The HABP2 G534E variant is an unlikely cause of familial non-medullary thyroid cancer

Ruta Sahasrabudhe¹, Jacob Stultz¹, John Williamson¹, Paul Lott¹, Ana Estrada², Mabel Bohorquez², Claire Palles³, Guadalupe Polanco-Echeverry¹,⁴, Emma Jaeger³, Lynn Martin³, Maria Magdalena Echeverry², Ian Tomlinson³ on behalf of TCUKIN⁵, Luis G. Carvajal-Carmona¹,²,⁴

¹Genome Center and Department of Biochemistry and Molecular Medicine, School of Medicine, University of California, Davis, USA
²Grupo de Citogenética, Filogenia y Evolución de Poblaciones, Facultades de Ciencias y Facultad de Ciencias de la Salud, Universidad del Tolima, Ibagué, Colombia
³Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom
⁴Fundación de Genómica y Genética Molecular, Colombia
⁵Members of TCUKIN are listed in the Acknowledgments

Abbreviated Title: HABP2 G534E does not cause thyroid cancer

Key terms: HABP2, G534E, rs7080536, non-medullary thyroid cancer

Word count: 1685

Number of figures and tables: 4

Corresponding Author and person to whom reprint requests should be addressed:

Luis G. Carvajal-Carmona, PhD

University of California, Davis

UC Davis Genome Center, 451 Health Sciences Drive

Davis, California 95616, USA

Phone: +1 530-752-9654, Fax: +1 530-754-9658

Email: lgcarvajal@ucdavis.edu

Disclosure statement: The authors have nothing to disclose.
ABSTRACT

Context: A recent study reported the non-synonymous G534E (rs7080536, allele A) variant in the HABP2 gene as causal in familial non-medullary thyroid cancer (NMTC).

Objective: The objective of this study was to evaluate the causality of HABP2 G534E in the TCUKIN study, a multi-center population based study of NMTC cases from the British Isles.

Design and setting: A case-control analysis of rs7080536 genotypes was performed using 2,105 TCUKIN cases and 5,172 UK controls.

Participants: Cases comprised 2,105 NMTC cases. Patients sub-groups with papillary (N=1,056), follicular (N=691) and Hurthle cell (N=86) TC cases were studied separately. Controls comprised 5,172 individuals from the 1958 Birth Cohort (58C) and the National Blood Donor Service (NBS) study. The controls had previously been genotyped using genome-wide SNP arrays by the Wellcome Trust Case Control Consortium study.

Outcome Measures: Association between HABP2 G534E (rs7080536A) and NMTC risk was evaluated using logistic regression.

Results: The frequency of HABP2 G534E was 4.2% in cases and 4.6% in controls. We did not detect an association between this variant and NMTC risk (OR=0.896, 95% CI: 0.746-1.071, P=0.233). We also failed to detect an association between HABP2 G534E and cases with papillary (1056 cases, G534E frequency= 3.5%, OR=0.74, P=0.017), follicular (691 cases, G534E frequency= 4.7%, OR=1.00, P=1.000) or Hurthle cell (86 cases, G534E frequency= 6.3%, OR=1.40, P=0.279) histology.

Conclusions: We found that HABP2 G534E is a low-to-moderate frequency variant in the British Isles and failed to detect an association with NMTC risk, independent of histological type. Hence, our study does not implicate HABP2 G534E or a correlated polymorphism in familial NMTC and additional data are required before using this variant in NMTC risk assessment.

INTRODUCTION
Non-medullary thyroid cancer (NMTC) is the most common endocrine malignancy. As incidence rates are growing at an annual 5% in the U.S. (1), this malignancy will soon become the third most commonly diagnosed cancer among American women and is now the second most common cancer among U.S. Hispanic women (1-3). NMTC is also one of the few cancer types for which a genetic risk is higher than the risk conferred by lifestyle and environmental exposures (4). Genome-wide association (GWA) and candidate studies have identified low penetrance NMTC variants on chromosomes 9q22, 14q13, 2q35, 8p12, 8q24 and 14q13 (5-8) while linkage studies of familial NMTC have suggested that genomic regions on 1q21, 2q21 and 19p13 harbor highly penetrant variants (9-11), although no causal mutations have been identified conclusively in the latter regions.

In a recent study, Gara et al (12) identified a missense variant (G534E or rs7080536 allele A, rs7080536A) in the HABP2 gene, on chromosome 10q25, that showed complete co-segregation in an kindred of seven affected NMTC cases from unknown ethnicity including six with a papillary histology and one with a follicular adenoma. After several functional experiments, which included enhanced colony formation and increased cell migration and reporting a significant frequency difference, at p<0.001 between the data from the thyroid cancer TCGA study and a “multi-ethnic” database, the authors concluded that HABP2 G534E was causal of familial NMTC. The HABP2 gene was identified as a plasma-hyaluronan-binding protein with serine protease activity that plays a role in coagulation and fibrinolysis (13-15). The HABP2 G534E variant is also known as Marburgh I polymorphism, which has shown to reduce HABP2 activity (16) and which has been implicated as a risk factor in cardiovascular diseases (17), in progression of carotid stenosis (18) and in venous thromboembolism in some but not all studies (19,20). Gara et al thus suggested a new role for HABP2 as a familial NMTC gene, a finding that could be of potential great importance for risk assessment and for personalized prevention and treatment of this increasingly common malignancy. Intrigued by the potential importance of HABP2 G534E for NMTC risk, we decided to investigate this variant in a large multi-center population-based study of NMTC in the British Isles (21,22).
MATERIALS AND METHODS

Study Samples

2,105 NMTC cases were recruited through a multi-center Thyroid Cancer genetics UK and Ireland (TCUKIN) study (21,22). All cases had histologically confirmed NMTC and were of northern European ancestry. After completion of a brief questionnaire, cases donated ~10ml of blood for DNA isolation. The Southampton and South West Hampshire Research Ethics Committee approved the TCUKIN research protocol. For the present study, we also used previously published genotype data from 5,172 UK population controls, which included 2,673 participants in the 1958 Birth Cohort and (58C) and 2,499 donors of the National Blood Donor Service (NBS) (21).

Genotyping

All cases were genotyped for HABP2 G534E/rs7080536A using KASP chemistry (LGC genomics, UK) following the manufacturer’s protocol (genotyping probes are shown in Supplemental Table 1). Two of the HABP2 G534E homozygous and two of the heterozygous cases detected by KASP genotyping were verified by Sanger sequencing (Supplemental Figure 1). Call rates for genotyping were >99%; concordance between KASP genotyping and Sanger sequencing was 100% and the visual inspection of the genotype clusters did not reveal obvious technical issue. HABP2 G534E/rs7080536A was in Hardy Weinberg equilibrium in both cases and controls (data not shown).

Statistical analysis

Association testing: Logistic regression methods implemented in PLINK (23) and R (24) were used to obtain association statistics (odd ratios, ORs, and two sided P-values) as reported previously (21,22). We also carried out case-control and case-only analyses between HABP2 G534E and clinical characteristics.
including age of onset and histological subtype (available in 1,833 cases). These analyses were carried out using the chi-square test and R.

Haplotype analysis
In order to investigate the origin of the HABP2 G534E variant, we genotyped four HABP2 G534E homozygous and 18 heterozygous carriers at four closely linked SNPs (rs10787491, rs932650, rs10885478 and rs1885434), which covered a 6kb region (chr10:113,584,808-113,590,838) centered around HABP2 G534E. Haplotype reconstruction was carried out using Haplovie (25).

RESULTS

Association between HABP2 G534E and NMTC risk
We genotyped 2,105 cases from the British Isles for the HABP2 G534E variant and used publically available genotype data from 5,172 population matched controls as reported before (21,22,26-28). Allele counts, odds ratios and two sided allelic P values are shown in Table 1. The population frequency of HABP2 G534E was found to be 4.6%, which suggest that this is a low-to-moderate frequency variation in the general population of the British Isles. As shown in Table 1, HABP2 G534E was more common in controls than in cases and we failed to detect a significant association between this variant and NMTC risk in our study (OR=0.896, 95% CI: 0.746-1.071, P=0.233, Table 1) despite having considerable statistical power to detect variants associated with the expected familial effect of HABP2 G534E (not shown).

Association between HABP2 G534E and histological subtype and age of diagnosis
Histologically, NMTC can be divided into main subtypes: Papillary TC (PTC), the most common subtype which accounts for ~80% of all the cases and follicular TC (FTC), which accounts ~10-20% of cases (29,30) and also includes cases with follicular variant histology (30). Hurthle cell carcinoma is also a rare histological subtype commonly found in NMTC families (31). In our study, we evaluated the association...
between the HABP2 G534E and these three NMTC subtypes. As shown in Table 2, we failed to detect significant associations between HABP2 G534E and increased risk of any of the three NMTC subtypes tested in our study. Unexpectedly, we detected a protective effect of this variant in PTC (P=0.017, OR=0.74, Table 2) which further disagrees with Gara et al findings. Next, we examined if the HABP2 G534E heterozygotes or homozygotes had an earlier age of onset as the expectation for highly penetrant mutations is an anticipation of the disease. We did not detect differences in HABP2 G534E carriers and non-carriers in our study (average age for heterozygous = 48.3 years, G534E homozygous = 46.2 years, non-carriers = 46.7 years, P>0.273 for all comparisons, Table 3). Therefore, the stratification of our study by histological type and age of onset also failed to detect an effect of HABP2 G534E on NMTC risk.

Haplotype analysis

An intriguing possibility that could explain our failure to replicate Gara et al’s findings is that HABP2 G534 could have multiple independent origins and that one of the HABP2 G534 bearing haplotypes could have a second cryptic casual mutation in a second gene in the same region for NMTC. To assess whether HABP2 G534 had single or multiple origins, we evaluated closely linked markers in the region. Table 4 shows the haplotypes reconstructed in the 22 individuals (4 homozygous and 18 heterozygous) with this variant in our study. We found that all HABP2 G534E carriers shared the same 2.5 kb core haplotype defined by three markers (HABP2 G534E/rs7080536A, rs10885478G, rs1885434G, Table 4). This finding is important because it suggests that the G534E allele has a single origin as it only happens in a unique haplotype.

DISCUSSION

NMTC is the most common endocrine malignancy and one of the few cancers where the variance in the risk explained by a genetic predisposition (53%) is more important than the risk explained by lifestyle, the environment or chance (32). Although several regions harboring highly penetrant NMTC variants have
been identified so far, no mutated genes have been conclusively found in such regions (5-11). Only the recent study by Gara et al reported HABP2 G534E as a causal variant in an NMTC family. As these findings can have potentially important implications in NMTC risk assessment and management, we investigated the role of this HABP2 variant in our large population-based study but failed to confirm a causal role of HABP2 G534E in NMTC.

The frequency of HABP2 G534E has been reported to range from 1.6% to 4.3% in European populations (33). Our study confirmed these previous findings as we observed a population frequency of 4.6% in the British Isles controls. Our data are also consistent with a previous British study on cardiovascular disease where HABP2 G534E population frequency was found to be 4.3% (17). We can therefore conclude that HABP2 G534E is a low-to-moderate frequency variant that segregates in populations of European ancestry. This finding may explain why Gara et al reported a frequency of ~4% in cases from the NMTC TCGA study, most of whom were likely to be white/European Americans (30). As a control group, Gara et al used data from a “multi-ethnic database” where HABP2 G534E frequency was ~0.07%. This database was most likely the 1000 Genomes Study, which contains populations of both European and non-European ancestry and where HABP2 G534E frequencies are affected by the very different ancestries in the sample. Thus, the significant P-value reported by Gara et al could have been the result of population stratification rather than that of an association with the disease. The population frequency analyses from our study therefore suggest that HABP2 G534E is a low-to-moderate frequency variant that is highly unlikely to be a highly penetrant, cancer-causing mutation. This conclusion is further supported by our association analyses, where we failed to detect associations between HABP2 G534E and increased NMTC risk using both case-control and cases-only analyses, showing that it is not even a low-penetrance thyroid cancer gene. Although we cannot exclude the possibility that the G534E variant reported as causal in the Gara et al study resides in a very rare haplotype harboring additional cryptic mutations that cause NMTC, our haplotype analyses suggested a single origin of this variant in populations of European ancestry.
ancestry. While further haplotype analyses should also be carried out in Gara et al’s family, our finding of a single haplotype does not support the notion of multiple independent origins of G534E.

In summary, our study suggests that HABP2 G534E is very unlikely to cause familial NMTC. At the population level, we also found that this variant does not increase NMTC risk. We therefore suggest that careful replication studies should be performed and great caution should be taken when assessing NMTC risk among carriers of any HABP2 variant.

ACKNOWLEDGMENTS

We are grateful to all individuals who participated in the TCUKIN study. We are also grateful to National Cancer Research Institute’s Thyroid Cancer Subgroup and National Cancer Research Network for supporting the TCUKIN study. LGCC receives funding from the University of California Davis, The V Foundation for Cancer Research, and The National Institute On Aging (UC Davis Latino Aging Research Resource Center, award number P30AG043097) and The National Cancer Institute (Paul Calabresi Career Development Award for Clinical Oncology K12 at UC Davis, award number K12CA138464) of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. LGCC, ME and IT received funding from the FP7 CHIBCHA Consortium. We thank Wellcome Trust Case-Control Consortium in making control data publicly available. The Wellcome Trust Centre for Human Genetics is funded by the Wellcome Trust (Grant number 075491/Z/04). IT receives funding from Cancer Research UK and the European Commission. AE is supported by Programas Doctorales Becas COLCIENCIAS. AE, ME, MB and LGCC receive support from the Research Office from University of Tolima (Projects 400111 and 204360113).
Collaborators in the TCUKIN study include Laura Moss, Velindre Cancer Centre, Cardiff CF14 2TL, United Kingdom; Christopher Scrase, The Ipswich Hospital, Ipswich, IP4 5PD, UK; Andrew Goodman, Royal Devon & Exeter Hospital, Exeter, EX2 5DW, UK; Radu Mihai, John Radcliffe Hospital, OX3 9DU, UK; James Gildersleve, Royal Berkshire Hospital, Reading, RG1 5AN, UK; Catherine Lemon, Mount Vernon Hospital, Northwood, HA6 2RN, UK; Antony Robinson, Royal United Hospital, Bath, BA1 3NG, UK; Caroline Brammer, Newcross Hospital, Wolverhampton, WV10 0QP, UK; Georgina Gerrard, St. James University Hospital, Leeds. LS9 7TF, UK; Hisham Mehanna, Institute of Head and Neck Studies and Education, University Hospitals of Coventry and Warwickshire, Walsgrave, CV2 2DX, UK; Matthew Beasley, Bristol Hematology and Oncology Centre, Bristol, BS2 158ED, UK; Hosahalli K. Mohan & Sue Clarke, Guy’s Hospital, London, SE1 9RT, UK; Kate Goodchild, Luton & Dunstable Hospital, Luton, LU4 0DZ, UK; Jonathan Wadsley, Weston Park Hospital, Sheffield, S10 2SJ, UK; Abdel Hamid, Scunthorpe General Hospital, Scunthorpe, DN15 7BH, UK; Danielle Power, St. Mary’s Hospital, London, W2 1NY, UK; Elena Macias, Kent and Canterbury Hospital, CT1 3NG, UK; Jerry Sharp, Royal Derby Hospital, Derby, DE22 3NE, UK; Mr. Andrew Coatsworth, York Hospital, York YO31 8HW, UK; Hamish Courtney, Royal Victoria Hospital, Belfast, BT12 6BA, UK; Stephen Whitaker & Katie Wood, Royal Surrey County Hospital, Guildford, GU2 227XX, UK; James McCaul, Bradford Royal Infirmary, Bradford, BD9 6RJ, UK; Christopher Ashford, Worcestershire Royal Hospital, Worcester, WR5 1DD, UK; Tom Roques & Craig Martin, Norfolk and Norwich University Hospital NHS Trust, Norwich, NR4 7UY, UK; Vivienne Loo, Broomfield Hospital, Chelmsford, CM1 7ET, UK; Jennifer Marshall, Southampton General Hospital, Southampton, SO16 266YD, UK; Amy Roy, Derriford Hospital, Plymouth, PL6 8DH, UK; Joanna Simpson, The Royal Sussex County Hospital, Brighton, BN2 5BE, UK; Nick Rowell, Maidstone Hospital, Maidstone, ME16 9QQ, UK; Mr. Edward Babu, Hillingdon Hospital, Uxbridge, UB8 3NN, UK; Narayanan Srihari, Royal Shrewsbury Hospital, Shrewsbury, SY3 8XQ, UK; Mr. Simon Ellenbogen, Tameside General Hospital, Ashton-under-Lyne, OL6 9RW, UK; Paul Ryan, Medway Maritime Hospital, Gillingham, ME7 5NY, UK; Arshad Jamil, University Hospital North Staffs, Stoke on Trent, ST4 6QGU; Terri P McVeigh, National
REFERENCES


variants associated with low TSH levels and thyroid cancer risk. Nature genetics 2012; 44:319-322


23. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics 2007; 81:559-575


Cancer Genome Atlas Research N. Integrated genomic characterization of papillary thyroid carcinoma. Cell 2014; 159:676-690


Hoppe B, Tolou F, Dorner T, Kiesewetter H, Salama A. Gene polymorphisms implicated in influencing susceptibility to venous and arterial thromboembolism: frequency distribution in a healthy German population. Thrombosis and haemostasis 2006; 96:465-470