Screening of Graves' Disease Susceptibility Genes by Whole Exome Sequencing in a Three-generation Family

Zhuoqing Hu

Department of Endocrinology, The Second Affiliated Hospital of Guangzhou Medical University

Wei Li

Huizhou Health Science Polytechnic

Miaosheng Li

Department of Endocrinology, Affiliated Hospital of Guangdong Medical University, Zhanjiang

Hao Wei

Department of Endocrinology, Affiliated Hospital of Guangdong Medical University, Zhanjiang

Zhihui Hu

Department of Endocrinology, Affiliated Hospital of Guangdong Medical University, Zhanjiang

Yanting Chen

Department of Endocrinology, Affiliated Hospital of Guangdong Medical University, Zhanjiang

Ai Luo

Department of Endocrinology, The Second Affiliated Hospital of Guangzhou Medical University

Wangen Li (✉ liwg660@126.com)

Department of Endocrinology, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou 510220, China

<https://orcid.org/0000-0002-6138-7320>

Research article

Keywords: Graves' disease, susceptibility genes, three generations, Whole Exome Sequencing

Posted Date: June 5th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-33213/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Graves' disease(GD) has a tendency for familial aggregation, but it is uncommon to occur in more than two generations. Little is known about susceptibility genes for GD in the three-generation family.

Methods: DNA were extracted from three-generation familial GD patient with a strong genetic background in a Chinese Han population. We utilize the Whole Exome Sequencing(WES) to screen and the Sanger sequencing confirmed potential disease-causing genes.

Results:In this case study, there are five patients with Graves' disease(GD) from a three-generation family. We firstly revealed the SNVs of *MAP7D2*(c. 452C>T: p. A151V), *SLC1A7*(c. 1204C>T: p. R402C), *TRAF3IP3*(c. 209A>T: p. N70I), *PTPRB*(c. 3472A>G: p. S1158G), *PIK3R3*(c. 121C>T: p. P41S), *DISC1*(c. 1591G>C: p. G531R) were associated with the familial GD. The Sanger sequencing confirmed these variations and more importantly is that the PolyPhen-2 score showed that the variants in *TRAF3IP3*, *PTPRB*, *PIK3R3* are more likely to change protein functions.

Conclusion: The *MAP7D2*, *SLC1A7*, *TRAF3IP3*, *PTPRB*, *PIK3R3*, *DISC1* may be the candidate susceptibility genes for familial GD from a three generations family.

Background

Graves' disease is a thyroid-related autoimmune disorder which is caused by complex interaction between susceptibility genes and multiple environmental factors. Previous familial and twin studies showed that Graves' disease is associated with genetic factors in 79% cases[1], which influences the familial clustering in GD. Family linkage analysis, candidate gene method and genome-wide association analysis(GWAS) have identified a greater number of Graves' disease susceptibility loci. However, all the variants that have been discovered still account little for the heritability of GD. Therefore, different approaches were applied in this case study to identify the more susceptibility loci.

In addition to twins group, the familial GD is the ideal object of study on contribution of genetic heritability to complex disease. Findings from family linkage analysis indicated that the 5q31-q33, 6p, 7q, 8q, 10q, 12q, 14q and 20q regions were related to GD susceptibility[2–4]. Subsequent researches verified the linkage of familial GD to *HLA* gene and *CTLA-4* gene. Whereas, linkage analysis has not addressed the need for fine gene mapping in complex diseases, and this method requires a large familial sample size to localize the pathogenic genes by observing whether the genetic markers are co-segregated with the disease. The Whole Exome Sequencing makes it possible to explore susceptibility genes even with a small sample size. Moreover, it is an efficient strategy for fine mapping to determine the exact location of variant[5, 6]. The study is about one three-generation familial GD patient with a strong genetic background in a Chinese Han population, and to utilize WES to screen more disease-causing genes.

Methods

Study participants

A three-generation family from Zhanjiang, Guangdong Province had been targeted. There were 25 people in the family tree including 13 males and 12 females (Fig 1.). Five members were diagnosed with GD(Case 2 had type 2 diabetes simultaneously). Two healthy people in the first-degree relatives of the family were set as control group at the same time. Both the cases and control group were confirmed based upon medical history, physical examination, and results of thyroid function examination. All the five GD patients had typical clinical manifestation of hyperthyroidism such as heart and hands trembling. Four of the patients resorted to drug (Methimazole) for anti-thyroid treatment and one other patient resorted to I^{131} for anti-thyroid treatment. As of the month before the article was submitted, we confirmed that the control group was still free of Graves' disease. Table 1 has the details of the participants.

Thyroid function examination and Susceptibility Gene Screening

Detection of thyroid hormones and TRAb by chemiluminescence method(Cobase411) was used and the reference range of indicators are: FT3: 2.3-6.8pmol/L; FT4: 10-23.5pmol/L; TSH: 0.34~4.0mIU/L; TRAb: 0~1.75IU/L. For preparation of DNA, DNA extraction kit (Tiangen Biochemical Technology Co., Ltd.) was used.

The Agilent SureSelect Human All Exome Kit(Agilent) was used for exon capture. The sequencing processes are exon capture, hybrid library cleaning and purification, PCR amplification of exon DNA library, library quality detection, sequencing, and data analysis. The

exome region was sequenced by illumine hiseq2500, and GATK standard procedure was adopted to calibrate the initial data. Quality control of raw reads was performed with fastqc. Transition and Transversion (SNV) and Insertion and Deletion (InDel) of each sample were detected through VarScan and GATK HaplotypeCaller. Sanger sequencing was used to confirm the genotype variant of 6 genes(*MAP7D2*, *SLC1A7*, *TRAF3IP3*, *PTPRB*, *PIK3R3*, *DISC1*) that were in all the participants.

Results

The quality of raw data

The sequencing quality Q value was used to evaluate the sequencing error rate of the base. The base quality value Q20 indicates that the error rate is 1%. Similarly, the base quality value Q30 indicates that the error rate is 0.1%. In this case study, the data reveals that the sequencing quality of the seven samples are high (table 2).

The base average coverage depth of all samples are larger than 100× which means that the detected SNV is reliable.

SNV/InDel detection and annotation of 7 samples

The SNV/InDel locus that was discovered with both VarScan and GATK methods was to be of high quality and if not discovered by both VarScan and GATK methods, the locus was medium quality. There were 144169 high quality SNV/InDel locus involved in this study(table 3).

The variations carried by the patient were identified as susceptibility genes of GD in the three-generation family, including *MAP7D2*(c. 452C>T: p. A151V), *SLC1A7*(c. 1204C>T: p. R402C), *TRAF3IP3*(c. 209A>T: p. N70I), *PTPRB*(c. 3472A>G: p. S1158G), *PIK3R3*(c. 121C>T: p. P41S), *DISC1*(c. 1591G>C: p. G531R). All SNV/InDel sites in the case study were rare variants according to several known population databases(Freq_Alt1000, Kaviar_20150923, ESP6500, gnomAD). Additionally, Chinese Millionome Database(CMDB, <https://db.cngb.org/cmdb/>) were applied for the comparative analysis. CMDB contain considerable variation and their allele frequency information from 141,431 unrelated healthy Chinese individuals(Phase I results). SNV function analysis indicated that all the SNV of *MAP7D2*, *SLC1A7*, *TRAF3IP3*, *PTPRB*, *PIK3R3*, *DISC1* were nonsynonymous variations. The genotype of the 6 genetic variations from the 7 samples by Sanger sequencing were the same with the results of the whole exome sequencing(Figure 2). The PCR primers of six genetic variants were shown in Table 5. Based on PolyPhen-2 prediction and the amino acids conservation analysis in orthologous species, the rs555004337 in *TRAF3IP3*, rs186466118 in *PTPRB*, and rs115181807 in *PIK3R3* are likely to affect the protein function(Table 4 and figure 3). These genes involved in the Biological Process, Molecular Function, Cellular Component, and KEGG pathway were showed in additional excel file.

Discussion

In this case study, it presents a rare familial GD case in 5 patients in a three-generation family. The five patients are consistent with the general characteristic of GD patients which is that the GD is prone to attack women at the age of 30-60 [7]. All the members of the three-generation family came from the same district of Zhanjiang city and they have been living in similar environment which guarantees the consistency of environmental factors in this case study. Although the etiology of GD is complex and clear identification of potential factors for GD has not been completed, it is widely recognized that the genetic determinants contribute to the risk of GD such as *HLA*, *CTLA-4*, *PTPN22*, and *CD40*[8]. However, no variation of these former identified genes were found, but the following variations of *MAP7D2*, *SLC1A7*, *TRAF3IP3*, *PTPRB*, *PIK3R3*, *DISC1*, and *SUPT20HL* were found in this familial GD. Furthermore, the variations in *PTPRB*, *PIK3R3*, and *TRAF3IP3* were predicted that they may alter the functions of the encoded protein. Familial GD of multigeneration is important for heritable studies because it avoids the genetic heterogeneity factor, therefore, this case study may explain the genetic cause of the familial clustering of GD.

MAP7D2 (MAP7 domain containing 2) was located on X chromosome. *MAP7D2* is specifically expressed in human brain tissue which has impact on the behavioral traits and cognition in human. *MAP7D2* is also associated with sex-biased mental illnesses[9]. Previous studies also show that gender predisposition to GD is associated with X chromosome inactivation (XCI) migration[10]. Thus, the study on *MAP7D2* is likely to provide important general information about the reason why women are more vulnerable to GD. *SLC1A7* and *DISC1* are also susceptibility genes for mental illnesses. In one research, Keith A. Young et al. discovered that *DISC1* gene plays a vital role in post-traumatic stress disorder (PTSD) severity of US military veterans[11]. In another research, Fujita K, et al. revealed that *SLC1A7* gene expression in peripheral blood leukocytes is responsible for the association between socioeconomic status and depressive

mood in healthy adults[12]. We speculated that *SLC1A7* and *DISC1* may work in regulating the symptoms of GD such as nervousness, irritability and irritability.

The protein encoded by *PTPRB* belongs to the family of protein tyrosine phosphatases(PTP). The activation of PTK(protein tyrosine kinase) was regulated through the binding the SH2 domains of PI3K(*PIK3R3* gene encode). The balance of tyrosine protein phosphorylation was regulated by PTP/PTK which participated in cell signal transduction, and regulates cell growth, differentiation, metabolism, transcription, immune responses, etc. Researches have demonstrated that the significant role of PTKs and PTPs is to modulate the tyrosine phosphorylation-dependent signaling pathways which is critical for the effector of NK cell and Neutrophil cell[13,14]. *TRAF3IP3* is highly expressed in CD34+CD38+CD7+ common lymphoid progenitors (CLPs) and on the other hand, CD34+CD38+CD7+ cells have the capacity to differentiate into B/NK/T cell, which implies *TRAF3IP3* possibly plays a role in lymphoid development[15]. The over-activation of the T/B cell was regulated by *CTLA-4* and *CD40* gene variants which has been confirmed in the pathogenesis of autoimmune diseases including GD. Therefore, further study is necessary to figure out whether the variations of *PTPRB*, *PIK3R3* and *TRAF3IP3* are involved in the dysfunction of thyroid autoimmune.

Conclusion

To summarize, the WES was applied to establish an association between *MAP7D2*, *SLC1A7*, *TRAF3IP3*, *PTPRB*, *PIK3R3*, *DISC1* genes and the familial GD of a three-generation family. The findings in this cast studies are clues for further study and more verification and function researches are needed to explore these genes related to GD susceptibility.

Tables

Tab 1. The basic clinical characteristics of the object of study

	Case 1	Case 2	Case 3	Case 4	Case 5	Con 1	Con 2
Year (2015)	death	56	54	52	17	45	21
Sex	female	female	female	female	male	male	female
Year of initial diagnosis	1990	2006	2007	1992	2014	-	-
Smoking	-	-	-	-	-	-	-
Drinking	-	-	-	-	-	-	-
Clinical manifestation							
Heart	+	+	+	+	+	-	-
Proptosis	+	+	-	+	+	-	-
Hands tremble	+	+	+	+	+	-	-
Goiter	-	-	+	+	+	-	-
Serological test							
FT3 (pmol/L)	3.26	4.76	6.86	2.39	9.85	4.49	4.70
FT4 (pmol/L)	16.57	17.60	21.61	7.71	26.50	16.45	16.94
TSH (mUL/L)	0.005	0.118	0.005	62.10	0.072	1.18	0.740
TRAb (IU/L)	-	2.38	4.13	-	5.75	1.27	1.04
Treatment	Thiamazole	Thiamazole	Thiamazole	¹³¹ I	Thiamazole	-	-

+: positive; -: negative

Tab 2. Raw data sequencing data statistics

	R1				R2			
	Total	Reads	Q20	Q30	Total	Reads	Q20	Q30
	Reads	length	(%)	(%)	Reads	length	(%)	(%)
Case1	59415807	35-151	99.98	97.51	59415807	35-151	99.6	92.65
Case2	43894852	35-151	99.98	97.60	43894852	35-151	99.67	93.46
Case3	71415598	35-151	99.97	97.61	71415598	35-151	99.64	93.18
Case4	74526096	35-151	99.98	97.58	74526096	35-151	99.64	93.41
Case5	80885417	35-151	99.98	97.57	80885417	35-151	99.62	93.09
Con1	80289875	35-151	99.98	97.55	80289875	35-151	99.47	92.06
Con2	69913521	35-151	99.98	97.50	69913521	35-151	99.66	93.88

"Q20 (%)" and "Q30 (%)" respectively indicate the ratio that the sequencing quality value Q in the raw data is not lower than Q20 and Q30.

Tab 3. The SNV/InDe locus number statistics in 7 samples

Category	SNV(N)	InDel(N)
Both	125581	18588
Only Varscan	51959	12118
Only GATK	5337	2250

Table 4. The details of SNV/InDel information

Gene	<i>MAP7D2</i>	<i>SLC1A7</i>	<i>TRAF3IP3</i>	<i>PTPRB</i>	<i>PIK3R3</i>	<i>DISC1</i>
Information						
Position	X:20074820	1:53556483	1:209933593	12:70956666	1:46546408	1:231906773
exon	4	8	3	14	2	6
SNP ID	rs750367268	-	rs555004337	rs186466118	rs115181807	rs56229136
Ref allele	G	G	A	T	G	G
Alt allele	A	A	T	C	A	C
Function	nonsynonymous	nonsynonymous	nonsynonymous	nonsynonymous	nonsynonymous	nonsynonymous
PolyPhen-2 score	0.139(benign)	-	0.729(possibly damaging)	0.521(possibly damaging)	1(probably damaging)	0.059(benign)
Freq_Alt (1000)	-	0.000199	0.001797	0.004193	-	0.000599
Kaviar_20150923	0.0000194	0.0000388	0.0006015	0.0015006	-	0.0003428
ESP6500	-	-	-	-	-	0.000077
gnomAD	0.000011	0.000028	0.0006	0.0019	0.000022	0.0004
CMDB	-	-	-	0.0038	0.0051	-
Genotype						
Case 1	G/A	G/A	A/T	T/C	G/A	G/C
Case 2	G/A	G/A	A/T	T/C	G/A	G/C
Case 3	G/A	G/A	A/T	T/C	G/A	G/C
Case 4	G/A	G/A	A/T	T/C	G/A	G/C
Case 5	A/A	G/A	A/T	T/C	G/A	G/C
Con 1	G/G	G/G	A/A	T/T	G/G	G/G
Con 2	G/G	G/G	A/A	T/T	G/G	G/G

CMDB: Chinese Millionome Database; -: No found.

Table 5. The PCR primers of six genetic variants

Gene: variant	primers
<i>MAP7D2</i> : c. 452C>T	F-ATCTGAAAGTGGTGCCTCTGAA R-TAGCCTAGCCGCATTGTTTACT
<i>SLC1A7</i> : c. 1204C>T	F-GTTGACCTGGGCGATGAAGA R-AAACACCTCCCTCATAGGAAGAAC
<i>TRAF3IP3</i> : c. 209A>T	F-AACAGGTGCTTGGAGGTCATC R-AGCACACAGCAGTATGTCCCTT
<i>PTPRB</i> : c. 3472A>G	F-GGAAACTAAGGACCAACCAAGG R-CACTGCATTTCCTCCCTCA
<i>PIK3R3</i> : c. 121C>T	F-GCATTCTAGTTACCTTGAAATATCC R-CATACCTTGGTTAGTGAGCTGCT
<i>DISC1</i> : c. 1591G>C	F-GGAAATAGAGGAGCAAGAGCAG R-CAGACTGCTTGGGAAATGTTTAG

Abbreviations

GD: Graves' disease; WE: Whole Exome Sequencing

Declarations

Ethics approval and consent to participate

Informed written consent was obtained from all participants. In this study, no cases was under the age of 16. Ethical approval was provided by the the Ethics Committee of the Affiliated Hospital of Guangdong Medical University.

Consent for publication

Written informed consent for publication of the case was obtained for each participant. In our study, one case was under 18 and the written informed consent for publication was obtained from his parents.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Our research was financially supported by grants from China Postdoctoral Science Foundation Grant(Grant NO. 2019 M652859). The funding bodies only provided the financial means to buy reagents for experiments, and the funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Author contribution

Designed the study: ZQH

Analyzed data: WL, MSL, HW

Collected data: YTC, ZHH

Manuscript editing: WL

Wrote the paper: ZQH, WGL

Supervision: WGL, AL

All authors read and approved the final manuscript.

Acknowledgements

Not Applicable.

References

1. Brix, TH.; Kyvik, KO.; Christensen, K.; Hegedüs, L. Evidence for a major role of heredity in Graves' disease: a population-based study of two Danish twin cohorts. *J Clin Endocrinol Metab.* 2001; 86: 930-934.
2. Sakai, K.; Shirasawa, S.; Ishikawa, N.; Ito, K.; Tamai, H.; Kuma, K.; Akamizu, T.; Tanimura, M.; Furugaki, K.; Yamamoto, K.; Sasazuki, T. Identification of susceptibility loci for autoimmune thyroid disease to 5q31-q33 and Hashimoto's thyroiditis to 8q23- q24 by multipoint affected sib-pair linkage analysis in Japanese. *Hum Mol Genet.* 2001; 10: 1379-1386.
3. Jin, Y.; Teng, W.; Ben, S.; Xiong, X.; Zhang, J.; Xu, S.; Shugart, YY.; Jin, L.; Chen, J.; Huang, W. Genome-wide scan of Graves' disease: evidence for linkage on chromosome 5q31 in Chinese Han pedigrees. *J Clin Endocrinol Metab.* 2003; 88: 1798-1803.
4. Tomer, Y.; Ban, Y.; Concepcion, E.; Barbesino, G.; Villanueva, R.; Greenberg, DA.; Davies, TF. Common and unique susceptibility loci in Graves and Hashimoto diseases: results of whole-genome screening in a data set of 102 multiplex families. *Am J Hum Genet.* 2003; 73: 736-747.
5. Glessner, JT.; Bick, AG.; Ito, K.; Homsy, J.; Rodriguez-Murillo, L.; Fromer, M.; Mazaika, E.; Vardarajan, B.; Italia, M.; Leipzig, J.; DePalma, SR.; Golhar, R.; Sanders, SJ.; Yamrom, B.; Ronemus, M.; Iossifov, I.; Willsey, AJ.; State, MW.; Kaltman, JR.; White, PS.; Shen, Y.; Warburton, D.; Brueckner, M.; Seidman, C.; Goldmuntz, E.; Gelb, BD.; Lifton, R.; Seidman, J.; Hakonarson, H.; Chung, WK. Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of nucleotide variation array and exome sequence data. *Circulation Research.* 2014; 115: 884-896.
6. Brastians, PK.; Amaro, T W.; Manley, PE.; Jones, RT.; Dias-Santagata, D.; Thorner, AR.; Lawrence, MS.; Rodriguez, FJ.; Bernardo, LA.; Schubert, L.; Sunkavalli, A.; Shillingford, N.; Calicchio, ML.; Lidov, HG.; Taha, H.; Martinez-Lage, M.; Santi, M.; Storm, PB.; Lee, JY.; Palmer, JN.; Adappa, ND.; Scott, RM.; Dunn, IF.; Laws, ER Jr.; Stewart, C.; Ligon, KL.; Hoang, MP.; Van, Hummelen P.; Hahn, WC.; Louis, DN.; Resnick, AC.; Kieran, MW.; Getz, G.; Santagata, S. Exome sequencing identifies BRAF variations in papillary craniopharyngiomas. *Nature Genetics.* 2014; 46: 161-65.
7. Hemminki, K.; Li, X.; Sundquist, J.; undquist, K.; The epidemiology of Graves' disease: evidence of a genetic and an environmental contribution. *J Autoimmun.* 2010; 34: 307-313.
8. Tomer, Y. Mechanisms of autoimmune thyroid diseases: from genetics to epigenetics. *Annu Rev Pathol.* 2014; 9: 147-156.
9. Davies, W. Genomic imprinting on the X chromosome: implications for brain and behavioral phenotypes. *Ann N Y Acad Sci.* 2010; 1204 Suppl: E14-9.
10. Simmonds, MJ.; Kavvoura, FK.; Brand, OJ.; Newby, PR.; Jackson, LE.; Hargreaves, CE.; Franklyn, JA.; Gough, SC. Skewed X chromosome inactivation and female preponderance in autoimmune thyroid disease: an association study and meta-analysis. *J Clin Endocrinol Metab.* 2014; 99: E127-131.
11. Young, KA.; Morissette, SB.; Jamroz, R.; Meyer, EC.; Stanford, MS.; Wan, L.; Kimbrel, NA. 5-HTTLPR and *DISC1* risk genotypes for elevated PTSD symptoms in US military veterans. *World Psychiatry.* 2017; 16: 109-110.
12. Fujita, K.; Kuwano, Y.; Saijo, S. Negative perception of socioeconomic status with depressive mood down-regulates expression of PPBP and *SLC1A7* genes in peripheral blood leukocytes. *Cogent Psychology.* 2017; 4.
13. Vivier, E.; Nunès, JA.; Vély, F. Natural killer cell signaling pathways. *Science.* 2004; 306: 1517-1519.
14. Kruger, J.; Butler, JR.; Cherapanov, V.; Dong, Q.; Ginzberg, H.; Govindarajan, A.; Grinstein, S.; Siminovitch, KA.; Downey, GP. Deficiency of Src homology 2-containing phosphatase 1 results in abnormalities in murine neutrophil function: studies in moth-eaten mice. *J Immunol.* 2000; 165(10): 5847-5859.
15. Hoebeke, I.; De Smedt, M.; Stolz, F.; Pike-Overzet, K.; Staal, FJ.; Plum, J.; Leclercq, G. T-, B- and NK-lymphoid, but not myeloid cells arise from human CD34(+)CD38(-)CD7(+) common lymphoid progenitors expressing lymphoid-specific genes. *Leukemia.* 2007; 21: 311-319.

Figures

Figure 1
Pedgree of study family

Figure 2

The Sanger sequencing of variations for the 7 participants

Figure 3

Amino acids conservation in orthologous species for the variants.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement4.xlsx](#)