Poster #4616

FUNCTIONAL GENOMIC ANALYSES OF 21Q22.3 LOCUS IDENTIFY POTENTIAL FUNCTIONAL VARIANTS AND CANDIDATE GENE YBEY FOR BREAST CANCER RISK

CHRIS SHIDAL[1], XIANG SHU[1], JIE WU[1], JIFENG WANG[1], SHUYA HUANG[1], JOSHUA A. BAUER[2], XINGYI GUO[1], WEI ZHENG[1], XIAO-OU SHU[1], QIUYIN CAI[1]

[1] DIVISION OF EPIDEMIOLOGY, DEPARTMENT OF MEDICINE, VANDERBILT EPIDEMIOLOGY CENTER, VANDERBILT-INGRAM CANCER CENTER, VANDERBILT UNIVERSITY SCHOOL OF MEDICINE, NASHVILLE, TENNESSEE, 37203, USA

[2]DEPARTMENT OF BIOCHEMISTRY, VANDERBILT UNIVERSITY SCHOOL OF MEDICINE, NASHVILLE, TENNESSEE

VANDERBILT UNIVERSITY MEDICAL CENTER

ABSTRACT

We have previously identified SNP rs3541811, located in the 21q22.3 locus, to be associated with breast cancer risk in Asians. However, the underlying causal functional variants and gene(s) responsible for this association are unknown. In this study, we performed functional genomic analyses to identify potential functional variants and target genes that may mediate this association. Functional annotation for SNPs highly correlated with rs35418111, using epigenetic data from the Encyclopedia of DNA Elements (ENCODE) and Roadmap projects, showed evidence of promoter and/or enhancer activities of five putative functional SNPs, including rs35418111, rs2078203, rs8134832, rs57385578, and rs8126917. These SNPs were assessed for interactions with nuclear proteins by EMSA assay. Our results showed that compared to the reference allele, alternate allele rs2078203 (A → G) significantly increased DNA-protein interactions, while an opposite trend was observed for rs35418111 (G → A). Cis-expression quantitative loci (eQTL) analysis, using data from the Genotype-Tissue Expression (GTEx) project, The Cancer Genome Atlas (TCGA), Molecular Taxonomy of Breast Cancer International Consortium (METABRIC), and the Shanghai Breast Cancer Study, indicated that the risk allele rs35418111 is associated with a decreased expression of the YBEY (C21or57) gene, indicating its putative oncogenic function that may contribute to breast cancer risk. The gene YBEY is a relatively uncharacterized endoribonuclease in humans, which is thought to function in rRNA cleavage and maturation, similar to its bacterial homologues. We further investigated whether YBEY may have any biological effect in human breast cancers by knocking-down YBEY gene expression using siRNA in MCF-7, T47D, and MDA-MB231 cell lines. Transient knockdown of YBEY gene expression in three breast cancer cell lines consistently affected cell proliferation, colony formation, and migration/invasion, regardless of hormone receptor status. Our study provides support for a significant role for human YBEY gene in breast cancer pathogenesis and the association between the rs35418111/21q22.3 locus and breast cancer risk, which may be mediated through its correlated potential functional SNPs that regulate expression of the YBEY gene.

BACKGROUND

YBEY Function

-**Metalloprotein with endoribonuclease functions** (i.e. RNase) originally characterized in bacteria

-C21orf57 is the human homologue to bacterial Ybey gene

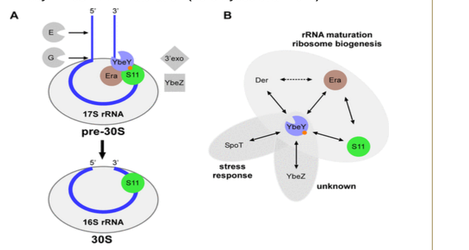
-Multiple splice variants

-In *E. coli*, Ybey functions in **16S ribosomal RNA (rRNA) maturation**

-Uncharacterized protein with **no known function in humans**

-Identified in 2018 (Zhang et al., Cancer Research) to be associated with breast cancer risk in Han Chinese women

Ybey in 16S rRNA maturation (Vercruysse et al. 2016)



Sequence alignment between E. Coli and human YBEY genes (Ghosal et al. 2017)

METHODS

Functional genomic analysis:

cis-eQTL analysis was performed using a linear regression model to identify potential causal genes associated with breast cancer risk

*(genotype and RNA-seq/microarray data available from SBCS, GTEx, TCGA, and METABRIC)

***In vitro* assays**: MB-MDA-231 [triple-negative], T47D [ER+], MCF-7 [ER+]

siRNA-mediated knock-down (k.d.) of gene of interest (GOI) – validated by qRT-PCR

Proliferation (alamarBlue) and colony formation (anchorage-dependent)

Migration/Invasion (Boyden chamber with/without growth factor reduced matrigel)

Electrophoretic mobility shift assay (EMSA)

RESULTS

Table 1. Identification of YBEY as a candidate gene for breast cancer risk

Candidate genes identified by expression quantative loci analysis (eQTL)

								SBCS (N=151)			GTEx (N=85)			TCGA (N=672)			Metabric (N=1904)		
SNP	Gene	Gene location	Chr	BP	Test	Other	Beta	SE	P	Beta	SE	P	Beta	SE	P	Beta	SE	P
rs35418111	**YBEY**	**Chr21:47706250-47717665**	**21**	**47856670**	**G**	**A**	**-0.33**	**0.12**	**9.6×10^{-3}**	**-0.84**	**0.21**	**1.0×10^{-4}**	**-0.3**	**0.10**	**5.2×10^{-3}**	**-0.2**	**0.02**	**3.7×10^{-20}**
rs8027365	SNUPN	Chr15:75890423-75918810	15	75808740	A	C	-0.17	0.08	0.02	-0.27	0.11	0.015	-0.2	0.06	9.6×10^{-5}	-0.1	0.008	4.6×10^{-10}
rs8027365	MAN2C1	Chr15:75647547-75660971	15	75808740	A	C	-0.21	0.03	2.0×10^{-11}	-0.68	0.10	3.2×10^{-9}	-0.1	0.07	0.04	-0.1	0.008	4.2×10^{-58}
rs11281251	LINC00886	Chr3:156465134-156534851	3	156519412	T	TTGTGAC	0.47	0.10	1.0×10^{-5}	0.77	0.09	3.9×10^{-13}	NA	NA	NA	NA	NA	NA
rs34331122	TBX1	Chr22:19744225-19771116	22	19762428	C	CTT	-0.26	0.09	5.5×10^{-3}	0.16	0.12	0.175	-0.2	0.04	2.0×10^{-4}	-0.1	0.01	2.8×10^{-17}
rs144145984	LOXL2	Chr8:23154701-23282841	8	23644003	CT	C	0.16	0.05	4.3×10^{-3}	NA	NA	NA	0.01	0.06	0.817	0.01	0.004	0.013
rs144145984	STC1	Chr8:23699427-23712320	8	23644003	CT	C	0.14	0.09	0.11	NA	NA	NA	0.17	0.06	3.4×10^{-3}	0.36	0.03	4.1×10^{-25}
rs2758598	SEMA4A	Chr1:156117156-156147543	1	156194339	G	A	-0.03	0.10	0.73	-0.13	0.05	0.021	-0.1	0.05	0.254	-0.03	0.01	0.0497
rs3790585	MUTYH	Chr1:45794834-45806142	1	46023356	A	T	0.02	0.07	0.81	0.15	0.13	0.251	0.21	0.08	8.4×10^{-3}	0.08	0.01	6.5×10^{-16}

Table 1: Novel susceptibility loci were previously identified (Shu et. al. Nature Comm, 2019). From these genes, we selected the human YBEY gene to evaluate a potential role in breast cancer risk. YBEY was significantly associated with breast cancer risk in SBCS, GTEx, TCGA, and METABRIC data. A negative beta coefficient suggests an inverse association (i.e. higher YBEY expression = less cancer risk)

Table 2. Putative causal SNPs associated with the human YBEY gene

ASN

Query SNP: rs35418111 and variants with r^2 >= 0.8

chr	pos (hg38)	LD (r^2)	LD (D')	variant	Ref	Alt	AFR freq	AMR freq	ASN freq	EUR freq	SiPhy cons	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed	NHGRI/EBI GWAS hits	GRASP QTL hits	Selected eQTL hits	GENCODE genes	dbSNP func annot
21	46420320	0.81	0.93	rs7275926	C	T	0.10	0.14	0.20	0.13			SPLN			7 altered motifs			74 hits	PCNT	intronic
21	46421683	0.81	0.93	rs13063115	G	A	0.07	0.14	0.20	0.13			IPSC, SPLN	IPSC		Smad3			76 hits	PCNT	intronic
21	46425328	0.81	0.94	rs8134832	A	G	0.55	0.24	0.20	0.23		BLD	5 tissues	BLD,BLD_THYM		9 altered motifs			47 hits	PCNT	intronic
21	46428994	0.83	0.94	rs71326328	G	A	0.08	0.14	0.20	0.13				IPSC		4 altered motifs			76 hits	PCNT	intronic
21	46430264	0.93	0.99	rs57385578	C	T	0.01	0.11	0.18	0.07			BRN, ADRL			ERalpha-a,NF-kappaB			45 hits	PCNT	intronic
21	46431882	0.84	0.94	rs2953522	G	A	0.07	0.14	0.20	0.13			4 tissues	ADRL,BRN		AhR::Arnt,Arnt,Pax-6			74 hits	PCNT	synonymous
21	46433109	0.84	0.94	rs2712582	C	T	0.07	0.14	0.20	0.13			BLD, ADRL			4 altered motifs			75 hits	PCNT	intronic
21	46435224	0.82	0.94	rs8126917	T	G	0.07	0.14	0.20	0.13			BRN, ADRL			9 altered motifs			76 hits	PCNT	intronic
21	46435963	0.83	0.94	rs17371795	A	G	0.07	0.14	0.20	0.13			BLD, THYM	IPSC,MUS				1 hit	75 hits	PCNT	synonymous
21	46436757	1	1	rs3541811	G	A	0.04	0.11	0.19	0.07			BLD, BRN, THYM	ADRL,PCNT,BRST		4 altered motifs			39 hits	PCNT	intronic
21	46437527	0.84	-0.95	rs7276499	C	T	0.90	0.75	0.88	0.81			5 tissues			5 altered motifs			27 hits	PCNT	intronic
21	46438861	0.83	-0.95	rs9981192	T	C	0.91	0.75	0.88	0.80						4 altered motifs			30 hits	PCNT	intronic
21	46442275	0.81	-0.93	rs10854482	T	C	0.89	0.75	0.88	0.81						4 altered motifs			34 hits	PCNT	intronic
21	46442711	0.83	-0.94	rs2839262	A	G	0.91	0.76	0.88	0.81			4 tissues	MUS					36 hits	PCNT	intronic
21	46451672	0.8	-0.94	rs1010111	G	A	0.73	0.70	0.80	0.73				KID,PLCNT		NRSF			16 hits	5.9kb 3' of PCNT	
21	46452693	0.82	-0.94	rs7280077	G	A	0.91	0.76	0.88	0.81			13 tissues	HRT,GI,GI					29 hits	6.2kb 5' of DIP2A	
21	46457092	0.82	-0.94	rs2130538	T	C	0.82	0.75	0.80	0.81		1 tissues	20 tissues	4 tissues		5 altered motifs			28 hits	1.5kb 5' of DIP2A	
21	46461396	0.81	-0.94	rs2078203	A	G	0.74	0.70	0.80	0.74		BLD	6 tissues	MUS		4 altered motifs			13 hits	DIP2A	intronic
21	46461722	0.81	-0.94	rs1981330	A	G	0.88	0.72	0.80	0.74		BLD	8 tissues			6 altered motifs			7 hits	DIP2A	intronic

EUR

Query SNP: rs35418111 and variants with r^2 >= 0.8

chr	pos (hg38)	LD (r^2)	LD (D')	variant	Ref	Alt	AFR freq	AMR freq	ASN freq	EUR freq	SiPhy cons	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed	NHGRI/EBI GWAS hits	GRASP QTL hits	Selected eQTL hits	GENCODE genes	dbSNP func annot
21	46430264	0.88	0.96	rs57385578	C	T	0.01	0.11	0.16	0.07			BRN, ADRL			ERalpha-a,NF-kappaB			45 hits	PCNT	intronic
21	46436757	1	1	rs3541811	G	A	0.04	0.11	0.19	0.07			BLD, BRN, THYM	ADRL,PLCNT,BRST		4 altered motifs			39 hits	PCNT	intronic

Table 2: Several SNPs in high LD with the index SNP (rs354) were identified and evaluated for potential causative function by EMSA. SNPs were prioritized based on the number of eQTL hits, motifs changed, and frequencies in the population of interest (ASN and EUR). Red rectangles represent variants that were evaluated by EMSA.

Figure 1. EMSA of potential causal SNPs associated with YBEY gene

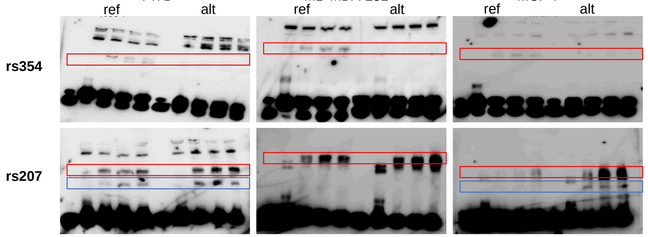


Figure 1: EMSA for SNPs associated with human YBEY. Three breast cancer cell lines were assayed. Red and blue rectangles represent differential signal intensities between reference and alternate sequence-bound nuclear proteins. Lane 1: Free probe Lane 2: Non-competing DNA + protein Lanes 3-5: Competitive DNA + protein.

YBEY SNV **rs354**18111 G → A
YBEY SNV **rs207**8203 A → G

Figure 2. YBEY knockdown inhibits breast cancer proliferation and colony formation

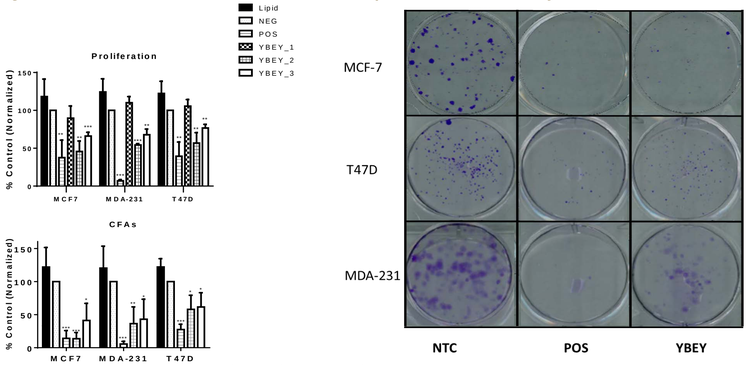


Figure 2: siRNA-mediated k.d. of YBEY inhibited proliferation and colony formation. Proliferation, assessed by alamarBlue assay, was significantly inhibited in all three cell lines following YBEY k.d. Colony formation was also significantly inhibited in breast cancer cell lines by 2 out of 3 siRNAs targeting YBEY mRNA. These data suggest a prime role of YBEY in tumor cell growth and cell cycling.

Figure 3. YBEY knockdown inhibits breast cancer migration and invasion

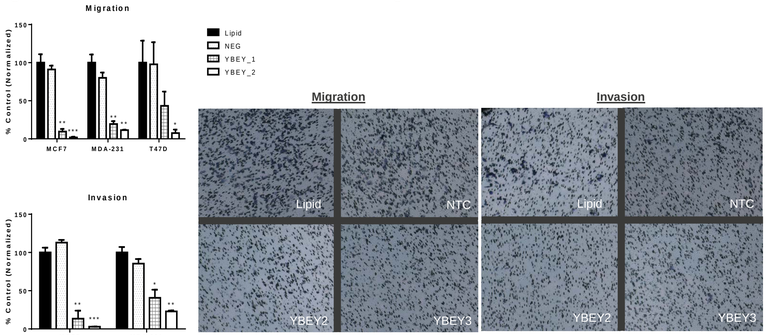


Figure 3: siRNA-mediated k.d. of YBEY inhibited migration and invasion. Boyden chamber assays were used to evaluate the ability of breast cancer cell lines to migrate through inserts with (invasion) and without (migration) Matrigel. Both siRNAs targeting YBEY mRNA were able to significantly disrupt cell migration and invasion in MCF-7, MB-MDA-231, and T47D cells. *T47D cells did not efficiently invade through Matrigel and were not included.

SUMMARY AND CONCLUSIONS

-Knockdown of YBEY in human breast cancer cell lines had a significant effect on cellular functions independent of hormone receptor status

-rs35418111 and rs2078203 may play a causal role in breast cancer risk

-Further investigations are necessary to elucidate the mechanisms by which YBEY influences carcinogenesis

ACKNOWLEDGMENTS

This work was supported by the Vanderbilt Institute for Clinical and Translational Research (VICTR) grant VR54254 to CS. CS is supported by the Vanderbilt Training Program in Molecular and Genetic Epidemiology of Cancer (MAGEC) program (T32CA160056).