University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Dugué, PA;Yu, C;McKay, T;Wong, EM;Joo, JE;Tsimiklis, H;Hammet, F;Mahmoodi, M;Theys, D;Kconfab,;Hopper, JL;Giles, GG;Milne, RL;Steen, JA;Dowty, JG;Nguyen‐dumont, T;Southey, MC

Title:
Vtrna2‐1: Genetic variation, heritable methylation and disease association

Date:
2021-03-01

Citation:
Dugué, P. A., Yu, C., McKay, T., Wong, E. M., Joo, J. E., Tsimiklis, H., Hammet, F., Mahmoodi, M., Theys, D., Kconfab, , Hopper, J. L., Giles, G. G., Milne, R. L., Steen, J. A., Dowty, J. G., Nguyen‐dumont, T. & Southey, M. C. (2021). Vtrna2‐1: Genetic variation, heritable methylation and disease association. International Journal of Molecular Sciences, 22 (5), pp.1-18. https://doi.org/10.3390/ijms22052535.

Persistent Link:
https://hdl.handle.net/11343/274489

License:
CC BY

International Journal of
Molecular Sciences

MDPI

Article

VTRNA2-1: Genetic Variation, Heritable Methylation and Disease Association

Pierre-Antoine Dugué [1,2,3], Chenglong Yu [1], Timothy McKay [1], Ee Ming Wong [1], Jihoon Eric Joo [4], Helen Tsimiklis [1], Fleur Hammet [1], Maryam Mahmoodi [1], Derrick Theys [1], kConFab [5], John L. Hopper [3], Graham G. Giles [1,2,3], Roger L. Milne [1,2,3], Jason A. Steen [1], James G. Dowty [3], Tu Nguyen-Dumont [1,4] and Melissa C. Southey [1,2,4,*]

1 Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC 3168, Australia; Pierre-Antoine.Dugue@monash.edu (P.-A.D.); chenglong.yu@monash.edu (C.Y.); tmck0012@student.monash.edu (T.M.); eeming.wong@monash.edu (E.M.W.); helen.tsimiklis@monash.edu (H.T.); fleur.hammet@monash.edu (F.H.); maryam.mahmoodi@monash.edu (M.M.); derrick.theys@monash.edu (D.T.); Graham.Giles@cancervic.org.au (G.G.G.); Roger.Milne@cancervic.org.au (R.L.M.); jason.steen@monash.edu (J.A.S.); tu.nguyen-dumont@monash.edu (T.N.-D.)
2 Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC 3004, Australia
3 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Parkville, VIC 3010, Australia; j.hopper@unimelb.edu.au (J.L.H.); jdowty@unimelb.edu.au (J.G.D.)
4 Department of Clinical Pathology, Melbourne Medical School, The University of Melbourne, Parkville, VIC 3010, Australia; ji.joo@unimelb.edu.au
5 Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer, Research Division, Peter MacCallum Cancer Centre, Melbourne, VIC 3000, Australia; heather.thorne@petermac.org
* Correspondence: melissa.southey@monash.edu

Abstract: *VTRNA2-1* is a metastable epiallele with accumulating evidence that methylation at this region is heritable, modifiable and associated with disease including risk and progression of cancer. This study investigated the influence of genetic variation and other factors such as age and adult lifestyle on blood DNA methylation in this region. We first sequenced the *VTRNA2-1* gene region in multiple-case breast cancer families in which *VTRNA2-1* methylation was identified as heritable and associated with breast cancer risk. Methylation quantitative trait loci (mQTL) were investigated using a prospective cohort study (4500 participants with genotyping and methylation data). The *cis*-mQTL analysis (334 variants ± 50 kb of the most heritable CpG site) identified 43 variants associated with *VTRNA2-1* methylation ($p < 1.5 \times 10^{-4}$); however, these explained little of the methylation variation ($R^2 < 0.5\%$ for each of these variants). No genetic variants elsewhere in the genome were found to strongly influence *VTRNA2-1* methylation. SNP-based heritability estimates were consistent with the mQTL findings ($h^2 = 0$, 95%CI: −0.14 to 0.14). We found no evidence that age, sex, country of birth, smoking, body mass index, alcohol consumption or diet influenced blood DNA methylation at *VTRNA2-1*. Genetic factors and adult lifestyle play a minimal role in explaining methylation variability at the heritable *VTRNA2-1* cluster.

Keywords: *VTRNA2-1*; nc886; MIR886; methylation quantitative trait loci; SNP-based heritability; rs2346018; breast cancer; prostate cancer

check for updates

Citation: Dugué, P.-A.; Yu, C.; McKay, T.; Wong, E.M.; Joo, J.E.; Tsimiklis, H.; Hammet, F.; Mahmoodi, M.; Theys, D.; k.; et al. *VTRNA2-1*: Genetic Variation, Heritable Methylation and Disease Association. *Int. J. Mol. Sci.* **2021**, *22*, 2535. https://doi.org/10.3390/ijms22052535

Academic Editors: Maria M. Sasiadek and Pawel Karpinski

Received: 10 February 2021
Accepted: 25 February 2021
Published: 3 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Mendelian-like inheritance of germline DNA methylation can be due to *cis*- or *trans*-acting genetic factors known as methylation Quantitative Trait Loci (mQTL) or epimutations (heritable change in gene activity that is not associated with a DNA mutation but rather with gain or loss of DNA methylation or other heritable modification of chromatin). Both can mimic germline pathogenic variants in their effect on gene function and disease

association and discriminating between the two possibilities (mQTL or epimutation) in specific genomic regions and disease context is often challenging.

We previously made a genome-wide assessment of heritable methylation using a family design [1]; probes were ranked by a methylation-heritability metric and 24 of the 1000 most heritable CpGs were identified to be associated with breast cancer risk in these families. Several CpGs within *VTRNA2-1* were among those that appeared to be most heritable and most strongly associated with breast cancer risk, including five within the gene promoter [1]. Evidence that DNA methylation can be transmitted from parent to offspring in the absence of a genetic explanation is scarce and controversial [2–4]. It is therefore important to assess whether any genetic variation may influence DNA methylation at the *VTRNA2-1* region, i.e., *cis*- or *trans*-acting mQTL.

VTRNA2-1 has been demonstrated to adopt two structurally and functionally distinct RNA conformations, one which strongly inhibits protein kinase R (PKR) and downstream eukaryotic translation initiation factor 2 subunit α (eIF2α) phosphorylation, and one which acts as a pseudo-inhibitor of PKR when competing with other double-stranded (ds) RNA molecules [5]. The ~2 kb region overlapping the *VTRNA2-1* locus has been reported to be polymorphic or atypically imprinted, and somatically acquires DNA methylation of the maternal allele in the majority of cases [6,7]. PKR is an interferon-induced kinase consisting of 551 amino acids that acts as an intracellular stress sensor, primarily associated with viral infection. dsRNA produced by viral replication binds to and activates PKR, causing dimerisation and subsequent phosphorylation of its substrate, eIF2α. Phosphorylation of eIF2α converts eIF2 from a substrate to an inhibitor of its GDP-GTP exchange factor eIF2B, inhibiting mRNA translation and arresting global protein synthesis [8,9]. In addition to eIF2α phosphorylation, PKR can activate the nuclear factor kappa-light-chain-enhancer of the activated B cells (NF-κB) signalling pathway, which is known to play an oncogenic role in tumourigenesis [10]. PKR is reported to act on the NF-κB pathway by inducing phosphorylation of IκBα, which interestingly requires the expression of *VTRNA2-1* to occur [9].

There is evidence that *VTRNA2-1* may act as a tumour suppressor and is a metastable epiallele [11,12]. The study by van Baak et al. used data from the Melbourne Collaborative Cohort Study (MCCS) to assess the association between methylation at metastable epialleles and cancer risk and concluded that methylation at *VTRNA2-1* was potentially associated with risk of lung cancer and B-cell lymphoma [13]. We also showed that *VTRNA2-1* promoter methylation was associated with prostate cancer risk, and these associations appeared stronger for aggressive disease [14].

Consistent with a familial aggregation of *VTRNA2-1* methylation not being due to genetic factors are observations that DNA methylation at this locus is sensitive to the perinatal environment, including factors such as season of conception [13,15] or maternal folic acid supplementation during pregnancy [16]. That *VTRNA2-1* methylation at birth was found to be associated with childhood overweight/obesity [17] may signal another plausible link with cancer risk in adulthood, in addition to the tumour suppressor role [15].

DNA methylation at *VTRNA2-1* has therefore been hypothesised to show the following pattern [7]: (1) The paternally-inherited allele seems to always be unmethylated, as observed in several studies [1,14], and (2) the maternally-inherited allele is methylated in ~75% and unmethylated in ~25% of individuals. Methylation at this locus is thought to be influenced by the aforementioned pre-/perinatal environmental factors, as well as by genetic variants, the latter possibly due the role of CTCF (transcription factor, CCCTC-binding) in imprinting via the influence of rs2346018. A recently published study of genome-wide mQTLs in 27,750 European participants [18] revealed relatively weak associations between genetic variants and methylation at cg26328633 (CpG site identified as part of the strong heritable cluster in our Australian families), but no data were available for neighbouring CpGs, and no apparent association was found for other single nucleotide polymorphisms (SNP) of interest within this region, including rs2346018. No mQTLs were

found for any of the most heritable *VTRNA2-1* DNA methylation marks in a previous genome-wide assessment that included ~2000 participants [19].

In this study, our aims were three-fold: first, to sequence the *VTRNA2-1* region to assess the presence of rare genetic variation at this locus; second, to conduct a genome-wide assessment of mQTLs and SNP-based methylation heritability in the *VTRNA2-1* region (previously identified heritable marks); third, to assess whether any genetic variants associated with DNA methylation in this region contribute to the previously observed associations with breast cancer risk.

2. Results

This study used data from (1) 179 participants of multiple-case breast cancer family studies to assess the presence of genetic variants in the *VTRNA2-1* heritable region, (2) 4500 participants in a prospective study to assess mQTLs and SNP-based heritability (genome-wide and focusing on *cis*-variants), and (3) 2141 participants in breast cancer family-based studies to adjust the *VTRNA2-1* results of our previous publication [1] for a nearby SNP.

The characteristics of participants in the prospective cohort (Melbourne Collaborative Cohort Study) and multiple-case family (Australian Breast Cancer Family Registry and Kathleen Cuningham Foundation Consortium for research into Familial Breast Cancer) studies (see Methods Section) included in this study are shown in Table 1. In the prospective cohort study, the majority were male, aged between 50 and 70 and never or former smokers. The distribution of methylation beta values at the *VTRNA2-1* region is shown in Figure 1b,c. Nine genetic variants were identified via sequencing (Figure 1a) in members of multiple-case breast cancer families. Of these, one was rare (identified in only one participant) and was excluded from further analysis. The other eight variants were used to estimate carrier probabilities in members of multiple-case breast cancer families (see Methods Section; *4.4.4. Associations with breast cancer risk*). All eight variants were available after genotype imputation in the prospective cohort study (see Methods Section; *4.2. Genetic and methylation data*) and therefore included in the mQTL analysis. None of these variants were found to directly overlap with the most heritable *VTRNA2-1* methylation site (cg06536614).

Table 1. Characteristics of the population-based and breast cancer family-based studies.

	Controls (*n* = 2272)	**Cases (*n* = 2228)**
Population-based study (MCCS)		
Age at blood draw (median [IQR])	59.8 [52.7–65.0]	60.3 [53.3–65.6]
Sex (female)	919 (40%)	901 (40%)
Country of birth		
Australia/NZ	1602 (70%)	1571 (69%)
UK/Northern Europe	161 (7%)	156 (7%)
Italy	311 (14%)	297 (13%)
Greece	198 (9%)	204 (9%)
Smoking status		
Never	1154 (51%)	1106 (49%)
Former	854 (38%)	888 (39%)
Current	263 (11%)	233 (10%)
BMI (kg/m2)	26.4 [24.1–29.1]	26.8 [24.4–29.7]
Alcohol consumption (g/day)	4.3 [0.0–17.1]	4.3 [0.0–17.1]
Alternate Healthy Eating Index 2010	63.5 [56.0–71.0]	64.0 [56.5–71.5]
Family-based studies (ABCFR/kConFab) [a]		
Sex (female)	123 (100%)	87 (100%)

Abbreviations: MCCS: Melbourne Collaborative Cohort Study; ABCFR/kConFab: Australian Breast Cancer Family Registry and Kathleen Cuningham Foundation Consortium for research into Familial Breast Cancer; IQR: interquartile range; BMI: body mass index. [a] Of these 210 participants, 179 had sequencing data available.

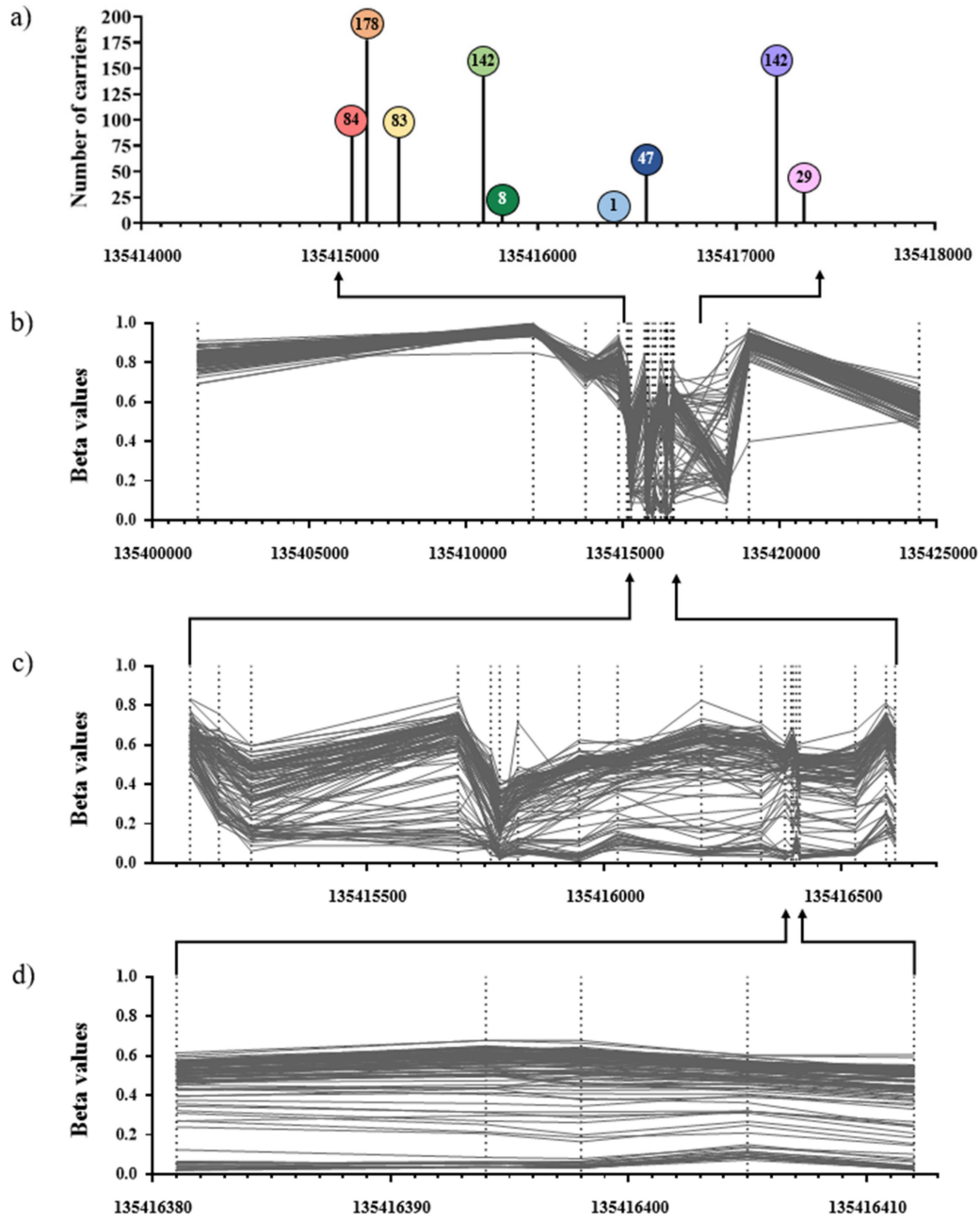


Figure 1. Genetic variants and DNA methylation at the *VTRNA2-1* region. Panel (**a**) shows the 9 variants identified via sequencing in the 179 participants of the multiple-case breast cancer family studies (ABCFR/kConFab). From left to right: rs62365993, rs7706795, rs2346018, rs2346019, rs34577747, rs1366231064, rs9327740, rs4976364, and rs74555710. Numbers within lollipops indicate the number of minor alleles observed in the sample. Panel (**b**)–(**d**) are from a random sample of 100 participants in the population-based study (MCCS), with coordinates: (**b**) 13401437–1354244 (26 CpGs), (**c**) 135415129–135416613 (19 CpGs), and (**d**) 135416381–135416412 (the 5 most heritable methylation marks).

Methylation values at the five most heritable methylation marks (cg06536614, cg26328633, cg25340688, cg26896946 and cg00124993) were highly correlated (in the MCCS all $r \geq 0.88$); we therefore focused on cg06536614, which was found to be the most heritable methylation mark [1]. The percentage methylation at cg06536614 was lower than 50% for 1946 (43%) of participants and lower than 60% for 4414 (98%) of them; 23% had less than 30% methylation (Figure 2).

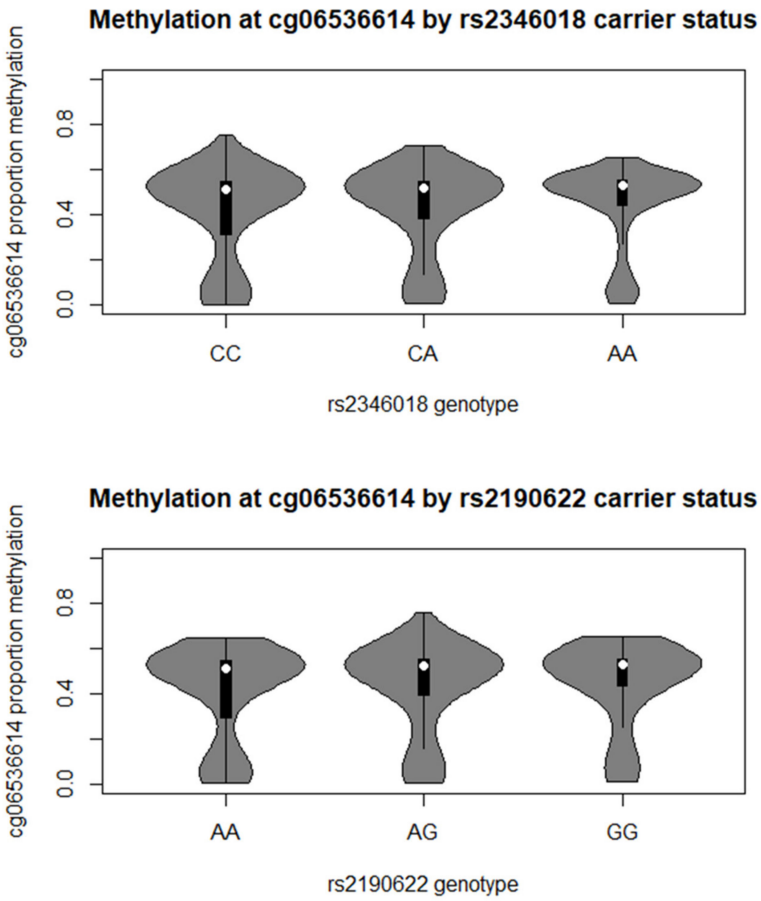


Figure 2. *VTRNA2-1* promoter percentage methylation distribution by carrier status at rs2346018 (CTCF binding site) and rs2190622 (strongest observed association). Carrier frequency (percentage): **rs2346018**: CC: 1969 (44%); CA: 1936 (43%); AA: 495 (11%); **rs2190622**: CC: 1838 (41%); CA: 2064 (46%); AA: 597 (13%).

The mQTL analysis was performed in 4500 participants in the prospective, population-based study (MCCS) for a total of 10,484,498 genetic variants; a *cis*-mQTL analysis was then carried out by focusing on genetic variants within 50 kb of the most heritable methylation mark (cg06536614). Genome-wide, we found no evidence that any included genetic variant was associated with *VTRNA2-1* methylation (all $p > 5 \times 10^{-9}$), with similar results obtained for the M-value (logit transformation of beta value) or RINT (rank-based inverse normal transformation, which was applied previously in the context of DNA methylation analyses and provides a Gaussian methylation distribution, which is not always the case for the M-values) transformation (Supplementary Table S1 and S2, respectively, showing the 100 CpGs with smallest *p*-values). Results from the *cis*-mQTL analysis, i.e., genetic variants within 50 kb of the CpG, are shown in Supplementary Table S3 (M-values) and Supplementary Table S4 (RINT-values). There were 43 variants with $p < 1.5 \times 10^{-4}$, indicating evidence of genetic influences on *VTRNA2-1* methylation (Table 2 and Figure 3). The strongest evidence of association was observed for rs2190622 ($p = 5 \times 10^{-6}$) (Figure 2), although this association was not substantially stronger than for other neighbouring SNPs. A significant association was also observed for rs2346018 ($p = 8 \times 10^{-5}$) (Figure 2), which was previously reported to modify methylation in this region [7]. Consistent with this, associations appeared stronger for the variants located closer to the CpG of interest (Supplementary Table S3 and Figure 3). Associations were qualitatively similar but appeared somewhat stronger for the RINT-transformed values (strongest hit: rs2190622, $p = 8 \times 10^{-8}$; rs2346018: $p = 8 \times 10^{-6}$ and Supplementary Table S4 and Figure 3). These variants appeared to explain very little of variation in *VTRNA2-1* methylation (Figure 2) with a variance explained ranging

from 0.33% to 0.47%. The findings were similar when restricting the analyses to MCCS participants selected as controls (not shown).

Table 2. Strongest mQTLs within 50 kb of cg06536614 (46 variants for the analysis of M-values).

Chromosome	Position	Variant	REF [a]	ALT [a]	BETA [a]	SE [a]	*P* [a]	R2 [a]	MAF [a]
5	135378781	rs3805700	A	G	0.093	0.024	1.2×10^{-4}	0.003	0.26
5	135402852	rs11956252	G	C	0.092	0.023	4.1×10^{-4}	0.004	0.32
5	135403529	rs6899012	G	A	0.093	0.022	3.9×10^{-4}	0.004	0.32
5	135403745	rs74634331	AGTG	AG	0.092	0.023	4.2×10^{-4}	0.004	0.32
5	135404173	rs9986124	G	T	0.093	0.023	4.1×10^{-4}	0.004	0.32
5	135404613	rs9986287	T	C	0.093	0.023	4.7×10^{-4}	0.004	0.32
5	135406459	rs7725702	C	G	0.097	0.022	1.2×10^{-4}	0.004	0.36
5	135406534	rs7725447	G	A	0.097	0.022	1.2×10^{-4}	0.004	0.35
5	135406658	rs2058043	A	G	0.098	0.022	1.0×10^{-4}	0.004	0.36
5	135406894	rs2058042	G	A	0.098	0.022	1.1×10^{-4}	0.004	0.35
5	135407572	rs4976470	A	G	0.099	0.022	8.0×10^{-4}	0.004	0.36
5	135408325	rs4976471	T	A	0.099	0.022	8.6×10^{-4}	0.004	0.36
5	135409014	rs6861956	T	C	0.096	0.022	1.5×10^{-4}	0.004	0.36
5	135410863	rs11742191	A	G	0.087	0.022	1.2×10^{-4}	0.003	0.33
5	135411281	rs11749522	C	T	0.087	0.022	1.1×10^{-4}	0.003	0.33
5	135412195	rs10079215	A	G	0.097	0.022	1.1×10^{-4}	0.004	0.36
5	135412675	rs35137944	A	G	0.097	0.022	1.0×10^{-4}	0.004	0.36
5	135413026	rs7724672	A	G	0.098	0.022	8.7×10^{-4}	0.004	0.36
5	135414280	rs2190622	A	G	0.100	0.022	4.6×10^{-4}	0.005	0.36
5	135414455	rs4246798	A	G	0.100	0.022	4.6×10^{-4}	0.005	0.36
5	135414510	rs4246799	G	A	0.100	0.022	5.3×10^{-4}	0.005	0.36
5	135414866	rs17169806	C	T	0.087	0.022	1.0×10^{-4}	0.003	0.33
5	**135415064**	**rs62365993**	**A**	**G**	**0.087**	**0.022**	**1.1** $\times 10^{-4}$	**0.003**	**0.33**
5	**135415300**	**rs2346018**	**C**	**A**	**0.089**	**0.023**	**8.2** $\times 10^{-4}$	**0.004**	**0.33**
5	**135415726**	**rs2346019**	**A**	**G**	**0.101**	**0.022**	**5.4** $\times 10^{-4}$	**0.005**	**0.36**
5	135417898	rs12653557	G	T	0.083	0.021	7.4×10^{-4}	0.004	0.49
5	135418032	rs917303	G	A	0.090	0.022	5.4×10^{-4}	0.004	0.34
5	135418717	rs4976472	G	C	0.083	0.021	6.6×10^{-4}	0.004	0.49
5	135419159	rs4976473	C	A	0.084	0.021	6.2×10^{-4}	0.004	0.49
5	135422443	rs11242311	T	C	0.086	0.021	3.9×10^{-4}	0.004	0.49
5	135422507	rs34835264	G	GA	−0.087	0.022	5.3×10^{-4}	0.004	0.50
5	135422598	rs11242312	G	A	0.086	0.021	4.2×10^{-4}	0.004	0.49
5	135422698	rs10900843	G	A	0.084	0.021	5.7×10^{-4}	0.004	0.49
5	135422738	rs10900844	A	G	0.086	0.021	4.2×10^{-4}	0.004	0.49
5	135422864	rs11242313	G	A	0.087	0.021	3.4×10^{-4}	0.004	0.49
5	135423029	rs11242314	T	C	0.086	0.021	4.4×10^{-4}	0.004	0.49
5	135424756	rs13186426	C	A	0.083	0.021	8.5×10^{-4}	0.004	0.48
5	135424847	5:135424847	A	AAT	0.083	0.021	7.0×10^{-4}	0.004	0.49
5	135424922	rs1465239	A	G	0.084	0.021	6.3×10^{-4}	0.004	0.49
5	135427371	rs1974552	T	A	0.089	0.022	8.1×10^{-4}	0.004	0.34
5	135429640	rs1558095	C	T	0.086	0.021	4.4×10^{-4}	0.004	0.49
5	135431590	rs1203219753	A	G	0.085	0.021	5.8×10^{-4}	0.004	0.49
5	135435140	rs1544486	C	T	0.087	0.021	$4.1 \times 10^{-4}5$	0.004	0.49

[a] Abbreviations: REF: Allele in the reference genome; ALT: Other allele found at that locus; BETA: Coefficient of the regression analysis of methylation on genetic variant; SE: standard error of BETA; *p*: *p*-value; R2: variance explained in methylation by genetic variant; MAF: Minor allele frequency. The three variants identified via sequencing that were available after OncoArray imputation are highlighted in bold (rs62365993, rs2346018, and rs2346019).

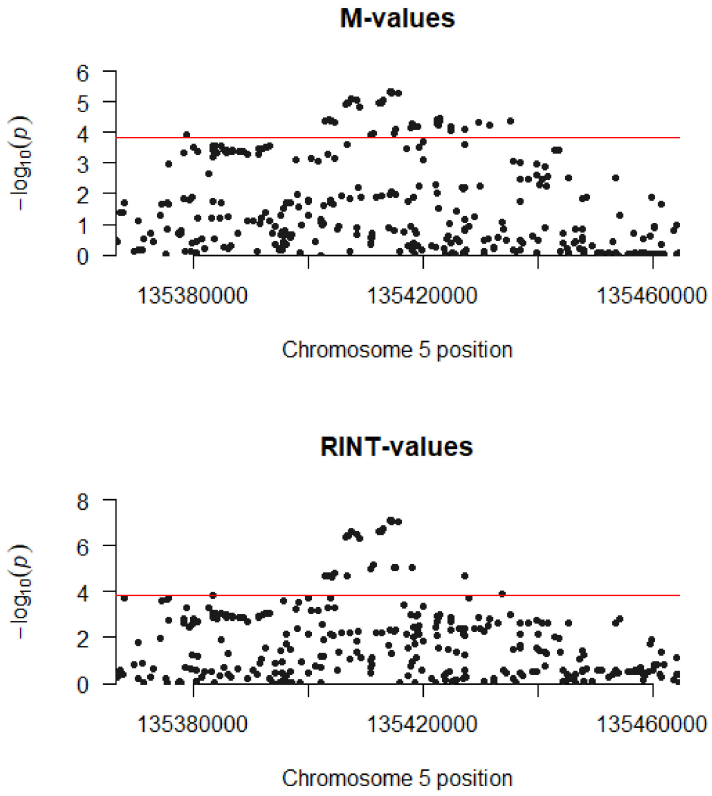


Figure 3. Manhattan plot for 334 genetic variants within 50 kb of cg06536614. The red line shows the Bonferroni threshold used to declare *cis*-mQTLs ($p = 1.5 \times 10^{-4}$); for the M-value analysis, the 46 significant associations are presented in Table 2. All variant names, positions, and quantitative results are shown in Supplementary Table S3 and S4.

Results from SNP-based heritability analyses (i.e., taking into account > 1M variants) are shown in Table 3. These were consistent with the mQTL analyses in showing minimal influence of SNPs on methylation in the region: $h^2 = 0$, 95%CI: −0.14 to 0.14. These results were virtually the same when the RINT transformation was used instead of M-values (Table 3) or restricting the analyses to MCCS participants selected as controls (not shown). Non-null or high heritability was observed for methylation sites distant from the heritable *VTRNA2-1* region (>2–8 kb, Table 3).

The associations between DNA methylation M-values and participant characteristics age, sex, country of birth and lifestyle factors (tobacco smoking, body mass index, alcohol consumption and healthy eating) are shown in Table 4. None of these factors showed an association with *VTRNA2-1* methylation; results were similar when considering only MCCS participants who were selected as controls or using RINT-values instead of M-values (results not shown). There was also no apparent influence of the sample proportion of white blood cells on DNA methylation (Table 4).

Table 3. SNP-based heritability (h^2) for 26 CpGs in the *VTRNA2-1* (*MIR886* in the Illumina HM450 annotation file) region, for M values and RINT values.

CpG	Chromosome	Position	Name	Location	Relation to Island	Enhancer	h^2 (M-Values)	95% CI (M-Values)	h^2 (RINT-Values)	95% CI (RINT-Values)
cg08836729	5	135401437				Yes	0	−0.14; 0.14	0	−0.14; 0.14
cg16402693	5	135412139			N_Shelf		0	−0.14; 0.14	0	−0.14; 0.14
cg17974054	5	135413810			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg11852404	5	135414858			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg16684184	5	135415129			Island		0	−0.14; 0.14	0.03	−0.11; 0.17
cg00308130	5	135415190			Island		0	−0.14; 0.14	0	−0.14; 0.14
cg15837280	5	135415258			Island		0	−0.14; 0.14	0	−0.14; 0.14
cg07158503	5	135415693			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg04515200	5	135415762			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg13581155	5	135415781			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg11978884	5	135415819			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg11608150	5	135415948			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg06478886	5	135416029			N_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg04481923	5	135416205	MIR886	Body	Island		0	−0.14; 0.14	0	−0.14; 0.14
cg18678645	5	135416331	MIR886	TSS200	Island		0	−0.14; 0.14	0	−0.14; 0.14
cg06536614	**5**	**135416381**	**MIR886**	**TSS200**	**Island**		**0**	**−0.14; 0.14**	**0**	**−0.14; 0.14**
cg26328633	**5**	**135416394**	**MIR886**	**TSS200**	**Island**		**0**	**−0.14; 0.14**	**0**	**−0.14; 0.14**
cg25340688	**5**	**135416398**	**MIR886**	**TSS200**	**Island**		**0**	**−0.14; 0.14**	**0**	**−0.14; 0.14**
cg26896946	**5**	**135416405**	**MIR886**	**TSS200**	**Island**		**0**	**−0.14; 0.14**	**0**	**−0.14; 0.14**
cg00124993	**5**	**135416412**	**MIR886**	**TSS200**	**Island**		**0**	**−0.14; 0.14**	**0**	**−0.14; 0.14**
cg08745965	5	135416529	MIR886	TSS1500	S_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg16615357	5	135416594	MIR886	TSS1500	S_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg18797653	5	135416613	MIR886	TSS1500	S_Shore		0	−0.14; 0.14	0	−0.14; 0.14
cg12897067	5	135418308			S_Shore		0.95	0.81; 1.09	0.76	0.62; 0.90
cg05631625	5	135419019			S_Shelf		0.14	0.00; 0.28	0.15	0.01; 0.29
cg01930756	5	135424444				Yes	0.09	−0.05; 0.23	0.1	−0.04; 0.24

The 5 CpGs in bold are those found to be most strongly heritable in our previous study Joo et al., Nat Commun, 2018 [1]. Name, Location, Relation to island, and Enhancer status are those provided by the Illumina HM450 annotation file.

Table 4. Association of non-genetic factors with *VTRNA2-1* blood DNA methylation (cg06536614) in the prospective, population-based study; 4500 participants in the Melbourne Collaborative Cohort Study (MCCS).

	Estimate [a]	95% CI	*p*-Value
Age (years)	−0.005	−0.012; 0.002	0.18
Sex (female)	0.047	−0.114; 0.207	0.57
Greece vs. Aus/NZ	−0.043	−0.250; 0.164	0.68
Italy vs. Aus/NZ	−0.068	−0.239; 0.102	0.43
Northern Europe vs. Aus/NZ	0.171	−0.049; 0.391	0.13
Current vs. never smoker	−0.031	−0.223; 0.162	0.76
Former vs. never smoker	−0.050	−0.176; 0.076	0.43
BMI (in kg/m2)	0.002	−0.012; 0.016	0.79
Alcohol consumption (g/day)	0.001	−0.003; 0.004	0.70
Healthy eating index	0.003	−0.003; 0.008	0.36
CD4 + T cells	−1.760	−6.108; 2.589	0.43
CD8+ T cells	−0.380	−4.054; 3.294	0.84
NK cells	−0.858	−5.242; 3.525	0.70
B cells	−2.616	−6.391; 1.159	0.17
Granulocytes	−1.852	−5.969; 2.264	0.38
Monocytes	−0.355	−4.699; 3.990	0.87

[a] Mixed linear regression model with methylation M-values as the outcome and mutually adjusted covariates modelled as fixed effects and technical variables study, assay plate and chip modelled as random effects.

Associations of *VTRNA2-1* methylation (cg06536614) with breast cancer risk were assessed in 2141 participants in multiple-case breast cancer families, using the same methods as in our previous publication [1]. As genotypes were not directly measured in all family members, we estimated carrier probabilities for genetic variants, using a method similar to that used to estimate methylation carrier probabilities [1], based on sequencing data for 179 participants and imputed OncoArray data for 23 participants (all multiple-case family members); the association of methylation with breast cancer risk was therefore assessed using the same models as previously, with additional adjustment for genetic variant carrier probabilities. The results "Not adjusted for SNPs" are the same as those presented in Joo et al. [1]. The association of cg06536614 methylation with breast cancer risk remained highly significant after adjustment for rs2346018 carrier probabilities, with *p* values ranging from 2×10^{-9} to 3×10^{-8} in the unadjusted model and from 8×10^{-9} to 1×10^{-7} in the adjusted model (Table 5). Similar results were obtained after adjustment for any of the other three variants significantly associated with cg06536614 (Table 2) for which carrier probabilities could be estimated (data not shown).

Table 5. Associations with breast cancer risk after adjustment for rs2346018 carrier probabilities [a].

CpG	Chromosome	Position	Δl [a]	Not Adjusted for SNPs [b]		Adjusted for rs2346018	
				Biased HR (95% CI) [c]	*p*-Value	Biased HR (95% CI) [c]	*p*-Value
cg06536614	5	135416381	143.6	3.1 (2.1–4.6)	7×10^{-9}	3.0 (2.0–4.3)	3×10^{-8}
cg00124993	5	135416412	108.0	3.2 (2.2–4.7)	2×10^{-8}	3.0 (2.0–4.4)	9×10^{-8}
cg26328633	5	135416394	107.5	3.2 (2.2–4.8)	2×10^{-8}	3.0 (2.0–4.5)	4×10^{-8}
cg25340688	5	135416398	105.9	3.2 (2.1–4.7)	3×10^{-8}	2.9 (2.0–4.3)	1×10^{-7}
cg26896946	5	135416405	92.1	3.6 (2.4–5.4)	2×10^{-9}	3.3 (2.2–5.0)	8×10^{-9}

[a] Δl: heritability metric: details of the methods used to calculate the heritability metric, carrier probabilities and Cox models for association with breast cancer risk are provided in [1]. [b] These are published and unpublished results from our study Joo et al., 2018 [1] and are presented here for comparison with the results adjusted for rs2346018. [c] While *p* values are unbiased, hazard ratios are biased by the ascertainment of families for this study [1], and the HR estimates are only included here to show that they are virtually unchanged by adjustment for rs2346018.

3. Discussion

Our study provides further evidence that DNA methylation at *VTRNA2-1* is minimally influenced by genetic factors, and thus, the Mendelian-like inheritance of germline DNA methylation at this locus is likely to be via a true epimutation mechanism rather than via a mQTL. Therefore, the "missing heritability" (approximately the difference between family-based and SNP-based heritability) appears to be substantial, which confirms the findings of Joo et al. [1]. Genetic variants are therefore unlikely to fully explain any associations between methylation at this locus and disease risk, including breast cancer as we found previously in the context of a multiple-case breast cancer family study.

Some genetic variants at this locus were statistically significantly associated with blood DNA methylation (carriers of the minor allele being less likely to show hypomethylation), but the effect sizes and variance explained were small (variance explained 0.4% to 0.5% for the strongest individual associations). Rs2346018 [6,7], which was previously implicated with methylation at this locus, showed a significant association with methylation in our data, thereby confirming it may exert a small influence. Although our study does not allow disentangling which of the identified SNPs might causally affect DNA methylation, it should be noted that (1) rs2346018 was one of the SNPs with smallest *p*-value among the 334 tested, and (2) many SNPs that appeared most strongly associated with methylation did not have a clear functional interpretation, for example rs2190622, the strongest observed association, is located in an intergenic region and has unknown regulatory function (Supplementary Table S5).

In their genome-wide assessment of imprinting in the methylome, Zink et al. concluded that *VTRNA2-1* is an example of a region with polymorphic imprinted methylation unrelated to SNP genotypes. In our study, a large number of participants (45%) had a percentage of methylation between 50% and 60%. Although this might appear to be inconsistent with maternal imprinting, this could indicate the limitation of the HM450 assay to measure DNA methylation with sufficient precision or DNA methylation at the paternal allele accumulated over the lifetime. However, technical validation using PyroMark (Pyrosequencing technical validation) produced similar methylation values, and we did not find evidence that *VTRNA2-1* methylation was influenced by factors for which our data shows widespread methylation changes, such as age, sex, country of birth, or other factors shown to strongly affect DNA methylation such as tobacco smoking, alcohol consumption or body mass index [20–22]. In this study, we did not have information on early-life factors. Although several factors in utero and early in life were shown to modulate DNA methylation at *VTRNA2-1*, none of the findings presented for season of conception [15], maternal folate during pregnancy [16], or gestational famine exposure [23] appear to fully explain methylation variation in this region; the mechanisms by which *VTRNA2-1* methylation is inherited therefore appear to be essentially non-genetic and only partially explained by the factors studied in the literature so far.

Previous studies of mQTLs and heritability have only provided a partial assessment of genetic influences on *VTRNA2-1* methylation. The study by Gaunt et al. [24] also did not report SNP-based heritability but assessed mQTLs. Various potential trans-mQTLs were identified but evidence of associations appeared weak (all $p > 3 \times 10^{-8}$, http://www.mqtldb.org/search.htm, accessed on 1 December 2020). The large meta-analysis by Min et al. only reported significant associations for cg26328633 based on 27,750 European participants; 12 SNPs located close to *VTRNA2-1* had a *p*-value ranging from 10^{-22} to 10^{-26} (http://mqtldb.godmc.org.uk/, accessed on 1 December 2020), but the effect sizes were virtually the same as those obtained in our study (beta ~0.1) so that the variance explained was likely similarly small. Results for rs2346018 or other variants of interest were not provided. In the study by McRae et al. [25], the family-based heritability (peripheral blood leukocytes) of our five most heritable marks associated with breast cancer was ~0.50. The study by van Dongen et al. [26] found very high twin-based heritability (~0.97) for the five *VTRNA2-1* CpGs but SNP-based heritability was reported as "NA" (due to convergence problems using GCTA, which might mean those values were in fact equal to

zero). The p value cut-off we used for detecting mQTLs genome-wide was conservative (strict Bonferroni correction, $p = 5 \times 10^{-9}$), but using other cut-offs commonly used in genome-wide association studies, such as $p = 5 \times 10^{-8}$, would have resulted in the same conclusion (M-values: one intergenic variant in chromosome 1, $p = 3 \times 10^{-8}$ and explaining 0.7% of methylation variability; RINT values: no variant with $p < 5 \times 10^{-8}$). It is possible that some true associations were not detected, but these would likely have a weak influence on DNA methylation. Other studies used various cut-offs to declare statistically significant mQTLs, e.g., $p = 10^{-14}$ in Gaunt et al. [24], $p = 10^{-11}$ in McRae et al. [19], and $p = 10^{-8}$ in Min et al. [18], but these studies carried out more tests because they investigated >400,000 CpGs.

Although findings are difficult to compare across studies, they appear to be consistent with ours in showing weak influences of genetic variants on *VTRNA2-1* methylation. Consistent with this, the associations of DNA methylation with breast cancer risk after adjustment for rs2346018 were only slightly attenuated. It should also be noted that none of rs2346018 or other sequenced genetic variants were found to be associated with breast cancer risk in the largest genome-wide association studies to date [27,28].

Although all methylation measures in this study were made on blood samples, it is worth noting that *VTRNA2-1* methylation has been implicated as playing a role in, or being influenced during, carcinogenesis. Fort et al. [29] sought to find direct association between *VTRNA2-1* transcript levels and methylation of its promoter in prostate tumour samples. Average *VTRNA2-1* promoter methylation was found to be substantially increased in both low-grade and metastatic tumour tissue compared with normal prostate tissue. Additionally, average *VTRNA2-1* promoter methylation appeared to correlate with Gleason score, clinical T-value and biochemical relapse [29]. The levels of *VTRNA2-1* transcript were found to be inversely correlated with average promoter methylation. The relationship between *VTRNA2-1* and cancer growth appears to be tissue specific, with several studies suggesting a tumour suppressive role, e.g., cholangiocarcinoma [30], oesophageal carcinoma [31], small cell lung cancer [32], gastric cancer [33] and acute myeloid leukaemia [34], and some suggesting an oncogenic role, e.g., in endometrial cancer [35] and thyroid cancer [36].

We conclude that the genetic and non-genetic factors we investigated play a minimal role in explaining variation in blood DNA methylation at *VTRNA2-1*, so these are unlikely to play a strong role in observed associations between *VTRNA2-1* methylation and disease risk. The mechanism of inheritance of DNA methylation in this region remains to be elucidated.

4. Materials and Methods

4.1. Data Sources

4.1.1. Prospective Cohort Study

The Melbourne Collaborative Cohort Study (MCCS) is a community-based study that recruited 41,513 participants in 1990–1994 [37]. Several nested case-control studies were conducted to assess associations between DNA methylation in blood and the risk of eight types of cancer. Incident cases were matched to controls on age, sex, country of birth and sample type (buffy coats/dried blood spots/peripheral blood mononuclear cells) using incidence density sampling [37,38]. We also used questionnaire-collected data on smoking and alcohol consumption [20–22], measures of body mass index and derived a healthy eating index using a validated 121-item food frequency questionnaire [39,40].

4.1.2. Multiple-Case Breast Cancer Families

A total of 210 individuals from 25 multi-generational multiple-case breast cancer families, including 20 from Kathleen Cuningham Foundation Consortium for research into Familial Breast Cancer (kConFab) and 5 from the Australian Breast Cancer Family Registry (ABCFR)), were included in this study [41–43]. Among these family members, there were 87 breast cancer cases and 123 unaffected relatives.

4.2. Genetic and DNA Methylation Data

The *VTRNA2-1* region (GRCh37, ch5:g.135414615-135417597) was screened in 179 of 210 individuals from a multiple-case breast cancer family by targeted-sequencing using a custom-designed HaloPlexHS panel (Agilent, Santa Clara, CA, USA). Libraries were prepared from blood-derived DNA according to the manufacturer's instructions and sequenced on a 2 × 150 bp high-output flow cell on the HiSeq3000 (Illumina, San Diego, CA, USA). Paired-end reads were aligned to the human reference genome GRCh37 using BWA-mem 0.7.17 [44]. Adapter sequences were removed, and unique molecular indices were marked for downstream read-deduplication using the Agilent Genomics NextGen Toolkit (Agilent, Santa Clara, CA, USA). Target coverage was calculated using bedtools v.2.27.1 [45], and variant calling was performed using VarDict v.1.7 [46]. Genetic variants filtering was performed as described previously [47]. Variant annotation was performed on variants with a read depth ≥30× and a variant allele frequency (VAF) ≥0.15, using VarSeq VSClinical v2.2 (Golden Helix Inc., Bozeman, MT, USA). Finally, an additional 23 participants in ABCFR/kConFab had genetic measures made using OncoArray, using the same method as for the MCCS for genotype imputation (see next paragraph).

Genome-wide genotyping was conducted on blood DNA samples from 12,584 MCCS participants using the Infinium OncoArray-500K BeadChip (Illumina, San Diego, CA, USA) [37,48]. Following previous standardised protocols [28], we imputed autosomal genotypes using the Michigan imputation server [49] and IMPUTE version 2 [50] with the 1000 Genomes Project dataset (phase 3) as the reference panel. The genotype probabilities from imputation were used to hard-call (uncertainty < 0.1) the genotypes for variants with an imputation info score > 0.3. For the current analysis, we included 4748 participants for whom DNA methylation data was also available. We then retained the hard-called variants with minor allele frequency >0.001, missing genotype rate <0.2 and Hardy–Weinberg equilibrium p-value > 10^{-6}. Furthermore, to avoid bias due to confounding by shared environment among close relatives, participants were removed based on relatedness, i.e., excluding one participant randomly selected from any pair with a genetic relationship ≥ 0.05 (4th-degree or closer relationship) [51,52]. This procedure also removed duplicated methylation samples (genetic relationship = 1) [38]. After these quality control steps, 4500 paired genetic-methylation samples were retained (including 2228 cancer cases and 2272 controls) and 10,484,498 genetic variants (including 9,551,474 SNPs) for the analysis.

For all samples, DNA methylation was measured using the HumanMethylation450 (HM450) BeadChip (Illumina, San Diego, CA, USA) using methods described previously [1,38,53]. We used methylation M-values as their distribution is usually closer to Gaussian than methylation beta-values [54]. As a sensitivity analysis, we performed a more direct normalization of beta-values using rank-based inverse normal transformation (RINT) which was applied previously in the context of DNA methylation analyses and provides a Gaussian methylation distribution, which is not the case for the M-values [55,56].

4.3. Technical Validation of Methylation Measures Using Pyrosequencing

Pyrosequencing (PSQ) conducted on the PyroMark Q48 (Qiagen, Hilden, Germany) was used to validate the methylation measures made on the HM450 assay. DNA was extracted as described previously [1] and bisulfite converted using Zymo (Zymo Research, Irvine, CA, USA). Forward, reverse and sequencing PCR primers were designed using PyroMark Assay Design Software version 2.0. (Forward primer sequence: 5Bios 5′-G/GGAGGAATTGAGAGTTTTTTTAGGATA-3′; Reverse primer sequence: 5′-CCTTCAA-AATAACACCAACTTATATTATCA-3′; Sequencing primer sequence: 5′-ACATAAAAAAA-TCAATAAACACC-3′) to target cg04481923 and were synthesised by Integrated DNA Technologies (Coralville, IA, USA). The EpiTect PCR control DNA set (Qiagen, Hilden, Germany), which includes a completely methylated and a completely unmethylated bisulfite-converted control DNAs, was used to generate a standard curve. The EpiTect control DNAs were mixed in known ratios (0, 0.25, 0.50, 0.75 and 1) and run with ten test samples, along with a non-converted DNA sample and a no-template control. The PyroMark PCR cycling

protocol was as follows: denaturation for 15 min at 95 °C, then 45 cycles of 30 s at 94 °C, 30 s at 56 C, and 30 s at 72 °C, then final extension for 10 min at 72 °C. For each sample, the raw percentage of methylation was determined at cg04481923 (*VTRNA2-1*) and calibrated using the EpiTect control DNAs standard curve.

4.4. Statistical Analysis

4.4.1. Assessment of mQTLs

After QC of methylation and genetic data, 4500 participants (including 2228 cancer cases and 2272 controls) and 10,484,498 genetic variants (including 9551,474 SNPs) were available for the analysis. We first removed factors that may confound DNA methylation values using linear mixed models with methylation M-values (or RINT-values) as the outcome and as covariates: age, sex, sample type, white blood cell proportions (estimated using the Houseman algorithm [57,58]) and 20 genetic principal components to account for population structure/ancestry as fixed effects; and as random effects: study, plate and slide of the assay. Our sample therefore included both cancer cases and controls. The inclusion of cancer cases may bias mQTL associations because of collider bias [59]; collider bias is usually considered to be small [60,61], and this may be particularly true in our setting because no strong associations of individual methylation markers with cancer risk were observed, but we nevertheless assessed consistency of associations in controls by analysing them separately.

As slightly over 10M genetic variants were tested, we used the Bonferroni correction for multiple testing and considered associations with a p-value less than $0.05/10^{-7} = 5 \times 10^{-9}$ to be potential true signals. Further, because *cis* acting genetic variants are considerably more likely than *trans* acting variants to influence DNA methylation, we considered all SNPs within 50 kb pairs of the methylation sites analysed. A total of 334 variants were identified, so we corrected the *cis*-mQTL analyses for multiple testing using the Bonferroni cut-off $p = 1.5 \times 10^{-4}$ (0.05/334).

4.4.2. SNP-Based Heritability

The univariate genome-based restricted maximum likelihood (GREML) method [62,63] was used to estimate the SNP-based heritability of methylation values in the sample of 4500 participants and a subsample of 2272 controls, respectively. The M-values (or RINT-values) after removing confounding effects were used as phenotypes in these analyses. We used only 1050,921 HapMap3 SNPs as they have been shown reliable and robust to bias in estimating SNP-based heritability and genetic correlations [64–66]. A genetic relationship matrix based on these SNPs was created and implemented in GREML. The heritability analyses were performed using the software GCTA [62].

4.4.3. Association of Non-Genetic Factors with VTRNA2-1 Methylation

We used mixed linear regressions similar to our previous publications [20,21] to assess the association of age, sex, BMI, smoking, alcohol consumption with methylation at *VTRNA2-1.* This analysis was undertaken using the same set of 4500 participants used for the genetic analyses, as well as separately in MCCS controls only.

4.4.4. Associations with Breast Cancer Risk

Cox proportional hazards survival analysis was used to test for associations between variants associated with *VTRNA2-1* methylation and breast cancer risk using all participants from the 25 multi-generational multiple-case breast cancer families (n = 2141). This analysis was based on the phenotype and relationships data of these 2141 participants, and the methylation and genetic data on 202 of them. Unobserved methylation and SNP data were replaced by estimated carrier probabilities using the methods presented in [1]. As the families in this study were ascertained because they each contained multiple breast cancer cases, and no adjustment for this ascertainment criterion was made, hazard ratio estimates are biased, but since the ascertainment criterion has no effect on the test statistic under the

null hypothesis, the *p* values for association with breast cancer are valid. These *p* values were based on the likelihood ratio test, not the Wald test, so variances for the hazard ratios were not needed and hence were not estimated using either standard maximum likelihood or robust variance. The same models as in Joo et al. [1] were performed, with additional adjustment for genetic variant carrier probabilities at rs2346018, which was one of the strongest mQTLs in this study (see Results), and previously reported to influence *VTRNA2-1* promoter methylation [7]). Similar results were obtained after adjustment for any of the other eight variants for which carrier probabilities could be estimated (data not shown).

Supplementary Materials: The following are available online at https://www.mdpi.com/1422-0067/22/5/2535/s1, Table S1. Genetic variants (genome-wide, 10,484,498 variants tested) with stronger evidence (top 100, ordered by P-value) of association with methylation M-values at cg06536614 (promoter of *VTRNA2-1*). Table S2. Genetic variants (genome-wide, 10,484,498 variants tested) with stronger evidence (top 100, ordered by P-value) of association with methylation RINT values at cg06536614 (promoter of *VTRNA2-1*). Table S3. Association between 334 genetic variants within 50kb of cg06536614 (promoter of *VTRNA2-1)* and methylation M-values at cg06536614 (promoter of *VTRNA2-1*). Table S4. Association between 334 genetic variants within 50kb of cg06536614 (promoter of *VTRNA2-1)* and methylation RINT values at cg06536614.

Author Contributions: Conceptualization, P.-A.D., C.Y., J.G.D., T.N.-D., and M.C.S.; methodology, P.-A.D., C.Y., J.G.D., T.N.-D., and M.C.S.; formal analysis, P.-A.D., C.Y., J.G.D.; resources, kConFab, J.L.H., G.G.G., R.L.M., and M.C.S.; data curation, T.M., E.M.W., J.E.J., H.T., F.H., M.M., D.T., J.A.S., and T.N.-D.; writing—original draft preparation, P.-A.D., C.Y., TM, J.G.D., T.N.-D., and M.C.S.; writing—review and editing, all authors; supervision, M.C.S.; project administration, M.C.S.; funding acquisition, J.L.H., G.G.G., R.L.M., and M.C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the U.S. National Institute of Health (grant number RO1CA-159868). The ABCFR was supported in Australia by the National Health and Medical Research Council, the New South Wales Cancer Council, the Victorian Health Promotion Foundation, the Victorian Breast Cancer Research Consortium, Cancer Australia, and the National Breast Cancer Foundation. The six sites of the Breast Cancer Family Registry (BCFR) were supported by grant UM1 CA164920 from the U.S. National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centres in the BCFR, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government or the BCFR. The MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 251553, 504711, 396414, 1074383, and by infrastructure provided by Cancer Council Victoria. The MCCS methylation studies were supported by the NHMRC grants 1011618, 1026892, 1027505, 1050198, 1043616, 1074383 and 1164455.TN-D is a National Breast Cancer Foundation (Australia) Career Development Fellow (ECF-17-001), M.C.S. is a National Health and Medical Research Council (NMHRC, Australia) Senior Research Fellow (APP1155163). This work was supported by an NHMRC Program grant (APP1074383), The National Breast Cancer Foundation (BRA-STRAP; NT-15-016), NHMRC European Union Collaborative Research Grant (APP1101400) and Monash University, Melbourne, Australia. kConFab is supported by a grant from the National Breast Cancer Foundation and previously by the National Health and Medical Research Council (NHMRC); the Queensland Cancer Fund; the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia; and the Cancer Foundation of Western Australia.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Human Research Ethics Committees of the University of Melbourne (1441955, 2014) and Monash University (21926, 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data will be made available upon reasonable request to the corresponding author.

Acknowledgments: We thank all the participants in this study, the entire team of Australian Breast Cancer Family Registry (BCFR-AU) and past and current investigators. We thank Heather Thorne, Eveline Niedermayr, Sharon Guo, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study for their contributions to this resource, and the many families who contribute to kConFab.

Conflicts of Interest: Authors have no conflicts of interest to declare.

References

1. Joo, J.E.; Dowty, J.G.; Milne, R.L.; Wong, E.M.; Dugue, P.A.; English, D.; Hopper, J.L.; Goldgar, D.E.; Giles, G.G.; Southey, M.C.; et al. Heritable DNA methylation marks associated with susceptibility to breast cancer. *Nat. Commun.* **2018**, *9*, 867. [CrossRef]
2. Heard, E.; Martienssen, R.A. Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* **2014**, *157*, 95–109. [CrossRef] [PubMed]
3. Jablonka, E.; Raz, G. Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* **2009**, *84*, 131–176. [CrossRef]
4. Horsthemke, B. A critical view on transgenerational epigenetic inheritance in humans. *Nat. Commun.* **2018**, *9*, 1–4. [CrossRef] [PubMed]
5. Calderon, B.M.; Conn, G.L. Human noncoding RNA 886 (nc886) adopts two structurally distinct conformers that are functionally opposing regulators of PKR. *RNA* **2017**, *23*, 557–566. [CrossRef] [PubMed]
6. Romanelli, V.; Nakabayashi, K.; Vizoso, M.; Moran, S.; Iglesias-Platas, I.; Sugahara, N.; Simón, C.; Hata, K.; Esteller, M.; Court, F.; et al. Variable maternal methylation overlapping the nc886/vtRNA2-1 locus is locked between hypermethylated repeats and is frequently altered in cancer. *Epigenetics* **2014**, *9*, 783–790. [CrossRef]
7. Carpenter, B.L.; Zhou, W.; Madaj, Z.; DeWitt, A.K.; Ross, J.P.; Grønbæk, K.; Liang, G.; Clark, S.J.; Molloy, P.L.; Jones, P.A. Mother–child transmission of epigenetic information by tunable polymorphic imprinting. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E11970–E11977. [CrossRef] [PubMed]
8. Sadler, A.J.; Williams, B.R.G. Structure and function of the protein Kinase R. In *Interferon: The 50th Anniversary*; Springer: Berlin/Heidelberg, Germany, 2007; Volume 316, pp. 253–292.
9. Golec, E.; Lind, L.; Qayyum, M.; Blom, A.M.; King, B.C. The noncoding RNA nc886 regulates PKR signaling and cytokine production in human cells. *J. Immunol.* **2019**, *202*, 131–141. [CrossRef]
10. Vlahopoulos, S.A. Aberrant control of NF-κB in cancer permits transcriptional and phenotypic plasticity, to curtail dependence on host tissue: Molecular mode. *Cancer Biol. Med.* **2017**, *14*, 254.
11. Rakyan, V.K.; Blewitt, M.E.; Druker, R.; Preis, J.I.; Whitelaw, E. Metastable epialleles in mammals. *Trends Genet.* **2002**, *18*, 348–351. [CrossRef]
12. Jirtle, R.L.; Skinner, M.K. Environmental epigenomics and disease susceptibility. *Nat. Rev. Genet.* **2007**, *8*, 253–262. [CrossRef]
13. Van Baak, T.E.; Coarfa, C.; Dugué, P.-A.; Fiorito, G.; Laritsky, E.; Baker, M.S.; Kessler, N.J.; Dong, J.; Duryea, J.D.; Silver, M.J.; et al. Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol.* **2018**, *19*, 1–20. [CrossRef] [PubMed]
14. Dugué, P.-A.; Dowty, J.G.; Joo, J.E.; Wong, E.M.; Makalic, E.; Schmidt, D.F.; English, D.R.; Hopper, J.L.; Pedersen, J.; Severi, G.; et al. Heritable methylation marks associated with breast and prostate cancer risk. *Prostate* **2018**, *78*, 962–969. [CrossRef] [PubMed]
15. Silver, M.J.; Kessler, N.J.; Hennig, B.J.; Dominguez-Salas, P.; Laritsky, E.; Baker, M.S.; Coarfa, C.; Hernandez-Vargas, H.; Castelino, J.M.; Routledge, M.N.; et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* **2015**, *16*, 118. [CrossRef]
16. Richmond, R.C.; Sharp, G.C.; Herbert, G.; Atkinson, C.; Taylor, C.; Bhattacharya, S.; Campbell, D.; Hall, M.; Kazmi, N.; Gaunt, T.; et al. The long-term impact of folic acid in pregnancy on offspring DNA methylation: Follow-up of the Aberdeen Folic Acid Supplementation Trial (AFAST). *Int. J. Epidemiol.* **2018**, *47*, 928–937. [CrossRef] [PubMed]
17. Van Dijk, S.J.; Peters, T.J.; Buckley, M.; Zhou, J.; Jones, P.A.; Gibson, R.A.; Makrides, M.; Muhlhausler, B.S.; Molloy, P.L. DNA methylation in blood from neonatal screening cards and the association with BMI and insulin sensitivity in early childhood. *Int. J. Obes.* **2017**, *42*, 28–35. [CrossRef]
18. Min, J.L.; Hemani, G.; Hannon, E.; Dekkers, K.F.; Castillo-Fernandez, J.; Luijk, R.; Carnero-Montoro, E.; Lawson, D.J.; Burrows, K.; Suderman, M. Genomic and phenomic insights from an atlas of genetic effects on DNA methylation. *medRxiv* **2020**. [CrossRef]
19. McRae, A.F.; Marioni, R.E.; Shah, S.; Yang, J.; Powell, J.E.; Harris, S.E.; Gibson, J.; Henders, A.K.; Bowdler, L.; Painter, J.N.; et al. Identification of 55,000 replicated DNA methylation QTL. *Sci. Rep.* **2018**, *8*, 1–9. [CrossRef]
20. Dugué, P.-A.; Jung, C.-H.; Joo, J.E.; Wang, X.; Wong, E.M.; Makalic, E.; Schmidt, D.F.; Baglietto, L.; Severi, G.; Southey, M.C.; et al. Smoking and blood DNA methylation: An epigenome-wide association study and assessment of reversibility. *Epigenetics* **2020**, *15*, 358–368. [CrossRef]
21. Dugué, P.; Wilson, R.; Lehne, B.; Jayasekara, H.; Wang, X.; Jung, C.; Joo, J.E.; Makalic, E.; Schmidt, D.F.; Baglietto, L.; et al. Alcohol consumption is associated with widespread changes in blood DNA methylation: Analysis of cross-sectional and longitudinal data. *Addict. Biol.* **2019**, e12855. [CrossRef]
22. Geurts, Y.M.; Dugué, P.-A.; Joo, J.E.; Makalic, E.; Jung, C.-H.; Guan, W.; Nguyen, S.; Grove, M.L.; Wong, E.M.; Hodge, A.M.; et al. Novel associations between blood DNA methylation and body mass index in middle-aged and older adults. *Int. J. Obes.* **2018**, *42*, 887–896. [CrossRef] [PubMed]

23. Finer, S.; Iqbal, M.S.; Lowe, R.; Ogunkolade, B.W.; Pervin, S.; Mathews, C.; Smart, M.; Alam, D.S.; Hitman, G.A. Is famine exposure during developmental life in rural Bangladesh associated with a metabolic and epigenetic signature in young adulthood? A historical cohort study. *BMJ Open* **2016**, *6*, e011768. [CrossRef]
24. Gaunt, T.R.; Shihab, H.A.; Hemani, G.; Min, J.L.; Woodward, G.; Lyttleton, O.; Zheng, J.; Duggirala, A.; McArdle, W.L.; Ho, K.; et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biol.* **2016**, *17*, 1–14. [CrossRef]
25. McRae, A.F.; Powell, J.E.; Henders, A.K.; Bowdler, L.; Hemani, G.; Shah, S.; Painter, J.N.; Martin, N.G.; Visscher, P.M.; Montgomery, G.W. Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol.* **2014**, *15*, R73. [CrossRef]
26. Van Dongen, J.; Nivard, M.; Willemsen, G.; Hottenga, J.J.; Helmer, Q.; Dolan, C.; Ehli, E.; Davies, G.; Van Iterson, M.; Breeze, C.E.; et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat. Commun.* **2016**, *7*, 11115. [CrossRef]
27. Zhang, H.; kConFab Investigators; Ahearn, T.U.; LeCarpentier, J.; Barnes, D.; Beesley, J.; Qi, G.; Jiang, O.; O'Mara, T.A.; Zhao, N.; et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* **2020**, *52*, 572–581. [CrossRef] [PubMed]
28. Michailidou, K.; Lindström, S.; Dennis, J.; Beesley, J.; Hui, S.; Kar, S.; Lemaçon, A.; Soucy, P.; Glubb, D.; Rostamianfar, A. Association analysis identifies 65 new breast cancer risk loci. *Nature* **2017**, *551*, 92. [CrossRef] [PubMed]
29. Fort, R.S.; Mathó, C.; Geraldo, M.V.; Ottati, M.C.; Yamashita, A.S.; Saito, K.C.; Leite, K.R.M.; Méndez, M.; Maedo, N.; Méndez, L.; et al. Nc886 is epigenetically repressed in prostate cancer and acts as a tumor suppressor through the inhibition of cell growth. *BMC Cancer* **2018**, *18*, 127. [CrossRef]
30. Kunkeaw, N.; Jeon, S.H.; Lee, K.; Johnson, B.H.; Tanasanvimon, S.; Javle, M.; Pairojkul, C.; Chamgramol, Y.; Wongfieng, W.; Gong, B.; et al. Cell death/proliferation roles for nc886, a non-coding RNA, in the protein kinase R pathway in cholangiocarcinoma. *Oncogene* **2013**, *32*, 3722–3731. [CrossRef]
31. Lee, H.-S.; Lee, K.; Jang, H.-J.; Lee, G.K.; Park, J.-L.; Kim, S.-Y.; Kim, S.-B.; Johnson, B.H.; Zo, J.I.; Lee, J.-S.; et al. Epigenetic silencing of the non-coding RNA nc886 provokes oncogenes during human esophageal tumorigenesis. *Oncotarget* **2014**, *5*, 3472–3481. [CrossRef]
32. Cao, J.; Song, Y.; Bi, N.; Shen, J.; Liu, W.; Fan, J.; Sun, G.; Tong, T.; He, J.; Shi, Y. DNA methylation-mediated repression of miR-886-3p predicts poor outcome of human small cell lung cancer. *Cancer Res.* **2013**, *73*, 3326–3335. [CrossRef] [PubMed]
33. Lee, K.-S.; Park, J.-L.; Lee, K.; Richardson, L.E.; Johnson, B.H.; Lee, H.-S.; Lee, J.-S.; Kim, S.-B.; Kwon, O.-H.; Song, K.S.; et al. nc886, a non-coding RNA of anti-proliferative role, is suppressed by CpG DNA methylation in human gastric cancer. *Oncotarget* **2014**, *5*, 3944–3955. [CrossRef] [PubMed]
34. Treppendahl, M.B.; Qiu, X.; Søgaard, A.; Yang, X.; Nandrup-Bus, C.; Hother, C.; Andersen, M.K.; Kjeldsen, L.; Möllgaard, L.; Hellström-Lindberg, E.; et al. Allelic methylation levels of the noncoding VTRNA2-1 located on chromosome 5q31.1 predict outcome in AML. *Blood* **2012**, *119*, 206–216. [CrossRef]
35. Hu, Z.; Zhang, H.; Tang, L.; Lou, M.; Geng, Y. Silencing nc886, a non-coding RNA, induces apoptosis of human endometrial cancer cells-1A in vitro. *Med. Sci. Monit.* **2017**, *23*, 1317–1324. [CrossRef]
36. Lee, E.K.; Hong, S.-H.; Shin, S.; Lee, H.-S.; Lee, J.-S.; Park, E.J.; Choi, S.S.; Min, J.W.; Park, D.; Hwang, J.-A.; et al. nc886, a non-coding RNA and suppressor of PKR, exerts an oncogenic function in thyroid cancer. *Oncotarget* **2016**, *7*, 75000–75012. [CrossRef]
37. Milne, R.L.; Fletcher, A.S.; MacInnis, R.J.; Hodge, A.M.; Hopkins, A.H.; Bassett, J.K.; Bruinsma, F.J.; Lynch, B.M.; Dugue, P.A.; Jayasekara, H.; et al. Cohort profile: The Melbourne collaborative cohort study (Health 2020). *Int. J. Epidemiol.* **2017**, *46*, 1757–1757i. [CrossRef]
38. Dugué, P.-A.; English, D.R.; MacInnis, R.J.; Jung, C.-H.; Bassett, J.K.; Fitzgerald, L.M.; Wong, E.M.; Joo, J.E.; Hopper, J.L.; Southey, M.C.; et al. Reliability of DNA methylation measures from dried blood spots and mononuclear cells using the human methylation 450k bead array. *Sci. Rep.* **2016**, *6*, 30317. [CrossRef]
39. Dugué, P.-A.; Hodge, A.M.; Brinkman, M.T.; Bassett, J.K.; Shivappa, N.; Hebert, J.R.; Hopper, J.L.; English, D.R.; Milne, R.L.; Giles, G.G. Association between selected dietary scores and the risk of urothelial cell carcinoma: A prospective cohort study. *Int. J. Cancer* **2016**, *139*, 1251–1260. [CrossRef] [PubMed]
40. Chiuve, S.E.; Fung, T.T.; Rimm, E.B.; Hu, F.B.; McCullough, M.L.; Wang, M.; Stampfer, M.J.; Willett, W.C. Alternative dietary indices both strongly predict risk of chronic disease. *J. Nutr.* **2012**, *142*, 1009–1018. [CrossRef]
41. Osborne, R.H.; Hopper, J.L.; Kirk, J.A.; Chenevix-Trench, G.; Thorne, H.J.; Sambrook, J.F.; Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer. kConFab: A research resource of Australasian breast cancer families. *Med. J. Aust.* **2000**, *172*, 463–464. [CrossRef] [PubMed]
42. John, E.M.; Hopper, J.L.; Beck, J.C.; Knight, J.A.; Neuhausen, S.L.; Senie, R.T.; Ziogas, A.; Andrulis, I.L.; Anton-Culver, H.; Boyd, N.; et al. The Breast Cancer Family Registry: An infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Res.* **2004**, *6*, R375–R389. [CrossRef]
43. Terry, M.B.; Phillips, K.A.; Daly, M.B.; John, E.M.; Andrulis, I.L.; Buys, S.S.; Goldgar, D.E.; Knight, J.A.; Whittemore, A.S.; Chung, W.K.; et al. Cohort profile: The breast cancer Prospective Family Study Cohort (ProF-SC). *Int. J. Epidemiol.* **2016**, *45*, 683–692. [CrossRef]

44. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]
45. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [CrossRef]
46. Lai, Z.; Markovets, A.; Ahdesmaki, M.; Chapman, B.; Hofmann, O.; McEwen, R.; Johnson, J.; Dougherty, B.; Barrett, J.C.; Dry, J.R. VarDict: A novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* **2016**, *44*, e108. [CrossRef]
47. Nguyen-Dumont, T.; Steen, J.A.; Winship, I.; Park, D.J.; Pope, B.J.; Hammet, F.; Mahmoodi, M.; Tsimiklis, H.; Theys, D.; Clendenning, M.; et al. Mismatch repair gene pathogenic germline variants in a population-based cohort of breast cancer. *Fam. Cancer* **2020**, *19*, 197–202. [CrossRef]
48. Amos, C.I.; Dennis, J.; Wang, Z.; Byun, J.; Schumacher, F.R.; Gayther, S.A.; Casey, G.; Hunter, D.J.; Sellers, T.A.; Gruber, S.B.; et al. The OncoArray Consortium: A network for understanding the genetic architecture of common cancers. *Cancer Epidemiol. Biomark. Prev.* **2017**, *26*, 126–135. [CrossRef]
49. Das, S.; Forer, L.; Schönherr, S.; Sidore, C.; Locke, A.E.; Kwong, A.; Vrieze, S.I.; Chew, E.Y.; Levy, S.; McGue, M.; et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **2016**, *48*, 1284–1287. [CrossRef] [PubMed]
50. Howie, B.N.; Donnelly, P.; Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **2009**, *5*, e1000529. [CrossRef] [PubMed]
51. Clarke, T.-K.; Adams, M.J.; Davies, G.; Howard, D.M.; Hall, L.S.; Padmanabhan, S.; Murray, A.D.; Smith, B.H.; Campbell, A.; Hayward, C.; et al. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N = 112 117). *Mol. Psychiatry* **2017**, *22*, 1376–1384. [CrossRef]
52. Marees, A.T.; De Kluiver, H.; Stringer, S.; Vorspan, F.; Curis, E.; Marie-Claire, C.; Derks, E.M. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *Int. J. Methods Psychiatr. Res.* **2018**, *27*, e1608. [CrossRef]
53. Joo, J.E.; Wong, E.M.; Baglietto, L.; Jung, C.-H.; Tsimiklis, H.; Park, D.J.; Wong, N.C.; English, D.R.; Hopper, J.L.; Severi, G.; et al. The use of DNA from archival dried blood spots with the Infinium Human Methylation 450 array. *BMC Biotechnol.* **2013**, *13*, 23. [CrossRef] [PubMed]
54. Du, P.; Zhang, X.; Huang, C.C.; Jafari, N.; Kibbe, W.A.; Hou, L.; Lin, S.M. Comparison of Beta-value and M-value Methods for quantifying methylation levels by microarray analysis. *BMC Bioinform.* **2010**, *11*, 587. [CrossRef]
55. Van Rooij, J.; Mandaviya, P.R.; Claringbould, A.; Felix, J.F.; Van Dongen, J.; Jansen, R.; Franke, L.; 't Hoen, P.A.C.; Heijmans, B.; Van Meurs, J.B.; et al. Evaluation of commonly used analysis strategies for epigenome- and transcriptome-wide association studies through replication of large-scale population studies. *Genome Biol.* **2019**, *20*, 235-14. [CrossRef] [PubMed]
56. Luijk, R.; Wu, H.; Ward-Caviness, C.K.; Hannon, E.; Carnero-Montoro, E.; Min, J.L.; Mandaviya, P.; Müller-Nurasyid, M.; Mei, H.; van der Maarel, S.M.; et al. Autosomal genetic variation is associated with DNA methylation in regions variably escaping X-chromosome inactivation. *Nat. Commun.* **2018**, *9*, 1–9. [CrossRef] [PubMed]
57. Houseman, E.A.; Accomando, W.P.; Koestler, D.C.; Christensen, B.C.; Marsit, C.J.; Nelson, H.H.; Wiencke, J.K.; Kelsey, K.T. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinform.* **2012**, *13*, 86. [CrossRef] [PubMed]
58. Jaffe, A.E.; Irizarry, R.A. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol.* **2014**, *15*, R31. [CrossRef] [PubMed]
59. Cole, S.R.; Platt, R.W.; Schisterman, E.F.; Chu, H.; Westreich, D.; Richardson, D.; Poole, C. Illustrating bias due to conditioning on a collider. *Int. J. Epidemiol.* **2010**, *39*, 417–420. [CrossRef]
60. Greenland, S. quantifying biases in causal models: Classical confounding vs collider-stratification bias. *Epidemiology* **2003**, *14*, 300–306. [CrossRef]
61. Pearce, N.; Richiardi, L. Commentary: Three worlds collide: Berkson's bias, selection bias and collider bias. *Int. J. Epidemiol.* **2014**, *43*, 521–524. [CrossRef]
62. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. [CrossRef] [PubMed]
63. Yang, J.; Benyamin, B.; McEvoy, B.P.; Gordon, S.D.; Henders, A.K.; Nyholt, D.R.; Madden, P.A.; Heath, A.C.; Martin, N.G.; Montgomery, G.W.; et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **2010**, *42*, 565–569. [CrossRef]
64. Consortium, I.H. Integrating common and rare genetic variation in diverse human populations. *Nature* **2010**, *467*, 52. [CrossRef]
65. Ripke, S.; O'Dushlaine, C.; Chambert, K.; Moran, J.L.; Kähler, A.K.; Akterin, S.; Bergen, S.E.; Collins, A.L.; Crowley, J.J.; Fromer, M.; et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat. Genet.* **2013**, *45*, 1150–1159. [CrossRef] [PubMed]
66. Bulik-Sullivan, B.; Finucane, H.K.; Anttila, V.; Gusev, A.; Day, F.R.; Loh, P.R.; ReproGen Consortium; Psychiatric Genomics Consortium; Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium; Duncan, L.; et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **2015**, *47*, 1236–1241. [CrossRef] [PubMed]