al., 2006; Silva et al., 2008). Functional genetic analyses have identified causal cancer genes and much effort has been made to determine their contribution to the tumourigenic phenotype.

Human cancers arise from the accumulation of numerous genetic and epigenetic alterations, which lead to dysregulation of protein-coding genes and interacting genes within a pathway (Schafer et al., 2009). Microarray studies assess abnormalities in copy number of specific genes (Hicks et al., 2006), expression (Perou et al., 2000), and methylation status (Feinberg and Tycko, 2004; Widschwendter and Jones, 2002). Genomic rearrangements, deletions, amplifications, and point mutations of genes regulating cell growth, apoptosis and DNA repair are responsible for unregulated proliferation (Vogelstein and Kinzler, 2004). As well, alterations in oncogenes and tumour-suppressor genes contribute to tumourigenesis (Davies et al., 2002; Friedberg, 2003; Nowell, 2002; Santarosa and Ashworth, 2004). Common targets for amplification and deletion include *ERBB2*, *MYC*, *CDKN2A*, *PTEN*, and *SMAD4* (Collins and Groudine, 1982; Hahn et al., 1996; Kamb et al., 1994; Li et al., 1997; Slamon et al., 1987; Steck et al., 1997). In breast tumours, genomic regions that are consistently abnormal have been termed "saw-tooth" and "firestorm regions" because they possess the highest frequencies of gains and losses of genomic sequences (Hicks et al., 2006). However, investigation of genes with little or no variation in copy number or expression has not been a focus of cancer studies, even though they may also contribute to maintenance of the tumour phenotype.

Confronted with frequent chromosome instability and gene mutation, some tumour cell lineages are surprisingly resilient to autophagy and apoptosis. We investigate the composition of the stable gene set in breast tumours which presumably contributes to their survival, regardless of whether they are derived from cancer stem cells or from source cells that have avoided inactivation of essential genes. We have characterized regions of breast cancer genomes that share stable copy number (Chin et al., 2007; Hicks et al., 2006) and exhibit levels of expression similar to matched normal tissues (Turashvili et al., 2007; Naderi et al., 2007). We address whether these stable regions encode essential gene products by determining if standard breast cancer chemotherapies kill cancer cells by depriving tumours of these functions.

2. Materials and methods

2.1. Definition of stable and unstable genomic regions in the breast cancer genome

Copy number and expression were analyzed from independent array comparative genomic hybridization datasets (aCGH): by a Representational Oligonucleotide Microarray Analysis (ROMA; GEO GPL7313) and a custom 30K 60-mer oligonucleotide array (GEO GPL5737). The ROMA platform contained approximately 85,000 probes with an approximately uniform genomic distribution (Lisitsyn et al., 1993; Lucito et al., 2003). The data consist of 2847 probes that detected autosomal deletions and amplifications in 243 primary breast carcinoma tissues (Hicks et al., 2006). The custom array contained 60-mer oligonucleotides representing 28,830 unique genes (van den Ijssel et al., 2005). In this aCGH platform, 1684 highly recurrent altered regions were found in 171 primary breast tumours (Chin et al., 2007).

Autosomal variations in copy number among multiple tumours were determined relative to a normal diploid male DNA (Hicks et al., 2006) or to a reference pool of 50 randomized tumours (Chin et al., 2007). In both studies, at least 10% of the tumours were required to display a consistent increase or decrease of at least one copy of the target locus. Neither study (Chin et al., 2007; Hicks et al., 2006) controlled for tumour subtype or heterogeneity. The requirement for ubiquitous genomic stability across all breast tumour subtypes is expected to identify common genomic intervals that are essentially unaltered in the most prevalent types of tumours (however, conclusions about stability in individual subtypes may not be valid).

ROMA probe IDs were ordered by genomic coordinate and hybridization copy number (based on NCBI Build 36/hg18 assembly). Adjacent probes with identical copy numbers (either increased or decreased) were grouped to form contiguous intervals with the same unstable genotypes. This same approach was taken to cluster intervals of gains or losses less than 105 kb apart using an independent dataset (Chin et al., 2007).

Stable genomic intervals were inferred by complementing genomic coordinates of clusters of tightly linked unstable intervals. Genes located within stable and unstable regions were determined by convolving the genomic coordinates of all known protein-coding genes (CCDS; build Hs36.3) with those of stable and unstable regions using the Galaxy metaserver (http://main.g2.bx.psu.edu). The history and results of the operations used to derive this and other genomic datasets are available on our laboratory website (http://cancer.cytognomix.org/stable). Genes that overlapped the interface between adjacent stable and unstable intervals were classified as unstable. Although conservative, this approach avoided false assignments of unstable regions as stable. Stable chromosomal regions were further characterized by comparing the cumulative stability across each chromosome to the frequency of recurrent cytogenetic abnormalities in breast cancer (Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer; http://cgap.nci.nih.gov/Chromosomes/Mitelman; $n = 5328$, Sept 2010 version).

2.2. Gene expression analysis with copy number data across breast cancer subtypes

To identify genes in stable regions with nominal expression, datasets for tumour and normal samples were compared. These consisted of either matched tumour and normal

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

marginal tissues from the same individual (Turashvili et al., 2007) or individual tumours compared with average expression in a pooled set of 50 tumours (Naderi et al., 2007). These comparisons were done to mitigate inter-individual sources of variability in differential gene expression.

Differentially expressed genes (the highest and lowest 10%) present in tumours relative to normal matched tissues in the Turashvili et al. (2007) data were identified with Oncomine (https://www.oncomine.org/). These genes were then categorized with Galaxy according to their genomic stability using data from Hicks et al. (2006). These results were replicated by analyzing the distribution of stable genomic intervals in an independent set of 113 matched primary breast tumours, for which both expression (Naderi et al., 2007) and copy number abnormalities (Chin et al., 2007) were available (Array Express: E-UCON-1). However, Chin et al. (2007) and Naderi et al. (2007) used a lower resolution microarray which detects genomic sequences at a lower density. Over- and under-expressed genes were similarly eliminated with GeneSpring GX Software (Agilent). *P*-values of log2 ratios (fold change) were adjusted using the Benjamini-Hochberg multiple testing correction ($p < 0.05$). The sets of stable genes and genes with nominal expression were indexed by HGNC (HUGO [Human Gene Organization] Gene Nomenclature Committee) symbol and joined with the Galaxy metaserver to identify common members of both sets. This set of stable genes should be independent of differentially genes expressed in a series of unrelated breast tumours. Subsequently, we analyzed differentially expressed genes in 5 independent series of paired breast tumour and matched normal specimens using the Student's *t* test ($p < 0.01$; GEO datasets GDS2739, GDS3716, GDS3324, GDS3139, GDS2635); and compared these results with the stably expressed gene set deduced from Turashvili et al. (2007) and Hicks et al. (2006). The deduced dually stable gene set was also compared with the spectrum of commonly found mutations that either abolish gene function or cause rearrangement detected by DNA sequencing of candidate cancer genes and breast tumour exomes. The data sources included the most prevalent genes mutated in breast tumours in the COSMIC database (http://www.sanger.ac.uk/perl/genetics/CGP/cosmic?action=byhist&sn=breast&s=3), the Cancer Gene Consensus database (http://www.sanger.ac.uk/genetics/CGP/Census/)(as of 10/12/2011) and from high throughput genome sequences of breast tumours (Stephens et al., 2009).

2.3. *Pathway and gene ontology analysis*

Protein-coding genes residing in stable regions were assigned to biochemical pathways with Webgestalt (http://bioinfo.vanderbilt.edu/webgestalt). Significant KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways enriched in stable genes in the complete genome (http://www.genome.jp/kegg/), were identified based on a hypergeometric test ($p < 0.01$) for enrichment scores $R > 1$ (Supporting Information Table S1). Genes were colour-coded with the Gene Map Annotator & Pathway Profiler (http://www.genmapp.org) based on whether the copy number was decreased, increased or diploid (eg. Figure 4, Supporting Information Figure S5).

Previously annotated cancer pathways (KEGG: hsa05200) were compared to those enriched in the stable gene set. Stable cancer-related genes were classified according to gene ontology [GO] (Ashburner et al., 2000) and relevant GO terms were annotated using the Database for Annotation, Visualization

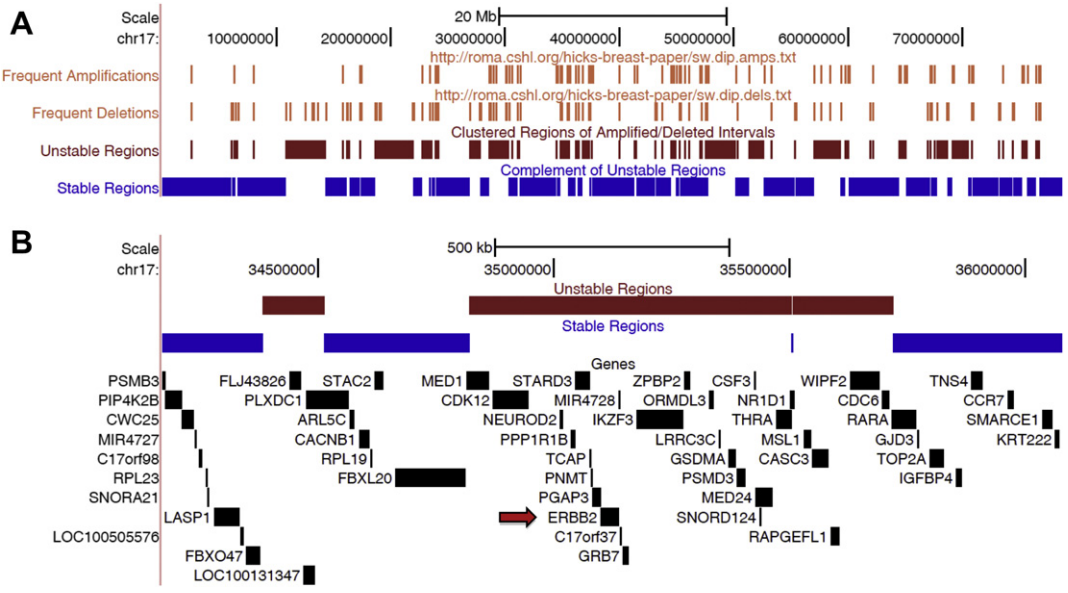

Figure 1 – Unstable and stable genomic regions mapped to chromosome 17. (A) Probes that detected frequent amplifications and deletions are displayed (orange) along the entire chromosome 17. Merging genomic intervals corresponding to these probes formed unstable regions (red). Stable regions (blue) were derived from the complement of unstable regions. (B) Unstable regions, stable regions, and genes (black) are displayed for chromosome 17(q12q21.1). The red arrow indicates the location of *ERBB2*, a gene within an unstable region and amplified in ~25% of breast cancers.

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

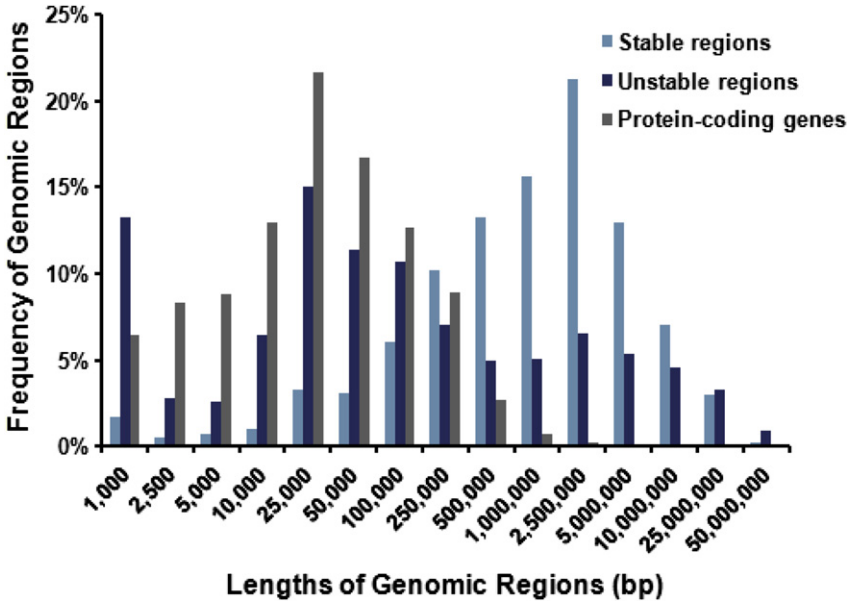

Figure 2 – Frequency and length distributions of unstable and stable genomic regions and protein-coding genes. Histogram indicates that the majority of unstable regions (dark blue) are ≤250 kb in length (bin 250,000 bp) and stable regions (turquoise) are longer than 250 kb (bin 250,000 bp). The *Y*-axis represents the number of stable or unstable regions, and genes as a percentage from each group. The *X*-axis depicts genomic lengths binned in units of 1000 bp. In comparison to stable regions, protein-coding genes (grey; consensus coding sequence project) have a smaller size distribution similar to that of unstable regions.

and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov). The Benjamini-Hochberg correction was applied to adjust for false positive stable gene assignments.

2.4. *Genes in stable genomic regions as targets for breast cancer therapy*

Stable gene products involved in metabolism of breast cancer therapeutic agents were retrieved from the Pharmacogenomics Knowledge Base (http://pharmgkb.org). For each agent, the KEGG pathways and the number of targets associated with stable genes were tabulated. The KEGG and PharmGKB results were linked to relate stable gene products with associated drugs and diseases. This process was repeated for all pathways that were significantly enriched ($p < 0.01$) in the KEGG table of drug targets. We also determined which drugs targeted the largest number of stable genes in the significantly enriched pathways.

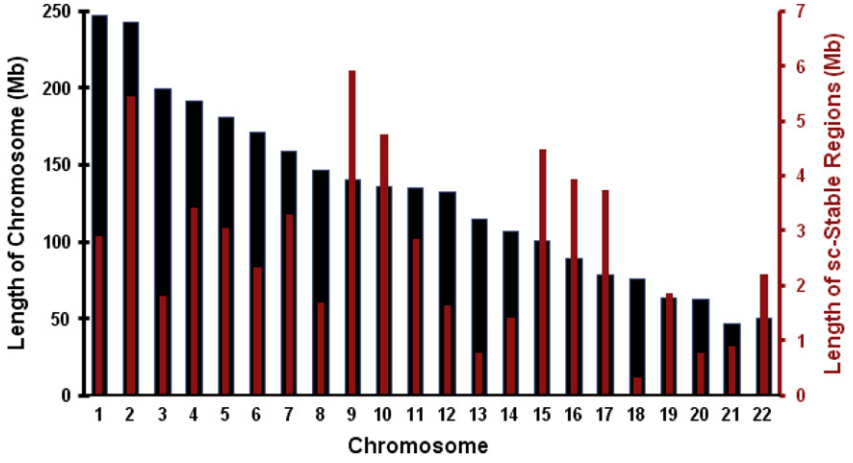

Figure 3 – Length distribution of single-copy (sc) stable genomic intervals per autosome. In the histogram, the lengths of stable regions including those within segmentally duplicated regions (red) are compared with overall chromosome length (black). While all chromosomes contain stable regions, the distribution of stable material is not related to chromosome length ($r^2 = 0.401$). On a per nucleotide basis, there is a higher degree of preservation of sc-genomic stability on chromosomes 2, 7, 9, 10, 15, 16, 17, 19 and 22. Stable regions of chromosome 13 and 18 appear to have more repetitive DNA (as represented by diminished length of sc-intervals) than the other chromosomes.

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

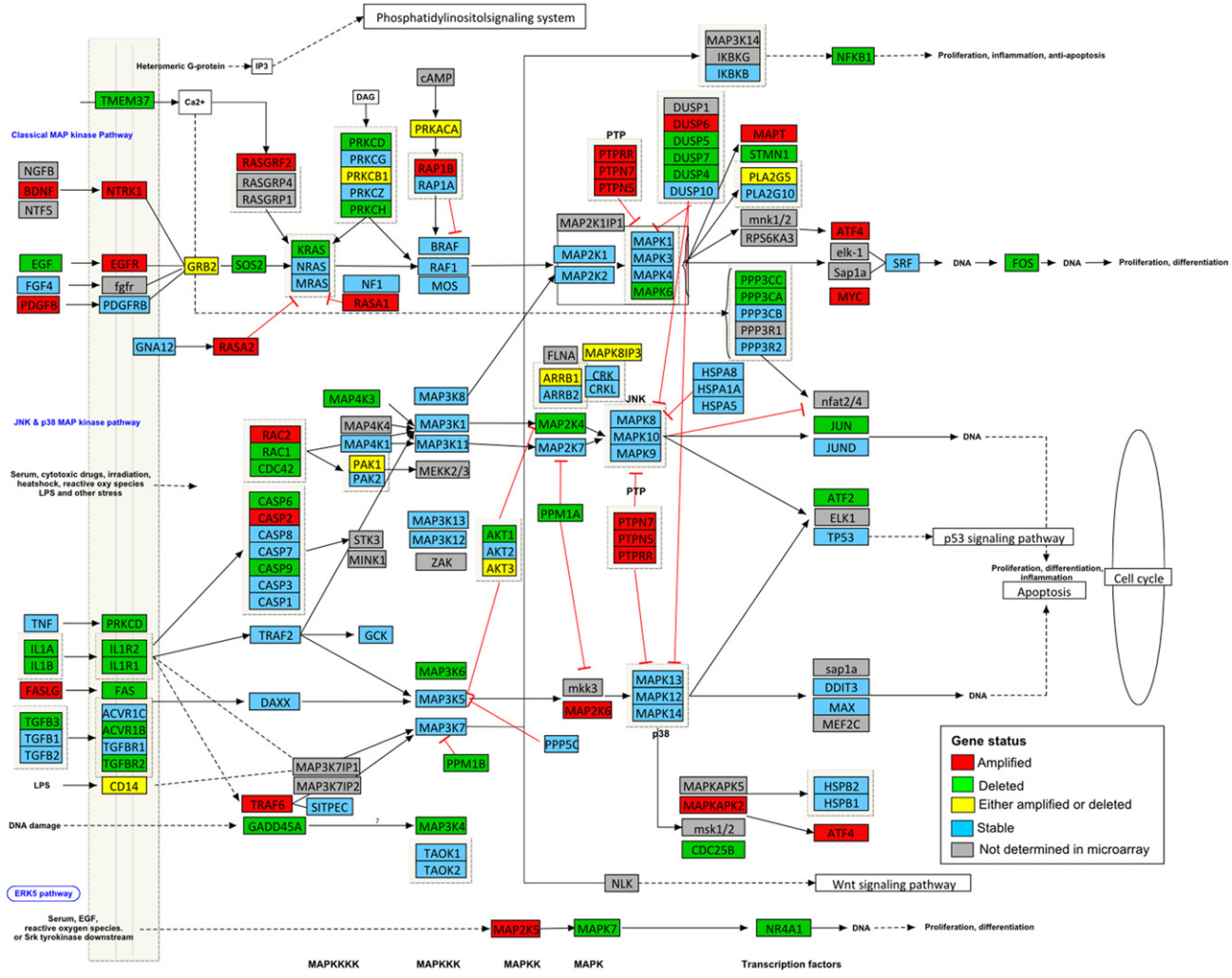

Figure 4 – MAPK signalling pathway with unstable and stable (copy number and expression) genes represented. The genes in the pathway are annotated as indicated in the figure. The preponderance of genes at the entry point (cell membrane) are either amplified and/or deleted, whereas stable genes are present downstream. This is consistent with abnormal signalling propagated through these downstream gene products. This figure was modified from the WikiPathways entry for MAPK. It used data from GENMAPP as indicated in the methods.

2.5. *Identification and confirmation of multi-copy sequences in stable regions*

Copy number changes were assessed in several multi-copy sequences predicted to be stable in breast cancer, since such changes would be more likely to occur in tumour genomes as a result of aberrant replication or recombination. Multi-copy, stable regions were identified using Galaxy by intersecting stable repeat-masked intervals with segmentally duplicated sequences in the genome (>500 bp long). The genomic copy number of each stable multi-copy interval was determined with mpiBLAST (www.mpiblast.org) and parsed to extract all divergent sequences with >70% and >80% similarity over >200 bp and >100 bp lengths, respectively. Multi-copy sequences which intersected unstable copy number sequence intervals in GEO GPL7313 were excluded from subsequent analyses. Stability of a subset of multi-copy regions from 5 different chromosomes was determined by quantitative PCR (Q-PCR) in 5 breast cancer cell lines [SKBR3 (Trempe, 1976), MCF7 (Soule et al., 1973), T47D (Keydar et al., 1979), HS578T (Hackett et al., 1977), and MDA-MB-231 (Cailleau et al., 1974)] and 10 control individuals (Supporting Information Table S2). *ERBB2* (chr17q12) which is amplified in SKBR3 (Xiao et al., 2009) served as a positive copy number control. The Pfaffl method (Pfaffl, 2001) was used to determine the copy number relative to diploid. Genomic stability was also assessed by fluorescence in situ hybridization (FISH) on metaphase chromosomes with synthetic multi-copy probes developed from stable regions (*DEFA1*-chr 8p23.1 and *P14KA*-chr 22q11.2, Supporting Information Table S2) (Khan et al., 2011; Knoll and Rogan, 2003) and BAC clones (Osoegawa et al., 2001) spanning stable chemotherapy gene targets from 3 different chromosomes: *FLT4* (RP11-179D12-chr 5q35), *GMPR2* (RP11-368G9-chr14q12), *CSF1R* (RP11-754J8-chr 5q33.1) (Knoll and Lichter, 2005). An ERBB2 BAC probe [RP11-94L15-chr 17p12,] in combination with a chromosome 17 centromere specific probe (CEP17) served as a positive control for copy number amplification (Wolff et al., 2007).

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

3. Results

3.1. *Stable regions in genomes of breast cancer patients*

We report stable regions in the breast cancer genome based on analysis of two independent datasets which mapped copy number alterations (Chin et al., 2007; Hicks et al., 2006), and filter the content of these regions for genes with nominal expression levels (see Section 3.2). From 2847 'representative or core' oligonucleotide probes that detected frequent deletions and amplifications in 243 primary breast carcinoma tissues, 766 contiguous genomic regions were derived and classified as unstable and 812 adjacent genomic intervals as stable (Figure 1 and Table 1).

Stable copy number regions comprise 56.3% of the entire genome or 1.6 Gb. The distribution of unstable genomic region lengths is on average smaller (<250 kb) than the stable genomic intervals (>250 kb). Typically, stable interval lengths ranged from 1 to 2.5 Mb (21.3%; 173 of total 812). The most frequent protein-coding genes in stable intervals were 10,001–25,000 bp in length (4310 of 19,856). Stable regions are therefore characterized by higher gene densities (Figure 2). Using the same approach with data from another study (Chin et al., 2007), we observed many of the same stable regions. That is, 828 unstable regions and 680 adjacent stable regions were deduced from the aCGH data. Of the 680 stable regions, 535 of them overlapped stable regions determined from Hicks et al. (2006), equivalent to 1.3 Gb of the genome (Table 1). Chromosomes with the largest fraction of stable regions (Figure 3) were associated with lower frequencies of cytogenetic abnormalities (for chromosomes 2, 9, 10 and 15: 3.1–3.5%) and those with lower stable region coverage had disproportionately higher frequencies of abnormalities (chromosomes 8, 13, 18, and 20) (Supporting Information Table S3).

Genomic architecture and paralogous structures often potentiate DNA rearrangements in tumour cells (Kolomietz et al., 2002; Pace et al., 2009). We found stable regions contain paralogous regions *in cis* that appear to be maintained as an ensemble. Multi-copy stable regions are annotated according to their genic (Table 2) and intergenic content. Multi-copy, paralogous sequences in stable regions (Figure 3) are present on every chromosome, but are proportionately less abundant on longer chromosomes and more common on shorter chromosomes. The highest percentages of paralogs are present on chromosomes 2 (9.1%), 9 (9.9%), 10 (7.9%), 15 (7.5%) and the lowest in chromosomes 13 (1.3%) and 18 (0.6%). Chromosome 18 was substantially underrepresented for stable regions relative to other F group chromosomes (ie. chromosomes 16 and 17) and chromosome 15 was overrepresented compared to other D group chromosomes (ie. chromosomes 13 and 14) (Supporting Information Table S3). The number of stable, *in cis* multi-copy sequences ranged from none on chromosome 1 to 356 on chromosome 16 (Supporting Information Table S4). Multi-copy families were more often distributed *in cis* (771) than *in trans* (360). There were 21 different multi-copy sequence families covering 5.7 Mb, with copy numbers ranging from 2 to 26 copies per haploid genome. Stability of 22, 18, 12, 6, and 3 copy sequence families *in cis* were tested in breast cancer cell lines and normal controls. Copy numbers of these sequence families were conserved in the cell lines by Q-PCR, with the exception of the 22-copy sequence family which was estimated to have 6–12 copies in HS578T and T47D (Supporting Information Figure S1). Members of this sequence family were 1234 bp in average length, organized *in cis* on chromosome 8 and distributed intergenically between *FAM90A* gene family members (Supporting Information Figure S2). *ERBB2* amplification in SKBR3 served as a positive control by Q-PCR (Supporting Information Figure S3) and metaphase FISH. The stability of other deduced multi-copy sequences (*DEFA1, P14KA*) was also demonstrated by metaphase FISH in all but one breast cancer cell line, despite differing modal chromosome numbers (Supporting Information Table S5; Figure S4).

Table 1 – Summary of aCGH datasets used to derive stable copy number regions.

	Dataset 1 (Hicks et al., 2006)			Dataset 2 (Chin et al., 2007)	
aCGH platform	ROMA (GEO: GPL7313)			Custom array (GEO: GPL5737)	
Total probe no.	85,000 (50-mer)			30,000 (60-mer)	
Copy number abnormalities	No. of probes			No. of regions	
	1772 (amp)	1864 (del)	782 (both)	1911 (gain)	1255 (loss)
Unstable regions[a]	766			828	
Stable regions	812			680	
Genes (CCDS) in stable regions[b]	9463			7692	
Stable regions common to both datasets	535 regions (1,348,553,559 bp)				
Stable genes common to both datasets	3859				

The number of probes and regions of frequent copy number abnormality are summarized by Dataset 1 and Dataset 2, respectively. Dataset 2 was used in our replication study. The numbers of unstable and corresponding stable regions are shown, as well as the number of protein-coding genes wholly contained within stable regions. The numbers of stable regions and protein-coding genes common to both datasets are summarized.

[a] Merged genomic coordinates of adjacent probes (Dataset 1) or regions (Dataset 2) with the same copy number abnormality (see Section 2); "amp" refers to amplification, "del" refers to deletion, and "bp" refers to base pair.

[b] Includes genes with and without corresponding Entrez IDs.

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

Table 2 – Stable copy number genes contained within *in cis* multi-copy regions.

Gene symbol	Gene description
3–5 Multi-Copy	
ABCC6	NM_001171: ATP-binding cassette, sub-family C, member 6; NM_001079528: URG7 protein isoform 2
ANAPC1	Anaphase promoting complex subunit 1
AQP7	Aquaporin 7
C2orf78	Hypothetical protein LOC388960
CCDC144A	Coiled-coil domain containing 144A
CCDC144B	Coiled-coil domain containing 144B
CCDC144C	Homo sapiens cDNA clone IMAGE: 4837395.
CNTNAP3	Cell recognition molecule CASPR3
DEFA1B	Alpha-defensin 1b
DEFA3	Defensin, alpha 3 preproprotein
DUB3	Deubiquitinating enzyme 3
FAM86B1	NM_001083537: hypothetical protein LOC85002; NR_003494: FAM86B1 protein
FLJ32679	Hypothetical protein LOC440321
FLJ36492	Homo sapiens cDNA, FLJ18738
GGT1	Gamma-glutamyltransferase 1 precursor
GOLGA6L1	Golgi autoantigen, golgin sub-family a-like
GOLGA8E	Golgi autoantigen, golgin sub-family a, 8E
GOLGA8G	Golgi autoantigen, golgin sub-family a, 8G
HERC2	Hect domain and RLD 2
HERC2P3	Putative uncharacterized protein (Fragment)
KIAA0393	Uncharacterized protein ENSP00000372404
LOC339047	Hypothetical protein LOC339047
LOC339240	Putative uncharacterized protein FLJ46089 precursor
LOC375133	Homo sapiens cDNA FLJ11279 fis, clone place1009444, highly similar to phosphatidylinositol 4-kinase alpha
LOC375133	Homo sapiens cDNA FLJ11279 fis, clone place1009444, highly similar to phosphatidylinositol 4-kinase alpha
LOC392196	Homo sapiens deubiquitinating enzyme 3 pseudogene (LOC392196), non-coding RNA
MBD3L2	Methyl-CpG binding domain protein 3-like 2
NOMO1	Nodal modulator 1
NOMO2	Nodal modulator 2 isoform 1
NOMO3	Nodal modulator 3
NPIP	Nuclear pore complex interacting protein
OVOS2	Ovostatin 2
PDXDC1	Pyridoxal-dependent decarboxylase domain
PI4KA	Phosphatidylinositol 4-kinase type 3 alpha
PKD1	NM_001009944: polycystin 1 isoform 1 precursor; NM_000296: polycystin 1 isoform 2 precursor
PLGLA	Homo sapiens plasminogen-like A (PLGLA), non-coding RNA
PLGLB1	Plasminogen-like B1
PLGLB2	Plasminogen-like B2
PRR20	Proline rich 20
RGPD1	RANBP2-like and GRIP domain containing 1
RGPD2	RANBP2-like and GRIP domain containing 2
RGPD5	RANBP2-like and GRIP domain containing 5 isoform
RGPD6	RANBP2-like and GRIP domain containing 6
RIMBP3	DKFZP434H0735 protein
RIMBP3B	RIMS binding protein 3B
RIMBP3C	RIMS binding protein 3C
RRN3P1	RNA polymerase I transcription factor homolog (*S. cerevisiae*) pseudogene 1
URG7	Homo sapiens up-regulated gene 7 (URG7) mRNA, complete cds
ZNF705D	Zinc finger protein 705D
6–8 Multi-Copy	
C2orf78	Hypothetical protein LOC388960
DKFZp434P211	Homo sapiens mRNA; cDNA DKFZp434P211
FLJ32679	Hypothetical protein LOC440321
GOLGA6L1	Golgi autoantigen, golgin sub-family a-like
GOLGA8E	Golgi autoantigen, golgin sub-family a, 8E
GOLGA8G	Golgi autoantigen, golgin sub-family a, 8G
LOC339047	Hypothetical protein LOC339047
NPIP	Nuclear pore complex interacting protein
PDXDC1	Pyridoxal-dependent decarboxylase domain
PKD1	NM_001009944: polycystin 1 isoform 1 precursor; NM_000296: polycystin 1 isoform 2 precursor
RGPD1	RANBP2-like and GRIP domain containing 1
RGPD2	RANBP2-like and GRIP domain containing 2

(*continued on next page*)

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

Table 2 – (*continued*)

Gene symbol	Gene description
RGPD5	RANBP2-like and GRIP domain containing 5 isoform
RGPD6	RANBP2-like and GRIP domain containing 6
USP18	Ubiquitin specific protease 18
9–11 Multi-Copy	
LOC339047	Hypothetical protein LOC339047
NPIP	Nuclear pore complex interacting protein
RANBP2	RAN binding protein 2
RGPD1	RANBP2-like and GRIP domain containing 1
RGPD2	RANBP2-like and GRIP domain containing 2
RGPD5	RANBP2-like and GRIP domain containing 5 isoform
RGPD6	RANBP2-like and GRIP domain containing 6
USP17	Ubiquitin specific peptidase 17
12–16 Multi-Copy	
LOC339047	Hypothetical protein LOC339047
NPIP	Nuclear pore complex interacting protein
NPIPL3	Nuclear pore complex interacting protein-like 3
PDXDC2	Homo sapiens cDNA: FLJ23482 fis, clone KAIA03142
RUNDC2B	RUN domain containing 2B

Multi-copy intervals are categorized by copy number, ie. 3–5, 6–8, 9–11 and 12–16 copies. More genes are associated with the lower number multi-copy intervals. Some genes contain several distinct multi-copy subfamilies giving rise to overlapping multi-copy sequences (e.g. *NPIP*, *GOLGA8E*). Many genes have not been assigned a function (e.g. *PDXDC2*, *URG7*).

3.2. *Stably expressed genes in stable copy number regions: major pathways and functions involved in tumour maintenance*

Within stable regions, there were 9463 complete CCDS genes (8083 with Entrez Gene IDs) and 7403 genes within unstable regions. Expression analysis of paired breast tumours and normal controls (Turashvili et al., 2007) demonstrated 5804 genes to be stable at both genomic and transcript levels. These dually stable genes were analyzed in established cellular pathways and gene ontologies, which suggested functions that are maintained in the majority of breast tumours. Interestingly, many of the same pathways containing stable genes are also disrupted by abnormalities in unstable or mutated genes. We suggest that this overlap may be relevant to tumour initiation and/or maintenance. The dually stable gene set was enriched for KEGG pathway hsa05200, which includes neuroactive ligand–receptor interaction ($p = 7.32e^{-14}$), cytokine–cytokine receptor interaction ($p = 2.54e^{-12}$), MAPK signalling pathway ($p = 1.76e^{-7}$), focal adhesion ($p = 3.55e^{-10}$), Jak-STAT signalling pathway ($p = 6.04e^{-8}$), cell cycle ($p = 2.64e^{-7}$), cell–cell/adherens junction ($p = 7.19e^{-9}$), and TGF-beta signalling ($p = 1.01e^{-7}$; Supporting Information Table S1). Other pathways with significantly enriched stable genes included oxidative phosphorylation ($p = 1.46e^{-14}$), Wnt signalling ($p = 7.30e^{-10}$), natural killer cell mediated cytotoxicity ($p = 2.58e^{-8}$), leukocyte transendothelial migration ($p = 3.53e^{-8}$), and cell adhesion ($p = 3.81e^{-5}$). Some interesting patterns emerge among the stable and unstable genes in the same pathways. In the MAPK pathway (Figure 4), we noted a higher propensity for unstable gene products encoding surface receptors and ligands (i.e. initiating sites of aberrant signalling), contrasting with many stable genes being distributed throughout the rest of these pathways (i.e. to propagate the abnormal signals). A similar pattern emerged for the Wnt signalling pathway (Supporting Information Figure S5).

Dually stable genes were classified according to biological processes to which they contribute. Cell adhesion and motility, signal transduction, transcriptional regulation, transport, cellular metabolism, and RNA metabolism are dysregulated in breast neoplasia (Sjoblom et al., 2006). Contrary to expectation, the pathways containing stable genes were not distinct from those containing dysregulated genes with abnormal copy number. The gene ontologies of 373 stable genes were significantly enriched for known cancer pathways. Cell surface receptor-linked signal transduction (42.9% of stable gene set; $p = 2.16e^{-41}$), followed by intracellular signalling cascade (31.9%; $p = 2.55e^{-32}$) and G-protein coupled receptor protein signalling (24.4%; $p = 1.35e^{-20}$) showed the most significant enrichment. Additionally, positive regulation of catalytic activity ($p = 2.59e^{-26}$), regulation of cell proliferation ($p = 1.35e^{-20}$), protein amino acid phosphorylation ($p = 5.84e^{-19}$), regulation of macromolecule metabolic process ($p = 1.90e^{-12}$), regulation of cell motion ($p = 1.58e^{-8}$), anti-apoptosis ($p = 5.08e^{-4}$), and regulation of epithelial cell proliferation ($p = 0.02$) were significantly enriched. These functions are associated with tumour cell survival, proliferation, repair, and regeneration.

3.2.1. *Replication of findings*

Results were replicated using an independent dataset of tumour-normal pairs analyzed for both copy number and expression (Chin et al., 2007; Naderi et al., 2007). Of the 7692 protein-coding genes with stable copy number, 3589 did not exhibit differences in expression. These genes were represented in 150 enriched pathways, 95 of which were shared by the Hicks et al. (2006) (n_1) and Chin et al. (2007) (n_2) datasets.

The most significant pathways common to both analyses included MAPK signalling ($n_1 = 74$, $n_2 = 67$), regulation of actin cytoskeleton ($n_1 = 64$, $n_2 = 50$), Wnt signalling ($n_1 = 52$, $n_2 = 40$), insulin signalling ($n_1 = 50$, $n_2 = 37$), VEGF signalling ($n_1 = 23$, $n_2 = 14$), apoptosis ($n_1 = 22$, $n_2 = 17$) and glycolysis/gluconeogenesis ($n_1 = 21$, $n_2 = 23$) (Supporting Information Table S6). This study revealed 325 gene ontologies enriched for stable genes (n_2; $p \leq 0.05$), with 183 gene ontologies shared with the above gene set (n_1) (Supporting information Table S7).

The composition of the dually stable gene set was supported by several other replicate genomic and expression breast cancer studies (http://cancer.cytognomix.org/stable). A high throughput, sequence-based genomic analysis (Stephens et al., 2009) identified 130 genes mutated in 2 or more tumours (~10% for $n = 24$). However, only 20 of these genes were present in the stable gene set (http://cancer.cytognomix.org/stable) (Supporting Information Table S8). The deduced dually stable gene set was also compared with the COSMIC and Cancer Gene Consensus Databases of mutations. Of the genes that are commonly mutated in breast cancer in these databases, 8 were present among the deduced set of stable copy number and expressed genes. Furthermore, differentially expressed genes in 5 independent series of paired breast tumour-normal GEO datasets (see Methods) identified 69 differentially expressed genes in 2 or more series among the 5804 stably expressed genes with normal copy number (Supporting information Table S8). After accounting for unstable or mutated genes present in the combined GEO, COSMIC, Cancer Gene Consensus, and high throughput tumour sequencing data, the number of dually stable genes deduced from Hicks et al. (2006) and Turashvili et al. (2007) was reduced by 1.6% ($n = 96$) to 5708. Intersecting the above data with Chin et al., 2007 and Naderi et al., 2007, the combined unstable or mutated gene datasets reduce the number of dually stable genes by 1.7% ($n = 66$) to 3793. The low numbers of unstable genes in the replicates overlapping the deduced stable sets suggest that this core set of stable genes will be present in the majority of breast tumours.

3.3. *Stable gene products as targets for breast cancer therapy*

A common set of dually stable genes is more likely to be functional in a plurality of tumours. Their products may represent potential therapeutic targets, since drugs acting upon them would be effective in a preponderance of tumours. These gene products are thus plausible targets for systemic breast tumour chemotherapies that are approved because of their efficacy in ablating tumours in the maximum number of patients. Genes encoding these targets would thus be maintained as an ensemble in the breast tumour genomes, since drugs would be expected to disrupt multiple functions or pathways containing them.

Of the 68 agents commonly used in breast cancer treatment, 24 target proteins encoded by dually stable genes (Table 3). Genomic stability of the 3 genes tested (*CSF1R*, *FLT4*, and *GMPR2*) was confirmed by metaphase FISH (Supporting information Table S5). Pathways containing each of these gene products included metabolic processes (e.g. purine, pyrimidine metabolism and glycolysis), inflammation (natural killer cell mediated cytotoxicity and gamma R-mediated phagocytosis) and cellular interactions (neuroactive ligand–receptor interactions, endocytosis, focal adhesion and cytokine–cytokine interaction). The context of these drug targets showed a broad distribution of biochemical pathways. In some instances, these targets occurred at the intersection of multiple pathways; for example, glycolysis and gluconeogenesis are enriched for stable gene products ($p = 0.000117$). Among these products, glucokinase is an investigational target for potential breast cancer therapy because of increased energy requirements of tumour cells (Lu and Huang, 2010; Pelicano et al., 2006; Warburg, 1956).

In some instances, multiple dually stable gene products are targets of the same drug and are members of the same pathways. For example, *FLT4*, *KDR*, *CSF1R*, and *RET* comprise four of nine known drug targets of sunitinib. *FLT4*, *KDR*, and *CSF1R* belong to signalling pathways characterized by cytokine–cytokine receptor interactions, focal adhesion, and endocytosis. These genes also contribute to vascular development, VEGF signalling and haematopoiesis, both of which are essential for tumour growth (Folkman, 2006). Inhibition of *KDR* decreases tumour growth and angiogenesis (Shao et al., 2004). *GMPR2* and *AMPD3* are also members of common nucleoside metabolic pathways which are targets of azathioprine and mercaptopurine. *GMPR2* and *AMPD3* are both components of purine and pyrimidine metabolism. F_c receptors encoded by *FCGR3B* and *FCGR2A* are targets of cetuximab and trastuzumab, and belong to the same pathways associated with natural killer cell mediated cytotoxicity and F_c gamma R-mediated phagocytosis. Anti-tumour effects of monoclonal antibodies are dependent on immune activation through these receptors (Clynes et al., 2000; Stavenhagen et al., 2007). Based on the independent replication dataset (Chin et al., 2007; Naderi et al., 2007), 5 of the 24 dually stable gene targets were also targets of venlafaxine (*HTR1A*), mercaptopurine (*GMPR2*), azathioprine (*GMPR2*), sertraline (*HTR1A*), fluvoxamine (*HTR1A*), sunitinib (*KDR*), and topotecan (*ABCG2*).

4. Discussion

If cancer is driven and maintained solely by abnormal tumour suppressor and oncogenes, then all other genes in a tumour should be mutable. We have deduced a set of genes with unaltered copy number and expression that encode products in multiple tumour-associated biochemical pathways. Stably expressed genes contained within these stable genomic regions, their assigned functions and pathways, and the susceptibility of cells to drugs that target their gene products are all consistent with selection for their preservation in the breast cancer genome. The large numbers of tumours analyzed, many of which exhibit defective DNA repair facilitating genome-wide mutation accumulation (Lahtz and Pfeifer, 2011), and our initial and replication analyses defining a consistent core set of dually stable genes, each mitigate against the possibility that these are bystander genes that have fortuitously escaped mutation. The present analysis considers whether this gene set encodes functions required for homeostatic tumour initiation and maintenance in a set of phenotypically heterogeneous tumours. Dysregulated signalling

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

Table 3 – Drugs used in breast cancer therapy that act on stable gene products.

Drug: No. of targets	Stable drug targets	Associated pathways
Azathioprine: 12	*GMPR2*	Purine metabolism
	AMPD3	Purine metabolism, metabolic pathways
Cetuximab: 12	*FCGR3B*	Natural killer cell mediated cytotoxicity, Leishmania infection, systemic lupus erythematosus
	FCGR2A	Fc gamma R-mediated phagocytosis, Leishmania infection, systemic lupus erythematosus
Cyanocobalamin: 11	*TCN2*	N/A
	MTRR	N/A
	MMACHC	N/A
Dasatinib: 10	*ABL1*	ErbB signalling pathway, cell cycle axon guidance, neurotrophin signalling pathway, pathogenic *E. coli* infection, pathways in cancer, chronic myeloid leukaemia, viral myocarditis
	YES1	Tight junction, adherens junction
Docetaxel: 2	*BCL2*	Apoptosis, neurotrophin signalling pathway, amyotrophic lateral sclerosis, pathways in cancer, colorectal/prostate/small cell lung cancer
Oestradiol: 4	*SHBG*	N/A
Oestrone: 5	*SHBG*	N/A
Fluvoxamine: 3	*HTR1A*	Neuroactive ligand–receptor interaction
Folic Acid: 5	*SLC25A32*	N/A
Gemcitabine: 3	*RRM1*	Pyrimidine metabolism, purine metabolism, glutathione metabolism, metabolic pathways
	CMPK1	
Irinotecan: 2	*TOP1MT*	*Panther pathways*: DNA replication
Losartan: 1	*AGTR1*	Calcium signalling pathway, neuroactive ligand–receptor interaction, vascular smooth muscle contraction, renin-angiotensin system
Mercaptopurine: 12	*GMPR2*	Purine metabolism
	AMPD3	Purine metabolism, metabolic pathways
Paclitaxel: 2	*BCL2*	Apoptosis, neurotrophin signalling pathway, amyotrophic lateral sclerosis, pathways in cancer, colorectal/prostate/small cell lung cancer
Sertraline: 3	*HTR1A*	Neuroactive ligand–receptor interaction
Sulfasalazine: 2	*ACAT1*	Butanoate metabolism, lysine degradation, pyruvate metabolism, benzoate degradation via CoA ligation, fatty acid metabolism, propanoate metabolism, valine, leucine and isoleucine degradation, synthesis and degradation of ketone bodies, tryptophan metabolism, terpenoid backbone biosynthesis, metabolic pathways
Sunitinib: 9	*FLT4*	Cytokine–cytokine receptor interaction, focal adhesion
	KDR	Cytokine–cytokine receptor interaction, focal adhesion, VEGF signalling pathway, endocytosis
	CSF1R	Haematopoietic cell lineage, cytokine–cytokine receptor interaction, endocytosis, pathways in cancer
	RET	*Panther pathways*: Endothelin signalling pathway, heterotrimeric G-protein signalling pathway-Gi alpha and Gs alpha mediated pathway, metabotropic glutamate receptor group I pathway
Topotecan: 3	*TOP1MT*	*Panther pathways*: DNA replication
	ABCG2	ABC transporters - General
Trastuzumab: 13	*FCGR3B*	Natural killer cell mediated cytotoxicity, Leishmania infection, systemic lupus erythematosus
	FCGR2A	Fc gamma R-mediated phagocytosis, Leishmania infection, systemic lupus erythematosus
Venlafaxine: 5	*HTR1A*	Neuroactive ligand–receptor interaction
Warfarin: 3	*VKORC1L1*	N/A

Drugs used in breast cancer treatment (from PharmGKB) along with the number of genes they target. Some of these drugs may not be used for killing tumour cells. Gene names of targets that are stable by copy number and expression data are reported along with associated KEGG biological pathways unless otherwise specified. N/A refers to not available.

still requires downstream genes, stable in copy number and expression, to propagate these signals and maintain basal cellular processes in cancer cells.

This study defines a subset of sequence families *in cis* in tumour cells that possess the original germline genomic structure. A plausible explanation for retention of these multiplex structures is that they encode functionally essential genes or contain regulatory sequences; and/or maintain genomic architectures required for tumour viability. We confirm stable copy numbers in multiple breast cancer cell lines by Q-PCR and FISH. All multi-copy regions tested were stable in copy number, with the exception of a 22-copy per haploid sequence in the HS578T and T47D cell lines (Supporting Information Figure S1). The 22-copy sequence is organized in four distinct clusters across 900 kb on chromosome 8p23.1 and contiguous loss of one or two of these clusters would result in a decrease

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

in copy number. Array comparative hybridization of HS578T and T47D breast cancer cell lines confirms these results (Kao et al., 2009). These deletions may be polymorphic or perhaps, all copies of this sequence family are not essential for somatic viability.

Stable genes which apparently do not undergo genomic rearrangement or do so at a low frequency are unlikely to be a chance observation, as the gene set is stable in both copy number analysis of a large number of tumours and in gene expression studies of breast cancer. 5804 genes occur in stable copy number regions and exhibit stable expression in most tumours. Furthermore, analysis of 5 additional GEO sets, COSMIC and Cancer Consensus gene databases and sequencing dataset (Stephens et al., 2009), decreased our stable gene set by <2%. These stable genes are more consistently represented in metabolism, transcription, RNA metabolic processes, cell communication, chromatin assembly or disassembly, plasma membrane organization and biogenesis, cellular component assembly and protein transport. Although some of the pathways (e.g. MAPK, Figure 4) containing stable genes, are also known to contain unstable mutation targets, we reconcile dysregulation of signalling proteins with the requirement for downstream stable genes to propagate abnormal signals and maintain basal cellular physiology. Mutations resulting in abnormal signalling at the inception of the pathway may dictate dysregulation and malfunction downstream, and therefore, stability of downstream genes in certain pathways may be critical for maintaining the tumour phenotype. Only a subset of genetic and epigenetic alterations are stable in the breast cancer genome (Feinberg and Tycko, 2004; Hicks et al., 2006; Perou et al., 2000; Widschwendter and Jones, 2002). Our findings support the idea that a common set of functional pathways are under selection to maintain tumour viability in all types of breast cancer. Stable genes tend to include those encoding responses to environmental stimuli. Primary ontologies describing these genes include cell recognition, intracellular protein transport, post-translational protein modification, cell cycle, ATP-binding and activity, drug response, membrane integrity, and signal transduction.

Certain anti-breast cancer therapies have been evaluated in the NCI-60 panel of cancer cell lines (Staunton et al., 2001) including four cell lines we studied (MCF7, MDA-MB-231, HS578T, and T-47D). Paclitaxel and its derivatives, oestrone, oestradiol, and topotecan are common therapies that act on multiple products encoded by dually stable genes which may, in part, explain why they are effective for inhibiting growth of these cell lines and for treating breast cancer.

Our data also support the suggestion that certain therapeutic agents approved for treating other diseases might also be good candidates for treatment of breast tumours. Antidepressants, in particular, have been shown to act on some of the gene products encoded by stable genomic regions. Venlafaxine targets five gene products, four of which are encoded by genes in stable regions (*HTR1A*, *SLC6A2*, *SLC6A4*, and *HTR1B*). Targeting 5-hydroxytryptamine 1A and 1B receptors (*HTR1A* and *HTR1B*), inhibits growth and induces apoptosis in cell lines for prostate (Siddiqui et al., 2006) and colorectal (Ataee et al., 2010) cancer. Metformin, an anti-diabetic drug, targets *PRKAB1* and is present within a stable interval, has anti-cancer activity against triple-negative breast cancer cell lines (Liu et al., 2009).

The effectiveness of these drugs may be related to preservation of gene structure and copy number of the multiple gene products that they target. Disruption of stable genes under selection would be expected to compromise key pathways needed to maintain tumours. Therapies that inhibit or inactivate multiple targets characterized by their stability in the genome and transcriptome may be an effective strategy to kill tumours. Judicious selection of stable targets in pathways that are intact in tumour cells, in which salvage (or alternative) pathways are mutated in these same cells, may provide a safer treatment approach. Redundant pathways could protect normal cells from toxic effects of these drugs, however tumour cells would remain susceptible. New therapies could focus on candidates that act on gene products encoded from within stable genomic regions (regardless of whether they were developed to treat breast cancer) and occur within key pathways required for tumour survival.

Acknowledgements

We gratefully acknowledge support from the University of Western Ontario, Schulich School of Medicine and Dentistry, Canada Foundation for Innovation, Ontario Research Fund (JHMK, PKR), Canada Research Chairs and Natural Sciences and Engineering Research Council of Canada (PKR), and Translational Breast Cancer Research Trainee Studentship (NIP)

This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET:www.sharcnet.ca) and Compute/Calcul Canada.

We thank Dr. M. Vincent (London Health Sciences Centre) for comments on the manuscript; Drs. A. Chambers (London Regional Cancer Program), J. Wiebe and C. Chakraborty (University of Western Ontario) for generously providing breast cancer cell lines; and the Cytogenetics Laboratory (London Health Sciences Centre) for providing samples for FISH probe validation.

Appendix. Supplementary material

Supplementary data related to this article can be found online at doi:10.1016/j.molonc.2012.01.001.

REFERENCES

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al., 2000. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat. Genet. 25, 25–29.

Ataee, R., Ajdary, S., Zarrindast, M., Rezayat, M., Shokrgozar, M.A., Ataee, A., 2010. Y25130 hydrochloride, a selective 5HT3 receptor antagonist has potent antimitogenic and apoptotic effect on HT29 colorectal cancer cell line. Eur. J. Cancer Prev. 19, 138–143.

Cailleau, R., Young, R., Olive, M., Reeves Jr., W.J., 1974. Breast tumor cell lines from pleural effusions. J. Natl. Cancer Inst. 53, 661–674.

Chin, S.F., Teschendorff, A.E., Marioni, J.C., Wang, Y., Barbosa-Morais, N.L., Thorne, N.P., Costa, J.L., Pinder, S.E., van de Wiel, M.A., Green, A.R., et al., 2007. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol. 8, R215.

Clynes, R.A., Towers, T.L., Presta, L.G., Ravetch, J.V., 2000. Inhibitory Fc receptors modulate in vivo cytoxicity against tumor targets. Nat. Med. 6, 443–446.

Collins, S., Groudine, M., 1982. Amplification of endogenous myc-related DNA sequences in a human myeloid leukaemia cell line. Nature 298, 679–681.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W., et al., 2002. Mutations of the BRAF gene in human cancer. Nature 417, 949–954.

Feinberg, A.P., Tycko, B., 2004. The history of cancer epigenetics. Nat. Rev. Cancer 4, 143–153.

Folkman, J., 2006. Angiogenesis. Annu. Rev. Med. 57, 1–18.

Friedberg, E.C., 2003. DNA damage and repair. Nature 421, 436–440.

Hackett, A.J., Smith, H.S., Springer, E.L., Owens, R.B., Nelson-Rees, W.A., Riggs, J.L., Gardner, M.B., 1977. Two syngeneic cell lines from human breast tissue: the aneuploid mammary epithelial (Hs578T) and the diploid myoepithelial (Hs578Bst) cell lines. J. Natl. Cancer Inst. 58, 1795–1806.

Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., et al., 1996. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 271, 350–353.

Hicks, J., Krasnitz, A., Lakshmi, B., Navin, N.E., Riggs, M., Leibu, E., Esposito, D., Alexander, J., Troge, J., Grubor, V., et al., 2006. Novel patterns of genome rearrangement and their association with survival in breast cancer. Genome Res. 16, 1465–1479.

Kamb, A., Gruis, N.A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S.V., Stockert, E., Day 3rd, R.S., Johnson, B.E., Skolnick, M.H., 1994. A cell cycle regulator potentially involved in genesis of many tumor types. Science 264, 436–440.

Kao, J., Salari, K., Bocanegra, M., Choi, Y.L., Girard, L., Gandhi, J., Kwei, K.A., Hernandez-Boussard, T., Wang, P., Gazdar, A.F., et al., 2009. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PLoS One 4, e6146.

Keydar, I., Chen, L., Karby, S., Weiss, F.R., Delarea, J., Radu, M., Chaitcik, S., Brenner, H.J., 1979. Establishment and characterization of a cell line of human breast carcinoma origin. Eur. J. Cancer 15, 659–670.

Khan, W.A., Knoll, J.H.M., Rogan, P.K., 2011. Context-based FISH localization of genomic rearrangements within chromosome 15q11.2q13 duplicons. Mol. Cytogenet. 4 (1), 15.

Knoll, J.H., Rogan, P.K., 2003. Sequence-based, in situ detection of chromosomal abnormalities at high resolution. Am. J. Med. Genet. A 121A (3), 245–257.

Knoll, J.H.M., Lichter, P., 2005. In situ hybridization to metaphase chromosomes and interphase nuclei. In: Haines, J.L., Korf, B.R., Morton, C.C., Seidman, C.E., Seidman, J.G., Smith, D.R. (Eds.), Current Protocols in Human Genetics. John Wiley & Sons, Inc., pp. 4.3.1–4.3.31.

Kolomietz, E., Meyn, M.S., Pandita, A., Squire, J.A., 2002. The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. Gene. Chromosomes. Cancer 35, 97–112.

Lahtz, C., Pfeifer, G.P., 2011. Epigenetic changes of DNA repair genes in cancer. J. Mol. Cell Biol. 3, 51–58.

Li, J., Yen, C., Liaw, D., Podsypanina, K., Bose, S., Wang, S.I., Puc, J., Miliaresis, C., Rodgers, L., McCombie, R., et al., 1997. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science 275, 1943–1947.

Lisitsyn, N., Lisitsyn, N., Wigler, M., 1993. Cloning the differences between two complex genomes. Science 259, 946–951.

Liu, B., Fan, Z., Edgerton, S.M., Deng, X.S., Alimova, I.N., Lind, S.E., Thor, A.D., 2009. Metformin induces unique biological and molecular responses in triple negative breast cancer cells. Cell Cycle. 8, 2031–2040.

Lu, W., Huang, P., 2010. Glycolytic pathway as a target for tumor inhibition. In: Bagley, R.G. (Ed.), The Tumor Microenvironment. Springer, pp. 91–118.

Lucito, R., Healy, J., Alexander, J., Reiner, A., Esposito, D., Chi, M., Rodgers, L., Brady, A., Sebat, J., Troge, J., et al., 2003. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. Genome Res. 13, 2291–2305.

Naderi, A., Teschendorff, A.E., Barbosa-Morais, N.L., Pinder, S.E., Green, A.R., Powe, D.G., Robertson, J.F., Aparicio, S., Ellis, I.O., Brenton, J.D., et al., 2007. A gene-expression signature to predict survival in breast cancer across independent data sets. Oncogene 26, 1507–1516.

Ngo, V.N., Davis, R.E., Lamy, L., Yu, X., Zhao, H., Lenz, G., Lam, L.T., Dave, S., Yang, L., Powell, J., et al., 2006. A loss-of-function RNA interference screen for molecular targets in cancer. Nature 441, 106–110.

Nowell, P.C., 2002. Tumor progression: a brief historical perspective. Semin. Cancer Biol. 12, 261–266.

Osoegawa, K., Mammoser, A.G., Wu, C., Frengen, E., Zeng, C., Catanese, J.J., de Jong, P.J., 2001. A bacterial artificial chromosome library for sequencing the complete human genome. Genome Res. 11, 483–496.

Pace 2nd, J.K., Sen, S.K., Batzer, M.A., Feschotte, C., 2009. Repair-mediated duplication by capture of proximal chromosomal DNA has shaped vertebrate genome evolution. PLoS Genet. 5, e1000469.

Pelicano, H., Martin, D.S., Xu, R.H., Huang, P., 2006. Glycolysis inhibition for anticancer treatment. Oncogene 25, 4633–4646.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al., 2000. Molecular portraits of human breast tumours. Nature 406, 747–752.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic. Acids Res. 29, e45.

Santarosa, M., Ashworth, A., 2004. Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way. Biochim. Biophys. Acta. 1654, 105–122.

Schafer, M., Schwender, H., Merk, S., Haferlach, C., Ickstadt, K., Dugas, M., 2009. Integrated analysis of copy number alterations and gene expression: a bivariate assessment of equally directed abnormalities. Bioinformatics 25, 3228–3235.

Shao, R., Bao, S., Bai, X., Blanchette, C., Anderson, R.M., Dang, T., Gishizky, M.L., Marks, J.R., Wang, X.F., 2004. Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression. Mol. Cell Biol. 24, 3992–4003.

Siddiqui, E.J., Shabbir, M., Mikhailidis, D.P., Thompson, C.S., Mumtaz, F.H., 2006. The role of serotonin (5-hydroxytryptamine1A and 1B) receptors in prostate cancer cell proliferation. J. Urol. 176, 1648–1653.

Silva, J.M., Marran, K., Parker, J.S., Silva, J., Golding, M., Schlabach, M.R., Elledge, S.J., Hannon, G.J., Chang, K., 2008. Profiling essential genes in human mammary cells by multiplex RNAi screening. Science 319, 617–620.

Sjoblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., et al., 2006. The consensus coding sequences of human breast and colorectal cancers. Science 314, 268–274.

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A., McGuire, W.L., 1987. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 235, 177–182.

Soule, H.D., Vazguez, J., Long, A., Albert, S., Brennan, M., 1973. A human cell line from a pleural effusion derived from a breast carcinoma. J. Natl. Cancer Inst. 51, 1409–1416.

Staunton, J.E., Slonim, D.K., Coller, H.A., Tamayo, P., Angelo, M.J., Park, J., Scherf, U., Lee, J.K., Reinhold, W.O., Weinstein, J.N., et al., 2001. Chemosensitivity prediction by transcriptional profiling. Proc. Natl. Acad. Sci. U.S.A. 98, 10787–10792.

Stavenhagen, J.B., Gorlatov, S., Tuaillon, N., Rankin, C.T., Li, H., Burke, S., Huang, L., Vijh, S., Johnson, S., Bonvini, E., et al., 2007. Fc optimization of therapeutic antibodies enhances their ability to kill tumor cells in vitro and controls tumor expansion in vivo via low-affinity activating Fc gamma receptors. Cancer Res. 67, 8882–8890.

Steck, P.A., Pershouse, M.A., Jasser, S.A., Yung, W.K., Lin, H., Ligon, A.H., Langford, L.A., Baumgard, M.L., Hattier, T., Davis, T., et al., 1997. Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. Nat. Genet. 15, 356–362.

Stephens, P.J., McBride, D.J., Lin, M.-L., Varela, I., Pleasance, E.D., Simpson, J.T., Stebbings, L.A., Leroy, C., Edkins, S., Mudie, L.J., et al., 2009. Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature 462, 1005–1010.

Trempe, G.L., 1976. Human breast cancer in culture. Recent Results Cancer Res., 33–41.

Turashvili, G., Bouchal, J., Baumforth, K., Wei, W., Dziechciarkova, M., Ehrmann, J., Klein, J., Fridman, E., Skarda, J., Srovnal, J., et al., 2007. Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. BMC Cancer 7, 55.

van den Ijssel, P., Tijssen, M., Chin, S.F., Eijk, P., Carvalho, B., Hopmans, E., Holstege, H., Bangarusamy, D.K., Jonkers, J., Meijer, G.A., et al., 2005. Human and mouse oligonucleotide-based array CGH. Nucleic. Acids Res. 33, e192.

Vogelstein, B., Kinzler, K.W., 2004. Cancer genes and the pathways they control. Nat. Med. 10, 789–799.

Warburg, O., 1956. On the origin of cancer cells. Science 123, 309–314.

Widschwendter, M., Jones, P.A., 2002. DNA methylation and breast carcinogenesis. Oncogene 21, 5462–5482.

Wolff, A.C., Hammond, M.E., Schwartz, J.N., Hagerty, K.L., Allred, D.C., Cote, R.J., Dowsett, M., Fitzgibbons, P.L., Hanna, W.M., Langer, A., McShane, L.M., Paik, S., Pegram, M.D., Perez, E.A., Press, M.F., Rhodes, A., Sturgeion, C., Taube, S.E., Tubbs, R., Vance, G.H., van de Vijver, M., Wheeler, T.M., Hayes, D.F., 2007. American society of clinical oncology/college of American pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J. Clin. Oncol. 25, 118–145.

Xiao, Y., Gao, X., Maragh, S., Telford, W.G., Tona, A., 2009. Cell lines as candidate reference materials for quality control of ERBB2 amplification and expression assays in breast cancer. Clin. Chem. 55, 1307–1315.

Please cite this article in press as: Nicole I. Park et al., Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy, Molecular Oncology (2012), doi:10.1016/j.molonc.2012.01.001