

# Physical Confirmation and Comparative Genomics of the Rat *Mammary carcinoma susceptibility 3* Quantitative Trait Locus

Saasha Le,\* Zachary C. Martin,\* and David J. Samuelson\*<sup>†,‡,1</sup>

\*Department of Biochemistry and Molecular Genetics, <sup>†</sup>James Graham Brown Cancer Center, and <sup>‡</sup>Center for Genetics and Molecular Medicine, University of Louisville School of Medicine, Kentucky 40292

ORCID ID: 0000-0003-4171-1362 (D.J.S.)

**ABSTRACT** Human breast and rat mammary cancer susceptibility are complex phenotypes where complete sets of risk associated loci remain to be identified for both species. We tested multiple congenic rat strains to physically confirm and positionally map rat *Mammary carcinoma susceptibility 3* (*Mcs3*)—a mammary cancer resistance allele previously predicted at *Rattus norvegicus* chromosome 1 (*RNO1*). The mammary cancer susceptible Wistar Furth (WF) strain was the recipient, and the mammary cancer resistant Copenhagen (Cop) strain was the *RNO1*-segment donor for congenics. Inbred WF females averaged 6.3 carcinogen-induced mammary carcinomas per rat. Two WF.Cop congenic strains averaged 2.8 and 3.4 mammary carcinomas per rat, which confirmed *Mcs3* as an independently acting allele. Two other WF.Cop congenic strains averaged 6.6 and 8.1 mammary carcinomas per rat, and, thus, did not contain *Mcs3*. Rat *Mcs3* was delimited to 27.8 Mb of *RNO1* from *rs8149408* to *rs105131702* (*RNO1*:143700228-171517317 of RGSC 6.0/rm6). Human genetic variants with *p* values for association to breast cancer risk below  $10^{-7}$  had not been reported for *Mcs3* orthologous loci; however, human variants located in *Mcs3*-orthologous regions with potential association to risk ( $10^{-7} < p < 10^{-3}$ ) were listed in some population-based studies. Further, rat *Mcs3* contains sequence orthologous to human *11q13/14*—a region frequently amplified in female breast cancer. We conclude that *Mcs3* is an independently acting mammary carcinoma resistance allele. Human population-based, genome-targeted association studies interrogating *Mcs3* orthologous loci may yield novel breast cancer risk associated variants and genes.

## KEYWORDS

rat QTL  
breast cancer  
susceptibility  
mammary  
carcinogenesis

Excluding skin cancer, breast cancer is the most commonly diagnosed cancer among women in the United States, accounting for 29% of cancers diagnosed, and 14% of cancer-related deaths in females (Siegel *et al.* 2016). The current population-based lifetime-risk of a woman being diagnosed with breast cancer is 12–13% (Siegel *et al.* 2016). Breast

cancer susceptibility is a complex trait controlled by genetic, epigenetic, and environmental factors. The genetic component of breast cancer susceptibility is comprised of a spectrum of high, moderate, and low penetrance risk associated mutations and variants (Pharoah *et al.* 2002). Genetic linkage analyses and genome-wide association studies (GWAS) have enabled the identification of several risk alleles, but a majority of susceptibility loci remain unidentified (Smith *et al.* 2006; Ioannidis *et al.* 2010; Mavaddat *et al.* 2010; Economopoulou *et al.* 2015). Furthermore, GWAS have yielded many potentially associated variants. These are polymorphisms with *p* values for association above a stringent genome-wide significance threshold level of  $1 \times 10^{-7}$ , which is commonly used to avoid false-positive associations (Thomas *et al.* 2005). Comparative genomics approaches that integrate relevant QTL of experimental organisms and human loci, containing potentially associated variants, may be useful to identify additional risk associated alleles. To serve this approach, mammary cancer susceptibility QTL identified in experimental organisms must be mapped to conserved

Copyright © 2017 Le *et al.*

doi: <https://doi.org/10.1534/g3.117.039388>

Manuscript received January 10, 2017; accepted for publication March 28, 2017; published Early Online April 5, 2017.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material is available online at [www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.039388/-/DC1](http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.039388/-/DC1).

<sup>1</sup>Corresponding author: Department of Biochemistry and Molecular Genetics, University of Louisville School of Medicine, HSC-A Bldg., Room 606, 319 Abraham Flexner Way, Louisville, KY 40292. E-mail: [david.samuelson@louisville.edu](mailto:david.samuelson@louisville.edu)

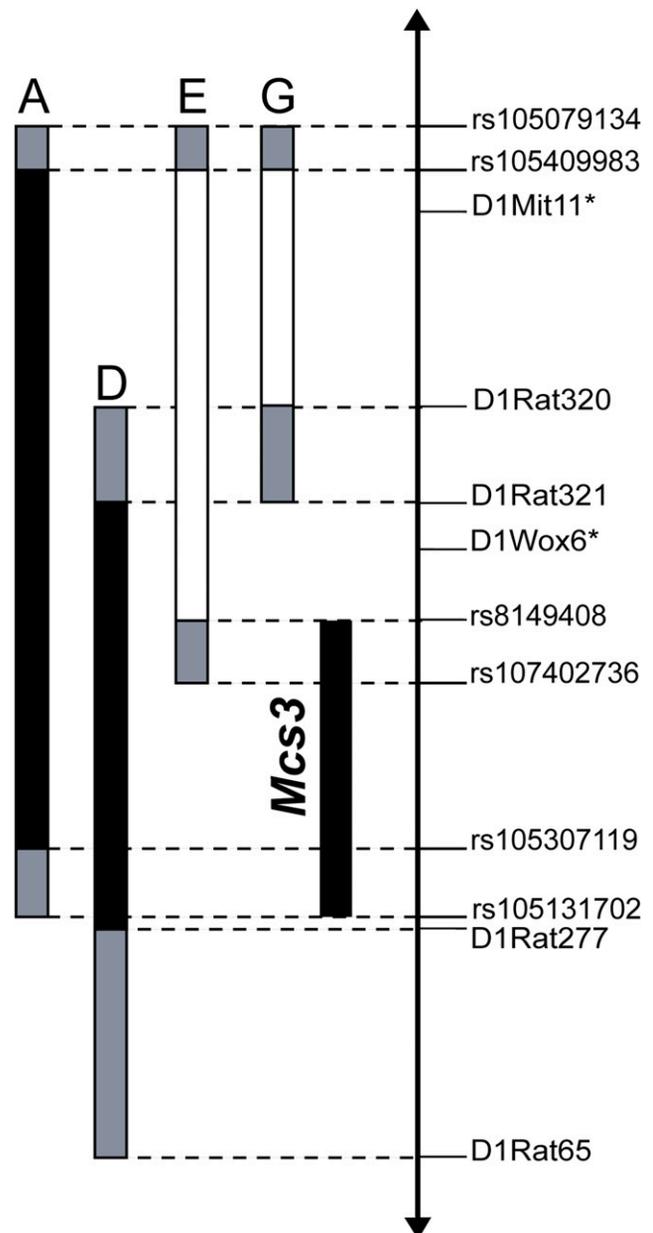
syntenic regions to effectively guide human genome-targeted association studies (Lesueur *et al.* 2005; Samuelson *et al.* 2007).

Laboratory rats (*Rattus norvegicus*) provide a good experimental model of female breast cancer, as rat mammary carcinomas are similar to human breast carcinomas with respect to histopathology and hormone responsiveness (Gould 1995). Evidence suggests that both rat mammary and human breast carcinomas originate from mammary ductal cells (Russo *et al.* 1983). Inbred rat strains that differ in susceptibility to 7,12-dimethylbenz[*a*]anthracene (DMBA)-, N-methyl-N-nitrosourea (NMU)-, or estrogen-induced mammary carcinogenesis have been used to predict locations of *Mammary carcinoma susceptibility* (*Mcs*) and *Estrogen-induced mammary cancer* (*Emca*) loci (Hsu *et al.* 1994; Shepel *et al.* 1998; Lan *et al.* 2001; Gould *et al.* 2004; Quan *et al.* 2006; Schaffer *et al.* 2006; Ren *et al.* 2013). Linkage analyses using crosses between the DMBA-induced mammary carcinoma susceptible Wistar-Furth (WF) and resistant Copenhagen (Cop) strains resulted in four predicted *Mcs* QTL named *Mcs1*, *Mcs2*, *Mcs3*, and *Mcs4* (Hsu *et al.* 1994; Shepel *et al.* 1998). Resistance-associated QTL, *Mcs1* and *Mcs2*, have been physically confirmed (Haag *et al.* 2003; Sanders *et al.* 2011). In this article, we report our congenic strain results that physically confirm rat *Mcs3*, and delimit this Cop resistance QTL to a 27.8 Mb segment of rat chromosome 1 (*RNO1*).

## MATERIALS AND METHODS

### Congenetic breeding and genotyping

Congenetic lines were made by adapting previously described methods to *RNO1* (Samuelson *et al.* 2003). Briefly, rats with selected Cop *RNO1* segments from the predicted *Mcs3* QTL were introgressed into a WF/NHsd genetic background by successive backcrossing to the WF strain. Congenic rat lines were maintained in an Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC)-approved facility on a 12-hr light/dark cycle, and provided LabDiet 5001 Rodent Diet (PMI Nutrition International) and water *ad libitum*. All animal protocols were approved by the University of Louisville Animal Care and Use Committee. Sequence information and locations of genetic markers defining the ends of COP alleles in each congenic line A, D, E, and G are available at the UCSC Genome Browser ([www.genome.ucsc.edu](http://www.genome.ucsc.edu)), the Rat Genome Database (<http://rgd.mcw.edu/>), or Supplemental Material, Table S1. Animals were genotyped as described in Samuelson *et al.* (2003). Briefly, tail clips were used for DNA extraction and genotyping by either gel electrophoresis or Sanger sequencing of PCR products. Informative genetic markers at respective congenic ends, and at 10 Mb intervals between ends, were used to determine genotypes. For microsatellite markers, each fast-PCR underwent denaturation of 95° for 10 sec, followed by 40 cycles of 94° for 0 sec and 63° for 8 sec, and an extension at 72° for 30 sec on an Applied Biosystems Veriti Fast Thermal Cycler. Amplified genomic DNA was run on a 3% high-resolution agarose gel. After electrophoresis, gels were stained in SYBR Gold and scanned using a Typhoon imager. Visible gel-bands of the appropriate size were analyzed along with DNA from homozygous Cop and WF, and (WF × Cop)<sub>F1</sub> (heterozygous) rats. In regions where informative microsatellite markers were limiting, Primer3 was used to design primers to PCR amplify genomic regions containing potentially WF/Cop informative single nucleotide variants (SNVs), identified using the SNPlotyper function available at the Rat Genome Database. Genomic DNA was extracted from spleen tissue, and PCR amplified. These reactions were cleaned using Promega Wizard SV Gel and PCR clean-up system. Samples were sequenced by the University of Louisville, Center for Genetics and Molecular Medicine, sequencing core using an ABI PRISM



**Figure 1** Map of WF.Cop congenic lines that delimit Rat *Mcs3* to 27.8 Mb of *RNO1*. WF.Cop congenic lines used to physically confirm and map *Mcs3* are represented by vertical bars that, respectively, define the Cop segment of *RNO1* introgressed into a WF/NHsd strain genetic background. Black bars represent Cop alleles in congenic lines with an *Mcs3*-associated mammary carcinoma resistant phenotype, whereas white bars represent congenic lines with a susceptible phenotype. Gray bars at the ends of congenic segments mark regions of unknown genotype. The vertical axis represents a segment of *RNO1* defined by relative locations of informative genetic markers shown as horizontal tick marks. Asterisks indicate the peak LOD score markers from the original QTL scan (Shepel *et al.* 1998).

7700 Sequence Detection System. Primer sequences that define the ends of *Mcs3* congenic lines A and E can be found in Table S2.

### Phenotyping

Female WF.Cop congenic and WF/NHsd (Envigo) rats were gavaged with a single oral dose of DMBA (65 mg DMBA/kg body mass) in

■ **Table 1 Mammary carcinoma multiplicity phenotypes of WF.Cop *RNO1* congenic and WF/NHsd females**

Strain	Genetic Markers Spanning Cop <i>RNO1</i> Maximal Segment	Phenotype Mean ± SD	N	p-Value <sup>a</sup>
A	rs105079134 to rs105131702	3.4 ± 2.2	30	0.019
D	D1Rat320 to D1Rat65	2.8 ± 2.3	19	0.015
E	rs105079134 to rs107402736	6.6 ± 3.4	25	0.935
G	rs105079134 to D1Rat321	8.1 ± 3.4	29	0.334
WF/NHsd	NA	6.3 ± 3.8	12	—

WF, Wistar Furth; Cop, Copenhagen; SD, standard deviation; NA, not applicable.

<sup>a</sup>p-values from Mann-Whitney *post hoc* tests comparing each congenic line phenotype to a susceptible WF/NHsd phenotype following a statistically significant Kruskal-Wallis test with *p*-value < 0.0001.

sesame oil at 50–55 d of age. Mammary carcinomas  $\geq 3 \times 3$  mm<sup>2</sup> were counted at 15 wk post-treatment.

### Comparative genomics

Human orthologous regions and transcripts mapping to the delimited rat *Mcs3* and human orthologous loci were identified using the UCSC Genome Browser. The *R. norvegicus* reference genome sequence version RGSC 6.0/rn6 and *Homo sapiens* version GRCh38/hg38 were used. Breast cancer associated genes were identified by searching NCBI/PubMed using the respective *Mcs3*-nominated gene name and breast cancer as search terms. Databases at NHGRI-EBI (Welter *et al.* 2014) and GWAS Central (Beck *et al.* 2014) were searched for variants located in *Mcs3* orthologous regions that had *p* values of  $1 \times 10^{-3}$  or less for association to breast cancer.

### Statistical analysis

Mammary carcinoma multiplicity data were analyzed by comparing congenic strain phenotypes to a susceptible WF/NHsd phenotype. First, mammary carcinoma multiplicity phenotypes for all lines were compared using the Kruskal-Wallis nonparametric test to protect for multiple comparisons. Following a significant Kruskal-Wallis test, which was *p* < 0.0001, select group comparisons were made by two-tailed Mann-Whitney nonparametric tests corrected for ties; *p* values  $\leq 0.05$  were considered statistically significant.

### Data availability

Strain WF.Cop D is available upon request. Table S1 contains sequence and location information for primers to amplify microsatellite markers on *RNO1*. Table S2 contains sequence and location information for primers to amplify *RNO1* SNV containing sequences. File S1 contains a list of *Mcs3*-nominated breast cancer susceptibility candidate genes.

## RESULTS

### Physical confirmation and positional mapping of *Mcs3* With WF.Cop congenics

The *Mcs3* QTL was predicted to exist on *RNO1* by the laboratory of M. N. Gould in a previous study (Shepel *et al.* 1998). Figure 1 is a map of WF.Cop congenic strains used in the present study to physically confirm and positionally map the location of *Mcs3*. Table 1 contains DMBA-induced mammary carcinoma susceptibility phenotypes (mean ± SD mammary carcinomas per rat) of the congenic lines depicted in Figure 1. Congenic lines A (*n* = 30) and D (*n* = 19) developed  $3.4 \pm 2.2$  and  $2.8 \pm 2.3$  mammary tumors per rat, respectively. Mammary cancer susceptible WF females (*n* = 12) developed  $6.3 \pm 3.8$  DMBA-induced mammary carcinomas per rat. Females from lines A and D developed significantly less mammary tumors per rat compared to susceptible WF control females (*p*-values = 0.019 and 0.015, respectively). Reduced susceptibility lines A and D contained partially overlapping segments of Cop *RNO1*, and were not different from each other with respect to DMBA-induced tumor multiplicity (*p*-value = 0.198). This suggests that these congenic lines likely contain the same QTL. Thus, the reduced susceptibility phenotypes of both lines, A and D, physically confirm the *Mcs3* locus.

Congenic lines E (*n* = 25) and G (*n* = 29) developed  $6.6 \pm 3.4$  and  $8.1 \pm 3.4$  mammary tumors per rat, respectively. Neither line was significantly different in tumor multiplicity compared to susceptible WF rats (*p*-values = 0.935 and 0.334, respectively). Thus, neither line E nor line G contains an independently acting mammary carcinoma susceptibility QTL with an effect on the WF susceptibility phenotype.

The region of overlap between Cop *RNO1* segments that indicated a presence of at least one independently acting *Mcs3* allele (lines A and D) in combination with nonoverlap between these segments and the Cop *RNO1* segment contained in line E, which did not indicate the presence of an independently acting Cop allele, was used to delimit the

■ **Table 2 Minimum-number or names of genes at human loci containing rat *Mcs3* orthologous sequence**

Human Locus	Position <sup>a</sup>	Gene Transcripts (Minimum Number or Name)		
		Human	Rat	Human/Rat Orthologs
15q25.2	15:83134545–84130720	6	5	5
15q25.1–15q25.2	15:80005820–82285404	22	13	13
11q13.4–11q13.5, 11q14.1–11q14.3	11:71915764–89617253	142	103	87
11p15.4	11:3609839–7196087	113	186	90
Xq28	X:154886349–154888061	<i>F8A1</i>	<i>F8a1</i>	1
10q23.33	10:94762624–94853260	<i>CYP2C19</i>	<i>Cyp2c6v1</i>	1
11p11.12	11:49146635–49208670	<i>FOLH1</i>	<i>Folh1</i>	1
2p16.3	2:51696250–51787033	<i>LOC730100</i>	0	0
Total		287	310	198

Rat *Mcs3* from SNVs rs8149408 to rs105131702 or *RNO1*:143700228–171517317 (RGSC 6.0/rn6).

<sup>a</sup>Human genome reference GRCh38/hg38.

■ **Table 3 Rat *Mcs3*-nominated human genes and amplified region correlated with breast cancer development**

ID	Name	Locus	RefSeq on Function	Reference
<b>Genes</b>				
<i>ILK</i>	<i>Integrin-linked kinase</i>	11p15.4	Serine/threonine kinase that regulates integrin-mediated signaling	Yang et al. (2013)
<i>IL18BP</i>	<i>Interleukin 18 binding protein</i>	11q13.4	Inhibitor of proinflammatory cytokine IL18	Yang et al. (2015)
<i>EMSY</i>	<i>EMSY, BRCA2 interacting transcriptional repressor</i>	11q13.5	Repression of BRCA2-mediated DNA repair	Hughes-Davies et al. (2003), Madjd et al. (2014)
<i>PAK1</i>	<i>p21 (RAC1) activated kinase 1</i>	11q13.5–q14.1	Serine/threonine kinase that regulates cell motility and morphology	Adam et al. (2000), Wang et al. (2005), Bostner et al. (2007)
<i>RSF1</i>	<i>Remodeling and spacing factor 1</i>	11q14.1	Histone chaperone involved in RSF chromatin-remodeling complex	Ren et al. (2014)
<i>GAB2</i>	<i>GRB2 associated binding protein 2</i>	11q14.1	Adapter protein involved in signal transduction with a known role in PI3K activation	Bocanegra et al. (2010)
<i>CYP2C19</i>	<i>Cytochrome P450 family 2 subfamily C member 19</i>	10q23.33	Monooxygenase involved in xenobiotic and estrogen metabolism	Justenhoven et al. (2009), Sangrajrang et al. (2009), Gan et al. (2011)
<b>Amplified region</b>				
11q13/14		Amplified in a subset of breast cancers that are typically ER+ and have a poor prognosis		Tsuda et al. (1989)

*Mcs3* QTL interval (Figure 1). Congenic *Mcs3*-resistance-associated lines A and D overlapped each other minimally from markers *D1Rat321* to *rs105307119*, and maximally from *D1Rat320* to *rs105131702*. Susceptible line E minimally overlapped resistance-associated line A from *rs105409983* to *rs8149408* and resistance-associated line D from *D1Rat320* to *rs8149408*. Considering the phenotypes and relevant regions of overlap and nonoverlap of lines A, D, and E, the *Mcs3* QTL was delimited to a 27.8 Mb region of *RNO1* that spans from SNVs *rs8149408* to *rs105131702*. Respectively, these SNVs physically mapped to *RNO1* base positions 143,700,228 and 171,517,317 of the rat reference genome build RGSC 6.0/rn6.

Results of the original QTL scan predicted two LOD score peaks on *RNO1* at markers *D1Mit11* and *D1Wox6* (Shepel et al. 1998). These peak markers are denoted with asterisks in Figure 1. Congenic line E had the Cop genotype at both peak markers, but did not have an *Mcs3*-associated resistance phenotype, as would be predicted based on the original QTL scan. The phenotype of congenic line G, which was derived from line E and contained the Cop allele at *D1Mit11* and the WF allele at *D1Wox6*, was measured to determine if epistatic QTL might be responsible for the susceptibility phenotype of line E. The susceptibility phenotypes of lines E and G were not significantly different when compared to each other ( $p$ -value = 0.182). Thus, we did not uncover evidence of an epistatic interaction to potentially explain the susceptibility phenotype of line E.

### Human orthologous regions to the rat *Mcs3* QTL

The delimited rat *Mcs3* QTL from *rs8149408* to *rs105131702* of *RNO1* was found to align to four syntenic regions of human chromosomes 11 and 15 and four other regions of the human genome that contained a single gene (Table 2). The UCSC genome browser (<https://genome.ucsc.edu>) was used to identify known genes at rat *Mcs3* and orthologous human loci (Kent et al. 2002; Karolchik et al. 2004). These genes are listed in File S1. There were 310 rat and 287 human genes annotated in the respective genomic intervals. At least 198 of these genes were in common or orthologous between these species.

Public databases, namely the NHGRI-EBI catalog (<https://www.ebi.ac.uk/gwas>) (Welter et al. 2014), GWAS Central ([www.gwascentral.org](http://www.gwascentral.org))

(Beck et al. 2014), NCBI PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>), and the Cancer Portal of Rat Genome Database (RGD) (<https://rgd.mcw.edu>), were searched to identify genes and genetic variation at *Mcs3* orthologous human genome regions that have been associated with breast carcinoma susceptibility and development. Table 3 contains a list of female breast cancer associated genes and amplified regions that have *Mcs3*-nominated rat orthologs. Rat *Mcs3* was found to contain sequence orthologous to human *11q13/14*, which contains multiple genes and is amplified in a subset of female breast carcinomas with poor prognosis (Tsuda et al. 1989). In addition to a functional correlation with breast cancer development, studies reporting genetic associations of *CYP2C19* variants with breast cancer susceptibility have been published (Justenhoven et al. 2009; Sangrajrang et al. 2009; Gan et al. 2011).

Genome-wide significant associations ( $p$  values  $< 10^{-7}$ ) between human genetic variants located in *Mcs3*-orthologous multigene regions and female breast cancer risk have not been reported, but potentially associated variants with  $p$  values  $< 10^{-3}$  have been reported (Table 4).

### DISCUSSION

Rat *Mcs3* was physically confirmed and delimited to a 27.8 Mb segment of rat chromosome 1. Rat *Mcs3* is the last of the known Cop rat mammary carcinoma resistance-associated QTL (*Mcs1-3*) to be physically confirmed. In their linkage analysis predicting the *Mcs3* QTL, Gould's group reported that *Mcs3* heterozygous females had, on average, a 42% reduction in mammary carcinoma multiplicity compared to the WF phenotype, and females homozygous for the *Mcs3* Cop allele had an 84% reduction in mammary tumor number (Shepel et al. 1998). Thus, they appropriately concluded that there was no dominance effect at the *Mcs3* QTL. While we did not test heterozygous females in our study, we observed that the *Mcs3* Cop allele reduced the mammary carcinoma susceptibility phenotype of the highly susceptible WF strain from 46 to 65% when homozygous. The discrepancy between homozygous genotypes in these two studies may be due to effects of Cop resistance alleles at other *Mcs* QTL present in the linkage analysis study, as the  $F_2$  females in that study were not required to be WF homozygous at *Mcs1* or *Mcs2* resistance-associated QTL. Another possibility is that

■ **Table 4 Rat *Mcs3*-nominated human variants potentially associated with breast cancer susceptibility**

Variant	Locus	Base Position (hg38)	Genomic Context/Position	p-Value for Association	Reference
rs1435808	2p16.3	chr2:52,036,315	Intronic/ <i>LOC730100</i>	0.000642	Hunter et al. (2007)
rs6733295		chr2:52,040,189	Intronic/ <i>LOC730100</i>	0.000959	Hunter et al. (2007)
rs4352262		chr2:53,374,593	Intergenic/between <i>LOC730100</i> and <i>ASB3</i>	0.000893	Hunter et al. (2007)
rs10168550		chr2:53,385,857	Intergenic/between <i>LOC730100</i> and <i>ASB3</i>	0.000381	Hunter et al. (2007)
rs7126870	11p15.4	chr11:3868829	Intronic/ <i>STIM1</i>	0.000898	Hunter et al. (2007)
rs1459952	11q14.1	chr11:81871697	Intergenic/between <i>LOC101928944</i> and <i>MIR4300HG</i>	0.000562	Kibriya et al. (2009)
rs3793949		chr11:83847695	Intronic/ <i>DLG2</i>	0.000702	Hunter et al. (2007)
rs11235127	11q14.2	chr11:87377502	Intergenic/between <i>THEM135</i> and <i>LOC105369423</i>	0.00005	Easton et al. (2007)
rs12269979		chr11:87960152	Intergenic/between <i>LOC105369423</i> and <i>RAB38</i>	0.000439	Hunter et al. (2007)
rs6495623	15q25.2	chr15:81848308	Intergenic/between <i>TMC3</i> and <i>MEX3B</i>	0.000871	Hunter et al. (2007)

*Mcs3* contains multiple independent QTL, and at least one of these was not contained in the congenic strains tested for our study.

The 27.8 Mb region of *RNO1* from SNVs *rs8149408* to *rs105131702* is the segment most likely to contain *Mcs3*; however, it is possible that *Mcs3* is more complex. For example, an interaction between one or more elements in the delimited region and elements outside of this region could be required to confer the *Mcs3*-associated phenotype. The delimited *Mcs3* segment is distal to both *D1Mit11* and *D1Wox6*, the markers with peak LOD scores from the original QTL scan, but within the predicted *Mcs3* QTL interval (Shepel et al. 1998). This is not the first time a rat *Mcs* locus has been mapped to a genomic segment that did not contain a peak marker of the original QTL scan. For example, neither *Mcs1b* nor *Mcs1c* QTL contain the peak LOD marker of the predicted *Mcs1* QTL (Haag et al. 2003; denDekker et al. 2012). The *Mcs3* QTL, as delimited here, overlaps the distal third of mouse *Mammary tumor susceptibility modifier 1 (Mtsm1)* (Koch et al. 2007). Interestingly, the mouse *Mtsm1* segment overlapping rat *Mcs3* does not contain the peak LOD score marker defining the *Mtsm1* QTL. Furthermore, it is worth noting that *Mcs3* had one of the lowest LOD scores of *Mcs* QTL identified in linkage analyses using WF females as the susceptible strain (Shepel et al. 1998; Lan et al. 2001). Thus, interval mapping effectively predicted rat *Mcs* QTL, and, as expected, a wide genomic region of interrogation was required in congenic studies to pinpoint the location of each QTL for fine mapping.

Most importantly, this work provides information that is translatable to the genetic component of human breast cancer by using a versatile experimental organism that is highly relevant to female breast cancer. Different segments of the rat *Mcs3* locus align to four multigene human syntenic regions on chromosomes 11 and 15, and four single-gene loci at *Xq28*, *2p16.3*, *10p23.33*, and *11p11.12*. The rat *Mcs3* locus contains 310 annotated genes, and the orthologous human regions contain 287 annotated genes. To reduce the number of *Mcs3*-nominated candidate susceptibility genes, it will be necessary to fine-map the *Mcs3* locus. However, the human orthologous regions to rat *Mcs3* could be targeted for deep genetic analysis to determine if these regions contain risk-associated variants. In support of this direction, population-based genetic association studies have identified potentially associated variants at *Mcs3* orthologous loci *2p16.3*, *11p15.4*, *11q14.1*, *11q14.2*, and *15q25.2* (Easton et al. 2007; Hunter et al. 2007; Kibriya et al. 2009). This provides strong rationale for additional human studies to identify risk associated variants at these loci. Human genome targeted association studies require considerable resources, as a high density of variants and a large sample size are required to properly test each targeted locus.

Guidance from both rat genomics and previous human genetic studies stands to increase the likelihood of finding positive associations. Thus, these human loci are excellent candidates for further genetic analyses to test a high density of variants for association to risk.

Discussion of the entire list of genes located within the *Mcs3* QTL, as currently defined, is not practical; however, it is notable that some *Mcs3*-nominated candidate susceptibility genes have known or suggested roles in breast cancer progression, diagnosis, or prognosis. This list includes *p21 (RAC1) activated kinase 1 (PAK1)*, *EMSY*, *BRCA2 interacting transcriptional repressor (EMSY)*, *remodeling and spacing factor 1 (RSF1)*, *GRB2 associated binding protein 2 (GAB2)*, *integrin-linked protein kinase (ILK)*, *interleukin-18-binding protein (IL18BP)*, and *cytochrome P450 family 2 subfamily C member 19 (CYP2C19)*.

The *EMSY*, *PAK1*, *RSF1*, and *GAB2* genes are located within the *11q13/14* breast cancer amplicon. Interestingly, the breast cancer oncogene *cyclin D1 (CCND1)*, which is also located within the *11q13/14* amplicon, is not contained within the rat *Mcs3* delimited region. The *EMSY* gene encodes a BRCA2 binding partner that silences a potential transcription activation domain of BRCA2, thereby repressing BRCA2-mediated DNA repair (Hughes-Davies et al. 2003; Moelans et al. 2010). Amplification of *EMSY* has been reported in 13% of breast cancers, and is associated with poor survival, especially in node negative breast cancer (Hughes-Davies et al. 2003). A recent study reported a significant positive association between *EMSY* expression and lymph node metastasis, as well as a larger tumor size (Madjd et al. 2014). The *PAK1* gene product has many roles in cancer, including breast cancer progression, development and maintenance of a metastatic phenotype of breast cancer cells, and a predictor of recurrence and tamoxifen resistance in postmenopausal breast cancer (Adam et al. 2000; Bostner et al. 2007; Kumar and Li 2016). Ectopic expression of activated Pak1 has been shown to induce mouse mammary tumors (Wang et al. 2005). High Rsf-1 expression is associated with breast cancer subtype and poor prognosis (Ren et al. 2014). Amplification of *GAB2*, independent of *CCND1*, has been observed in breast carcinoma samples (Daly et al. 2002; Bocanegra et al. 2010).

A direct interaction between *PAK1* and *ILK* proteins has been identified. The activity of *ILK* is regulated by *PAK1* phosphorylation (Acconcia et al. 2007). Higher transcript levels of *ILK* have been found in human breast cancer tissue compared to adjacent nondiseased breast tissue (Yang et al. 2013). In the same study, breast cancer patients with more intense immunostaining for *ILK* had lower 5-yr survival than patients with low *ILK* levels (Yang et al. 2013).

Interleukin-18-binding protein (*IL18BP*) regulates the activity of IL-18, which has been shown to be higher in breast carcinoma tissue compared to tissue from patients with benign breast disease (Srabovic *et al.* 2011). Another study suggests that IL-18 enhances breast cancer cell migration (Yang *et al.* 2015). Due to a potential role of IL-18 in cancer progression and metastasis, an IL18BP-Fc has been developed to antagonize the effects of IL-18 (Cao *et al.* 2008).

Because of an established role in estrogen metabolism, *CYP2C19* variants have been tested for association with breast cancer risk. There is evidence that the *CYP2C19\*3* variant may be associated with breast cancer risk in Chinese Han women (Gan *et al.* 2011). Neither the *CYP2C19\*17* variant, which results in a rapid metabolizer phenotype, nor the *CYP2C19\*2* variant were associated with breast cancer risk in population-based studies of women of European descent; however, *CYP2C19\*17* was associated with decreased risk in women using hormone replacement therapy for >10 yr (Justenhoven *et al.* 2009, 2012; MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk 2010). In a study of cancer recurrence in ER+ postmenopausal breast cancer cases, it was concluded that patients with a *CYP2C19\*2* allele may benefit more from tamoxifen therapy (Beelen *et al.* 2013).

In summary, rat *Mcs3* is an independently acting QTL located in a 27.8 Mb region on *RNO1* between *rs8149408* and *rs105131702*. Additional congenic studies and other genomic approaches will be necessary to reduce the *Mcs3*-nominated candidate susceptibility gene list and determine mechanisms of *Mcs3*-associated reduced susceptibility to mammary carcinomas. While an overwhelming list of *Mcs3*-nominated breast cancer susceptibility candidate genes was identified, a manageable number of orthologous human syntenic regions were found to warrant deeper analysis of these loci in human population-based genetic association studies. The fact that some of these regions contain variants potentially associated with breast cancer risk further supports the need to ultrafine-map these loci to determine if true positive associations to susceptibility exist.

## ACKNOWLEDGMENTS

The authors thank Jennifer Sanders, Huy T. Le, and Aaron Puckett for technical assistance. This work was supported by the National Institutes of Health grant R01CA137052, the University of Louisville, School of Medicine, Department of Biochemistry and Molecular Genetics, and the University of Louisville, School of Medicine Research Office.

## LITERATURE CITED

Acconcia, F., C. J. Barnes, R. R. Singh, A. H. Talukder, and R. Kumar, 2007 Phosphorylation-dependent regulation of nuclear localization and functions of integrin-linked kinase. *Proc. Natl. Acad. Sci. USA* 104: 6782–6787.

Adam, L., R. Vadlamudi, M. Mandal, J. Chernoff, and R. Kumar, 2000 Regulation of microfilament reorganization and invasiveness of breast cancer cells by kinase dead p21-activated kinase-1. *J. Biol. Chem.* 275: 12041–12050.

Beck, T., R. K. Hastings, S. Gollapudi, R. C. Free, and A. J. Brookes, 2014 GWAS Central: a comprehensive resource for the comparison and interrogation of genome-wide association studies. *Eur. J. Hum. Genet.* 22: 949–952.

Beelen, K., M. Opdam, T. M. Severson, R. H. Koornstra, A. D. Vincent *et al.*, 2013 *CYP2C19 2* predicts substantial tamoxifen benefit in postmenopausal breast cancer patients randomized between adjuvant tamoxifen and no systemic treatment. *Breast Cancer Res. Treat.* 139: 649–655.

Bocanegra, M., A. Bergamaschi, Y. H. Kim, M. A. Miller, A. B. Rajput *et al.*, 2010 Focal amplification and oncogene dependency of *GAB2* in breast cancer. *Oncogene* 29: 774–779.

Bostner, J., M. A. Waltersson, T. Fornander, L. Skoog, B. Nordenskjöld *et al.*, 2007 Amplification of *CCND1* and *PAK1* as predictors of recurrence and tamoxifen resistance in postmenopausal breast cancer. *Oncogene* 26: 6997–7005.

Cao, Q., W. Cai, G. Niu, L. He, and X. Chen, 2008 Multimodality imaging of IL-18-binding protein-Fc therapy of experimental lung metastasis. *Clin. Cancer Res.* 14: 6137–6145.

Daly, R. J., H. Gu, J. Parmar, S. Malaney, R. J. Lyons *et al.*, 2002 The docking protein *Gab2* is overexpressed and estrogen regulated in human breast cancer. *Oncogene* 21: 5175–5181.

denDekker, A. D., X. Xu, M. D. Vaughn, A. H. Puckett, L. L. Gardner *et al.*, 2012 Rat *Mcs1b* is concordant to the genome-wide association-identified breast cancer risk locus at human 5q11.2 and *MIER3* is a candidate cancer susceptibility gene. *Cancer Res.* 72: 6002–6012.

Easton, D. F., K. A. Pooley, A. M. Dunning, P. D. Pharoah, D. Thompson *et al.*, 2007 Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447: 1087–1093.

Economopoulou, P., G. Dimitriadis, and A. Psyrris, 2015 Beyond BRCA: new hereditary breast cancer susceptibility genes. *Cancer Treat. Rev.* 41: 1–8.

Gan, C. Q., X. Y. Wang, Y. D. Cao, W. X. Ye, H. Liu *et al.*, 2011 Association of *CYP2C19\*3* gene polymorphism with breast cancer in Chinese women. *Genet. Mol. Res.* 10: 3514–3519.

Gould, K. A., M. Tochacek, B. S. Schaffer, T. M. Reindl, C. R. Murrin *et al.*, 2004 Genetic determination of susceptibility to estrogen-induced mammary cancer in the ACI rat: mapping of *Emca1* and *Emca2* to chromosomes 5 and 18. *Genetics* 168: 2113–2125.

Gould, M. N., 1995 Rodent models for the study of etiology, prevention and treatment of breast cancer. *Semin. Cancer Biol.* 6: 147–152.

Haag, J. D., L. A. Shepel, B. A. Kolman, D. M. Monson, M. E. Benton *et al.*, 2003 Congenic rats reveal three independent Copenhagen alleles within the *Mcs1* quantitative trait locus that confer resistance to mammary cancer. *Cancer Res.* 63: 5808–5812.

Hsu, L. C., W. S. Kennan, L. A. Shepel, H. J. Jacob, C. Szpirer *et al.*, 1994 Genetic identification of *Mcs-1*, a rat mammary carcinoma suppressor gene. *Cancer Res.* 54: 2765–2770.

Hughes-Davies, L., D. Huntsman, M. Ruas, F. Fuks, J. Bye *et al.*, 2003 EMSY links the *BRCA2* pathway to sporadic breast and ovarian cancer. *Cell* 115: 523–535.

Hunter, D. J., P. Kraft, K. B. Jacobs, D. G. Cox, M. Yeager *et al.*, 2007 A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.* 39: 870–874.

Ioannidis, J. P., P. Castaldi, and E. Evangelou, 2010 A compendium of genome-wide associations for cancer: critical synopsis and reappraisal. *J. Natl. Cancer Inst.* 102: 846–858.

Justenhoven, C., U. Hamann, C. B. Pierl, C. Baisch, V. Harth *et al.*, 2009 *CYP2C19\*17* is associated with decreased breast cancer risk. *Breast Cancer Res. Treat.* 115: 391–396.

Justenhoven, C., O. Obazee, S. Winter, F. J. Couch, J. E. Olson *et al.*, 2012 The postmenopausal hormone replacement therapy-related breast cancer risk is decreased in women carrying the *CYP2C19\*17* variant. *Breast Cancer Res. Treat.* 131: 347–350.

Karolchik, D., A. S. Hinrichs, T. S. Furey, K. M. Roskin, C. W. Sugnet *et al.*, 2004 The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* 32: D493–D496.

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle *et al.*, 2002 The human genome browser at UCSC. *Genome Res.* 12: 996–1006.

Kibriya, M. G., F. Jasmine, M. Argos, I. L. Andrusis, E. M. John *et al.*, 2009 A pilot genome-wide association study of early-onset breast cancer. *Breast Cancer Res. Treat.* 114: 463–477.

Koch, J. G., X. Gu, Y. Han, A. K. El-Naggar, M. V. Olson *et al.*, 2007 Mammary tumor modifiers in *BALB/cj* mice heterozygous for *p53*. *Mamm. Genome* 18: 300–309.

Kumar, R., and D. Q. Li, 2016 Chapter IV - PAKs in human cancer progression: from inception to cancer therapeutic to future oncobiology,

- pp. 137–209 in *Advances in Cancer Research*, edited by Kenneth, D. T., and B. F. Paul. Academic Press, Cambridge, MA.
- Lan, H., C. M. Kendzioriski, J. D. Haag, L. A. Shepel, M. A. Newton *et al.*, 2001 Genetic loci controlling breast cancer susceptibility in the Wistar-Kyoto rat. *Genetics* 157: 331–339.
- Lesueur, F., P. D. Pharoah, S. Laing, S. Ahmed, C. Jordan *et al.*, 2005 Allelic association of the human homologue of the mouse modifier *Ptprj* with breast cancer. *Hum. Mol. Genet.* 14: 2349–2356.
- Madjd, Z., M. E. Akbari, A. H. Zarnani, M. Khayamzadeh, E. Kalantari *et al.*, 2014 Expression of EMSY, a novel BRCA2-link protein, is associated with lymph node metastasis and increased tumor size in breast carcinomas. *Asian Pac. J. Cancer Prev.* 15: 1783–1789.
- MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk, 2010 Genetic polymorphisms in phase I and phase II enzymes and breast cancer risk associated with menopausal hormone therapy in postmenopausal women. *Breast Cancer Res. Treat.* 119: 463–474.
- Mavaddat, N., A. C. Antoniou, D. F. Easton, and M. Garcia-Closas, 2010 Genetic susceptibility to breast cancer. *Mol. Oncol.* 4: 174–191.
- Moelans, C. B., R. A. de Weger, H. N. Monsuur, R. Vijzelaar, and P. J. van Diest, 2010 Molecular profiling of invasive breast cancer by multiplex ligation-dependent probe amplification-based copy number analysis of tumor suppressor and oncogenes. *Mod. Pathol.* 23: 1029–1039.
- Pharoah, P. D., A. Antoniou, M. Bobrow, R. L. Zimmern, D. F. Easton *et al.*, 2002 Polygenic susceptibility to breast cancer and implications for prevention. *Nat. Genet.* 31: 33–36.
- Quan, X., J. F. Laes, D. Stieber, M. Riviere, J. Russo *et al.*, 2006 Genetic identification of distinct loci controlling mammary tumor multiplicity, latency, and aggressiveness in the rat. *Mamm. Genome* 17: 310–321.
- Ren, J., Q. C. Chen, F. Jin, H. Z. Wu, M. He *et al.*, 2014 Overexpression of Rsf-1 correlates with pathological type, p53 status and survival in primary breast cancer. *Int. J. Clin. Exp. Pathol.* 7: 5595–5608.
- Ren, X., J. C. Graham, L. Jing, A. M. Mikheev, Y. Gao *et al.*, 2013 Mapping of *Mcs30*, a new mammary carcinoma susceptibility quantitative trait locus (QTL30) on rat chromosome 12: identification of *Fry* as a candidate *Mcs* gene. *PLoS One* 8: e70930.
- Russo, J., L. Tait, and I. H. Russo, 1983 Susceptibility of the mammary gland to carcinogenesis. III. The cell of origin of rat mammary carcinoma. *Am. J. Pathol.* 113: 50–66.
- Samuelson, D. J., J. D. Haag, H. Lan, D. M. Monson, M. A. Shultz *et al.*, 2003 Physical evidence of *Mcs5*, a QTL controlling mammary carcinoma susceptibility, in congenic rats. *Carcinogenesis* 24: 1455–1460.
- Samuelson, D. J., S. E. Hesselson, B. A. Aperavich, Y. Zan, J. D. Haag *et al.*, 2007 Rat *Mcs5a* is a compound quantitative trait locus with orthologous human loci that associate with breast cancer risk. *Proc. Natl. Acad. Sci. USA* 104: 6299–6304.
- Sanders, J., J. D. Haag, and D. J. Samuelson, 2011 Physical confirmation and mapping of overlapping rat mammary carcinoma susceptibility QTLs, *Mcs2* and *Mcs6*. *PLoS One* 6: e19891.
- Sangrajrang, S., Y. Sato, H. Sakamoto, S. Ohnami, N. M. Laird *et al.*, 2009 Genetic polymorphisms of estrogen metabolizing enzyme and breast cancer risk in Thai women. *Int. J. Cancer* 125: 837–843.
- Schaffer, B. S., C. M. Lachel, K. L. Pennington, C. R. Murrin, T. E. Strecker *et al.*, 2006 Genetic bases of estrogen-induced tumorigenesis in the rat: mapping of loci controlling susceptibility to mammary cancer in a Brown Norway × ACI intercross. *Cancer Res.* 66: 7793–7800.
- Shepel, L. A., H. Lan, J. D. Haag, G. M. Brasic, M. E. Gheen *et al.*, 1998 Genetic identification of multiple loci that control breast cancer susceptibility in the rat. *Genetics* 149: 289–299.
- Siegel, R. L., K. D. Miller, and A. Jemal, 2016 Cancer statistics, 2016. *CA Cancer J. Clin.* 66: 7–30.
- Smith, P., L. McGuffog, D. F. Easton, G. J. Mann, G. M. Pupo *et al.*, 2006 A genome wide linkage search for breast cancer susceptibility genes. *Genes Chromosomes Cancer* 45: 646–655.
- Srabovic, N., Z. Mujagic, J. Mujanovic-Mustedanagic, Z. Muminovic, and E. Cickusic, 2011 Interleukin 18 expression in the primary breast cancer tumour tissue. *Med Glas (Zenica)* 8: 109–115.
- Thomas, D. C., R. W. Haile, and D. Duggan, 2005 Recent developments in genomewide association scans: a workshop summary and review. *Am. J. Hum. Genet.* 77: 337–345.
- Tsuda, H., S. Hirohashi, Y. Shimamoto, T. Hirota, S. Tsugane *et al.*, 1989 Correlation between long-term survival in breast cancer patients and amplification of two putative oncogene-coamplification units: *hst-1/int-2* and *c-erbB-2/ear-1*. *Cancer Res.* 49: 3104–3108.
- Wang, R. A., H. Zhang, S. Balasenthil, D. Medina, and R. Kumar, 2005 PAK1 hyperactivation is sufficient for mammary gland tumor formation. *Oncogene* 25: 2931–2936.
- Welter, D., J. MacArthur, J. Morales, T. Burdett, P. Hall *et al.*, 2014 The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42: D1001–D1006.
- Yang, H. J., Y. B. Zheng, T. Ji, X. F. Ding, C. Zhu *et al.*, 2013 Overexpression of ILK1 in breast cancer associates with poor prognosis. *Tumour Biol.* 34: 3933–3938.
- Yang, Y., S. Cheon, M. K. Jung, S. B. Song, D. Kim *et al.*, 2015 Interleukin-18 enhances breast cancer cell migration via down-regulation of claudin-12 and induction of the p38 MAPK pathway. *Biochem. Biophys. Res. Commun.* 459: 379–386.

Communicating editor: D. W. Threadgill