

Whole-exome sequencing reveals genetic variants in *ERC1* and *KCNQ4* associated with complete hydatidiform mole in Chinese Han women

Yan Yu^{1,*}, Bingjian Lu^{2,*}, Weiguo Lu³, Shuang Li⁴, Xiuqin Li⁵, Xinyu Wang³, Xiaoyun Wan³, Yaxia Chen³, Suwen Feng³, Yao Jia⁴, Ru Yang⁴, Fangxu Tang⁴, Xiong Li⁴, Shulan Zhang⁵, Xinyan Wang⁵, Heng Wei⁵, Zhilan Peng⁶, Lin Lu⁶, Huizhen Zhong⁷, Linjun Zhao⁷, Zhangqian Huang⁸, Lin Lin⁹, Weihong Shen⁹, Yan Lu^{3,10}, Zhu Cao¹, Jian Zou¹, Yuejiang Ma¹, Xiaojing Chen¹, Qifang Tian³, Shiming Lu¹¹, Pengyuan Liu^{10,12}, Ding Ma⁴, Xing Xie³ and Xiaodong Cheng³

¹Key Laboratory of Women's Reproductive Health of Zhejiang Province, Women's Hospital School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China

²Department of Surgical Pathology, Women's Hospital School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China

³Department of Gynecologic Oncology, Women's Hospital School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China

⁴Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

⁵Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang, Liaoning, China

⁶Department of Obstetrics and Gynecology, West China Second Hospital of Sichuan University, Chengdu, Sichuan, China

⁷Department of Obstetrics and Gynecology, Ningbo Women and Children's Hospital, Ningbo, Zhejiang, China

⁸Department of Obstetrics and Gynecology, Shaoxing Women and Children Hospital, Shaoxing, Zhejiang, China

⁹Zhejiang University Hospital, Zhejiang University, Hangzhou, Zhejiang, China

¹⁰Institute for Translational Medicine School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China

¹¹Department of Clinical Laboratory, Women's Hospital School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China

¹²Sir Run Run Shaw Hospital School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China

*These authors have contributed equally to this work

Correspondence to: Xiaodong Cheng, **email:** chengxd@zju.edu.cn

Xing Xie, **email:** xiex@zju.edu.cn

Ding Ma, **email:** dma@tjh.tjmu.edu.cn

Pengyuan Liu, **email:** pylu@zju.edu.cn

Keywords: genomics, whole-exome sequencing, complete hydatidiform mole, pathogenesis

Received: March 31, 2017

Accepted: July 29, 2017

Published: September 08, 2017

Copyright: Yu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Complete hydatidiform mole (CHM) is a rare pregnancy-related disease with invasive potential. The genetics underlying the sporadic form of CHM have not been addressed previously, but maternal genetic variants may be involved in biparental CHM. We performed whole-exome sequencing of 51 patients with CHM and 47 healthy women to identify genetic variants associated with CHM. In addition, candidate variants were analyzed using single base extension and Matrix Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry in 199 CHM patients and 400 healthy controls. We validated candidate variants using Sanger sequencing in 250 cases and 652 controls, including 205 new controls. Two single nucleotide polymorphisms, c.G48C(p.Q16H) in *ERC1* and c.G1114A(p.G372S) in *KCNQ4*, were

associated with an increased risk of CHM ($p < 0.05$). These variants may contribute to the pathogenesis of CHM and could be used to screen pregnant women for this genetic abnormality.

INTRODUCTION

Complete hydatidiform mole (CHM) is a rare pregnancy-associated disease that can spread to distant sites [1]. Women with CHM have a 1,000-fold increased risk of gestational trophoblastic neoplasia (clinically aggressive lesions consisting of choriocarcinoma, placental site trophoblastic tumor, and epithelioid trophoblastic tumor), compared to women who have had a term pregnancy [2]. Hydatidiform mole exhibits an unbalanced geographic distribution [3-7]. The incidence is highest in Southeast Asia (3.89–13.99 per 1,000 pregnancies) and lowest in Latin America, North America, Europe, and Oceania (0.23–1.21 per 1,000 pregnancies). The incidence of CHM in Chinese women is approximately five per 1,000 pregnancies [6].

Cytogenetic and molecular pathology data indicate CHM has an androgenetic origin [8, 9]. Lack of maternal genomic imprinting also plays an important role in CHM [10]. Causal mutations in the maternal genes *NLRP7* (Nacht Domain-, Leucine-rich Repeat-, and PYD-containing protein 7) [11-13] and *KHDC3L/C6orf221* (KH domain containing 3 like) have been identified in women with familial, recurrent bi-parental CHM, which accounts for approximately 20% of all CHM [11]. Women with a history of CHM have a 5–40 times higher relative risk of recurrent CHM than other women, regardless of changes in sexual partner. Thus, maternal genetic rather than environmental factors may play a predominant role in CHM.

Next-generation sequencing is a fast and cost-effective method for generating genome-scale sequencing data and has contributed to the search for disease- and trait-related genetic variants [14, 15]. Whole-exome sequencing allows the discovery of low-frequency variants in individuals with familial, highly penetrant diseases, and those with complex quantitative traits [16-19]. We performed whole-exome sequencing to identify genetic variants that contribute to non-familial CHM in Chinese Han women. Candidate variants were further analyzed by single base extension (SBE) and Matrix Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) [20]. An overview of the experimental workflow is shown in Figure 1.

RESULTS

Whole-exome sequencing analysis

The initial screening set included 98 samples (51 CHM and 47 controls). On average, we generated

72,135,990 high-quality reads per sample to a mean depth of 67-fold exon coverage (Supplementary Table 2). We identified 398,594 candidate variants (Figure 2). Most were single nucleotide variants (SNVs). Single nucleotide substitutions or inversions were the predominant types of non-synonymous variants. The major types of nucleotide alterations were C>T and G>A transitions. These SNVs were first filtered using the 1000 genome project with a minor allele frequency (MAF) less than 0.02. After filtering, these rare variants were subject to annotation using the NCBI RefGene database (hg19). A total of 43,219 SNVs were non-silent mutations in coding sequence (CDS) regions (Figure 3A). We then used multiple software tools, including SIFT, PolyPhen, LRT, MutationTaster, MutationAssessor, FATHMM, MetaSVM, and MetaLR, to predict whether the non-silent variants affected protein function [21]. Deleterious variants were defined as those predicted to affect protein function by at least five of the above tools. As a result, 5,301 rare variants were predicted to be deleterious in 2,033 single-variant genes and 1,054 multiple-variant genes (Figure 3B). Candidate SNVs are shown in Supplementary Table 3.

Mass spectrometry analysis

Fisher's exact tests were used to test the association between the aforementioned 2,033 single-variant genes and CHM risk. The top 41 genes containing 41 variants were chosen for further validation. AssotesteR analysis was used to evaluate the association between the aforementioned 1,054 multiple-variant genes and CHM risk. The top 46 genes containing 127 variants were chosen for further validation. The top 168 rare SNVs were chosen from the 5,301 deleterious rare variants for MALDI-TOF mass spectrometry analysis. We hypothesized that these rare SNVs were the most likely to be associated with CHM. Four SNVs were excluded due to the difficulty of primer design or failures due to simultaneous detection of other loci. Thus, 164 SNVs were detected in 6 wells in the mass spectrometry analysis of samples from 199 new CHM patients and 400 new normal controls (Supplementary Table 4). Using mass spectrometry analysis, the minor alleles of 41 SNVs were not detected in either CHM cases or healthy controls. All sites except two were in Hardy-Weinberg equilibrium. Therefore, 121 SNVs were retained for subsequent association analysis. Two variants, *ERC1*c.G48C (p.Q16H) ($p = 0.013$) and *KCNG4* c.G1114A (p.G372S) ($p = 0.014$), were associated with an increased risk of CHM. Several other SNVs in *SDK2*, *KIAA1462*, *ZNF799*, *Cllorf21*, and *BBS9* were also associated with CHM risk (Supplementary Table 5).

Identification of candidate variants associated with CHM risk

We sequenced two sites, *ERC1* c.G48C(p.Q16H) and *KCNG4* c.G1114A(p.G372S) in 902 samples (250 CHM and 652 controls, including 205 new controls). Three CHM cases were excluded because the genomic DNA was not available at this stage. In this validation set, 98 subjects (51 CHM cases and 47 controls) were sequenced using both exome sequencing and Sanger methods. The SNV results of these two sites showed 100% concordance between the two sequencing methods. Similarly, the comparison between mass spectrometry and direct sequencing showed 99.8% concordance between these two sites. High concordance was also observed among the three methods ($R^2=0.9999$) (Supplementary Table 6). *ERC1* c.G48C(p.Q16H) was associated with an increased risk of CHM in all samples ($p<0.01$, odds ratio [OR]=2.69, Table 1), and there were 20 patients have this variant. *KCNG4* c.G1114A(p.G372S) was also associated

with an increased risk of CHM ($p<0.01$, OR=2.68, Table 2), and there were 13 patients with this variant, including one patient homozygous for the variant.

DISCUSSION

We identified genetic variants associated with CHM risk in a Chinese Han population using whole-exome sequencing. Variants were validated using mass spectrometry and Sanger sequencing. Previous studies have demonstrated that exome sequencing enables the discovery of rare variants associated with various cancers, heritable diseases, and complex quantitative traits [16-19]. We have demonstrated that whole-exome sequencing can be used to screen patients for rare disease-associated germline variants.

We identified two new variants, *ERC1* c.G48C(p.Q16H) and *KCNG4* c.G1114A(p.G372S), that were associated with risk of CHM in Chinese women. *ERC1* is a RIM-binding protein. RIMs are active zone proteins

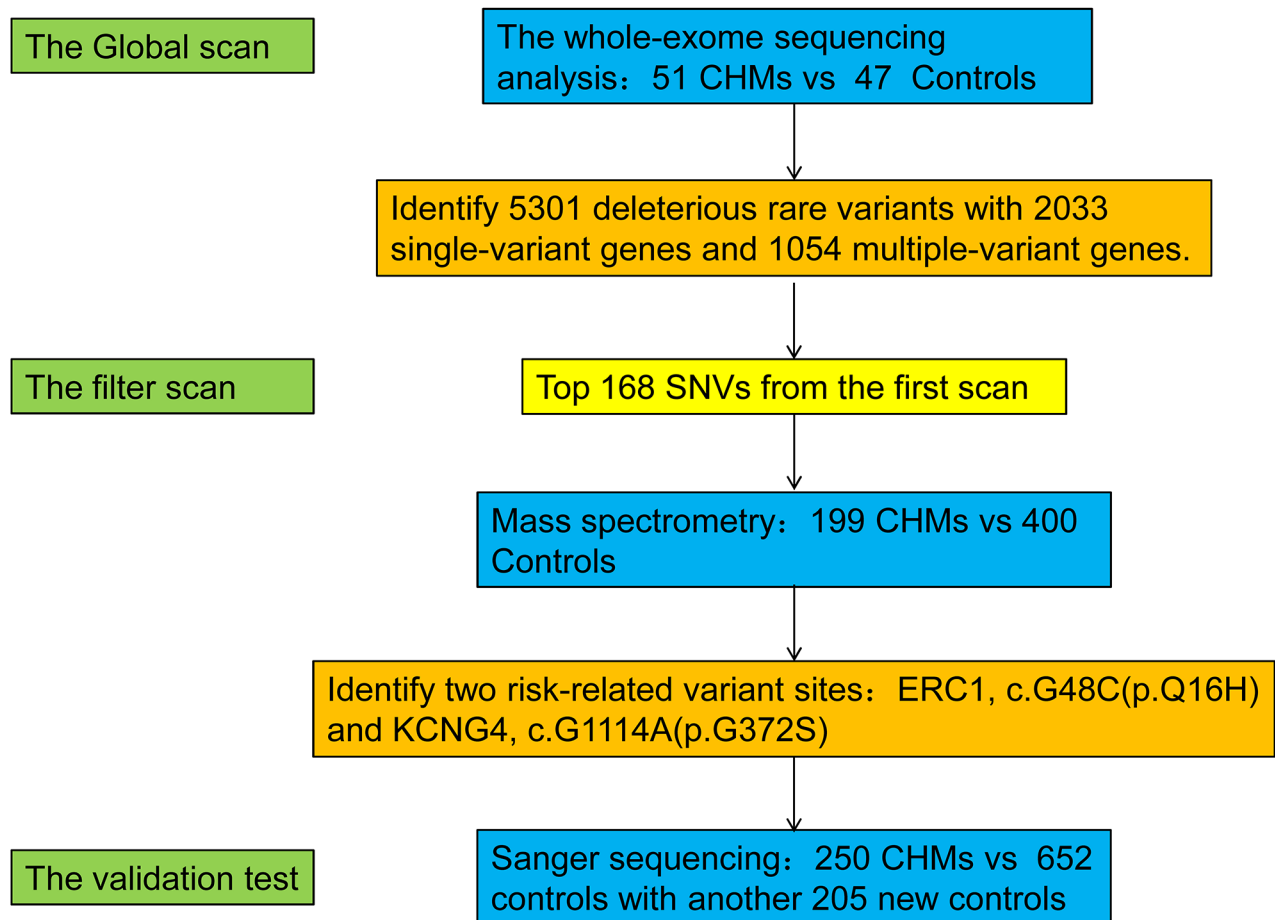


Figure 1: Summary of the study design and results. We performed whole-exome sequencing on samples from 51 CHM patients and 47 healthy women. Screening of candidate mutations by MALDI-TOF MS was performed in 199 CHM patients and 400 healthy women. Candidate polymorphisms were validated by Sanger sequencing in 250 cases and 652 controls, which included samples analyzed in the first and second screening and an additional 205 controls. We identified two SNPs: *ERC1* c.G48C(p.Q16H), and *KCNG4* c.G1114A(p.G372S) that were associated with an increased risk of CHM ($p<0.05$).

that regulate presynaptic neurotransmitter release. Immunocytochemical analysis and live cell imaging showed that YFP-conjugated ELKS translocated to the plasma membrane after antigen stimulation [22]. ELKS,

an essential regulatory subunit of the IKK complex, recruits Ikkappa-B-alpha to the IKK complex and regulates IKK activation [23]. ATM is also exported in a NEMO-dependent manner to the cytoplasm, where it promotes

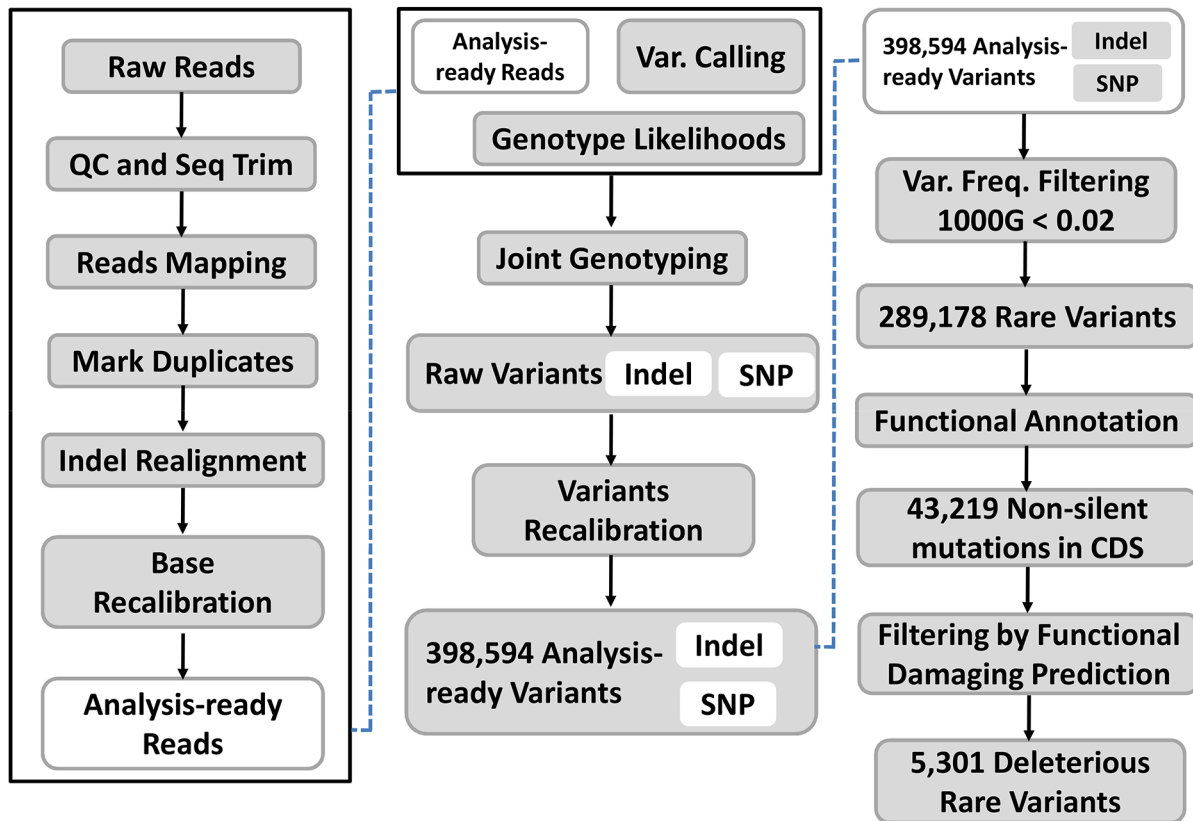


Figure 2: Analysis of whole-exome sequencing data. We identified 398,594 SNVs. There were 289,178 rare SNVs in 2,034 genes (MAF<0.02, in the 1000 genomes project (<http://1000genomes.org>)) and 43,219 non-silent mutations in CDS regions. Functional analysis indicated there were 5,301 rare variants predicted to be deleterious. MAF for each SNV was calculated according to the guideline from https://www.ncbi.nlm.nih.gov/projects/SNP/docs/rs_attributes.html.

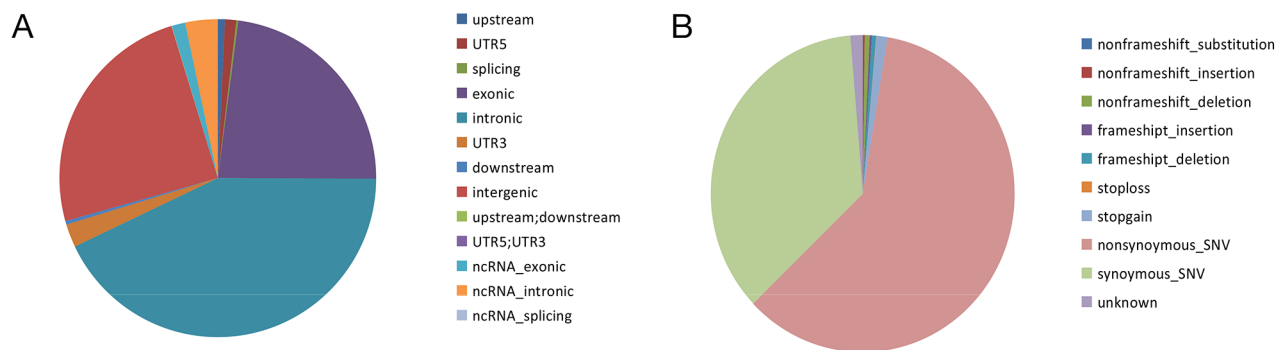


Figure 3: Distribution and functional analysis of rare SNVs. (A) Of the 289,178 rare SNVs identified, there were 2,035 SNVs in upstream regions; 3,367 SNVs in 5'UTRs; 560 SNVs in splice sites; 66,536 SNVs in exons; 123,938 SNVs in introns; 6,776 SNVs in 3'UTR; 1,085 SNVs in downstream regions; 71,210 SNVs in intergenic regions; 117 SNVs in upstream/downstream regions; 6 SNVs in 5'UTR/3'UTRs; 3,923 SNVs in ncRNA exons; 9,591 SNVs in ncRNA introns; 34 SNVs in ncRNA splice sites. (B) Functional analysis of the 43,219 SNVs in CDS regions revealed 7 SNVs that resulted in non-frameshift substitutions; 127 SNVs that resulted in non-frameshift insertions; 356 SNVs that resulted in non-frameshift deletions; 142 SNVs that resulted in frameshift insertions; 306 SNVs that resulted in frameshift deletions; 26 stop-loss SNVs; 764 stop-gain SNVs; 40,064 non-synonymous SNVs; 23,877 synonymous SNVs; and 867 SNVs of unknown consequence.

Table 1: Analysis of the association between *ERC1*(c.G48C) and CHM

ERC1(c.G48C)	CHM patients		Controls		p	OR (95% CI)
	n*	RAF/ALL	n*	RAF/ALL		
WES	51(51)	5/102	47(47)	2/94	0.26	
MassARRAY	198(199)	15/396	399(400)	12/798	0.013	2.58 (1.26 - +∞)
Sanger test	248(250)	20/496	651 (652)	20/1302	0.002	2.69(1.51 - +∞)

RAF=risk allele frequency; n=number.

* The number given in bracket represents for the total of samples for testing while that outside of the bracket indicates the total of the samples with interpretable results. Several cases failed in these detection systems.

Table 2: Analysis of the association between *KCNG4*(c.G48C) and CHM

KCNG4(c.G48C)	CHM patients		Controls		p	OR (95% CI)
	n*	RAF/ALL	n*	RAF/ALL		
WES	51(51)	3/102	47(47)	0/94	0.14	
MassARRAY	197(199)	11/394	395(400)	7/790	0.014	3.21(1.31 - +∞)
Sanger test	247(250)	14/494	650 (652)	14/1300	0.009	2.68(1.33 - +∞)

RAF=risk allele frequency; n=number.

* The number given in bracket represents for the total of samples for testing while that outside of the bracket indicates the total of the samples with interpretable results. Several cases failed in these detection systems.

IKK activation in an ELKS-dependent manner [24]. ATM- and NEMO-dependent ubiquitination of ELKS results in ubiquitin-dependent assembly of the TAK1/TAB2/3 and NEMO/IKK complexes, and IKK and NF- κ B activation in response to genotoxic stimuli [25]. SDCCAG8 was found to interact with centriolar satellite proteins (OFD1 and AZI1), members of the endosomal sorting complex (RABEP2 and ERC1), and with non-muscle myosin motor proteins (MYH9, MYH10, and MYH14) at the centrosome [26]. Mutations in ERC1 likely impact centrosome function. ERC1-RET fusions resulting from a translocation, t(10;12)(q11;p13), have also been observed in thyroid papillary carcinoma [27, 28].

Voltage-gated potassium (Kv) channels are the most structurally and functionally complex class of voltage-gated ion channels. These channels have multiple functions including regulation of neurotransmitter release, heart rate, insulin secretion, neuronal excitability, epithelial electrolyte transport, smooth muscle contraction, and cell volume. *KCNG4* encodes a member of the potassium channel, voltage-gated, subfamily G and functions as a modulatory subunit. Multiple alternatively spliced variants have been observed in normal and malignant tissue. Sequence analysis indicated Kv6.3 was a previously uncharacterized member of the Kv6 subfamily. The other splice variants were the first members of two unique subfamilies, Kv10.1 and Kv11.1. These channels did not produce K⁺ currents when expressed in mammalian cells, although they have all of the hallmarks of voltage-gated

K⁺ channel subunits. Kv6.3, Kv10.1, and Kv11.1 alone did not localize to the plasma membrane, but were retained in the endoplasmic reticulum [29]. Muller et al. found that Dlk1 suppressed Notch signaling and induced expression of the K(+) channel subunit *KCNG4* to modulate delayed rectifier currents [30]. This gene may affect the oocyte during the second meiotic division.

Sporadic CHM with androgenetic origin results from fertilization of an empty oocyte with paternal genes. But the precise mechanism underlying the empty oocyte is currently unclear. According to our results, *ERC1* and *KCNG4* are highly conserved genes associated with centrosome function. ERC1 interacts with SDCCAG8, a major component of the centriole. *KCNG4* may impact centrosome function by modulating the K⁺ current. We hypothesize that *ERC1*c.G48C(p.Q16H) and *KCNG4* c.G1114A(p.G372S) may play a role in the generation of an empty oocyte during the second meiotic division by altering centrosome function. However, there are no reliable methods to investigate empty oocytes *in vitro* or *in vivo*. This is partly because CHM have not been observed in animal models. However, a few partial hydatidiform moles have been described in cats and Friesian cows [31, 32].

We identified two single nucleotide polymorphisms(SNPs)in *ERC1*and *KCNG4* that were associated with an increased risk of CHM. Women with these two variants should be carefully monitored. A pathological examination of the product of pregnancy in

women with an inevitable abortion should be performed. Additional studies with larger sample sizes and long-term follow-up are necessary. Functional studies of ERC1 and KCNG4 in animal models of CHM are also required in order to clarify the pathogenesis of CHM.

MATERIALS AND METHODS

Study population and sample preparation

We analyzed 250 Chinese Han women with non-familial CHM (16–56 years old) and 652 age-matched healthy controls (21–45 years old) who were treated at several tertiary centers across China between November 2011 and July 2015. The mean age \pm standard deviation (SD) of the CHM patients and controls was 31.6 \pm 9.3 and 29.1 \pm 4.6 years, respectively. All subjects provided written consent. Clinical data was collected by chart review and self-reporting. CHM was histologically confirmed by a gynecological pathologist (B.L.) according to the consensus diagnostic criteria (WHO 2014). Women with familial, recurrent CHM were excluded from the study by chart review and follow-up visit. Healthy controls were defined as Han women who had at least one previous normal delivery of a healthy baby, no history of an abortion (including spontaneous and artificial) or abnormal pregnancy, and no family history of cancer or gestational trophoblastic disease. Each subject provided approximately 15mL of peripheral blood, which was stored in EDTA-anticoagulant tubes at -80°C until use. Genomic DNA was extracted using the Qiagen QIAamp DNA Blood Mini Kit according to the manufacturer's instructions. Genomic DNA was stored in Tris-EDTA buffer at -80°C until use.

Whole-exome sequencing and data analysis

Whole-exome sequencing was performed on genomic DNA from 51 patients and 47 controls (the initial screening group). Whole-exome capture libraries were constructed from approximately 3 μ g of purified genomic DNA following fragmentation, end repair, phosphorylation, and ligation. Ligated samples were hybridized to the exon capture arrays using Roche NimbleGen V2 (44.1 Mbp). Captured DNA was annealed at 95°C and single-stranded DNA amplified and sequenced on an Illumina HiSeq2000.

Exome sequencing data were analyzed using the Exome Analysis Pipeline (DNAnexus, <http://www.dnanexus.com>). Variant detection and genotyping were performed on exomes and the flanking 500 bp downstream of the 5'- or 3'-UTRs. A BAM (binary alignment map) file for each sample was generated from the alignment of the Illumina sequence reads to the human genome (hg19). SNV calls were evaluated and each variant annotated.

Mass spectrometry

We selected 168 rare SNPs that were identified by whole-exome sequencing and were predicted to be deleterious. We used the Sequenom MassARRAY technology platform with the iPLEX GOLD chemistry [33, 34] as a second-round filter. An additional 199 patients and 400 controls were tested. Genotyping analysis was performed using the commercial iPLEX Gold SNP genotyping kit and the MassARRAY platform according to the manufacturer's protocols. Mass determination was done with the MALDI-TOF mass spectrometer. The MassARRAY Typer 4.0 software was used for data acquisition. SNP genotypes were called after cluster analysis using the default setting. Genotype calls were reviewed manually to correct potential clustering artifacts. Assays with less than a 90% call rate within the same SpectroCHIP were excluded from further analysis.

Validation of selected candidate variant sites by Sanger sequencing

Two SNPs were selected for further validation by Sanger sequencing in 250 cases and 652 controls, including an additional 205 controls. Both non-synonymous SNPs were believed to be significant based on mass spectrometry. The primer sequences are shown in Supplementary Table 1. Genomic DNA was amplified by PCR and sequenced using an ABI 3730XL DNA analyzer.

Statistical analysis

Statistical analysis was performed using SPSS 20.0 (SPSS, Inc., Chicago, IL, USA). Fisher's exact tests and Chi-square tests were used to evaluate the associations between CHM and single-variant genes. The ORs and 95% confidence intervals (CIs) were calculated to estimate the relative risk of CHM. AssotesteR was used to detect the association between CHM and multi-variant genes. AssotesteR is a statistical package for R that is commonly used in genetic association studies of rare variants and binary (dichotomous) traits (<https://cran.r-project.org/web/packages/AssotesteR/index.html>). The threshold for statistical significance was one-sided p-values of 0.05, assuming that rare alleles increase the risk of CHM.

ACKNOWLEDGMENTS

This work was supported by an international cooperation project from the Zhejiang Science and Technology Bureau (2012C24029), a Science and Technology grant from the Zhejiang Financial Bureau (N20110531), a Science and Technology grant from the Zhejiang Financial Bureau (LY16H160010), and Projects in the National Science & Technology Pillar Program (2015BAI 13B05). We thank the following physicians

for assisting with sample and clinical data collection: (1) Yue Hu and Wenbing Chen, Department of Obstetrics and Gynecology, The 2nd Affiliated Hospital, Wenzhou Medical University, Wenzhou, Zhejiang, China; (2) Zhiming Hao, Department of Obstetrics and Gynecology, The Affiliated Hospital, School of Medicine, Ningbo University, Ningbo, Zhejiang, China; and (3) Xiaofeng Zhao, Zhejiang Provincial People's Hospital, Hangzhou, Zhejiang, China.

CONFLICTS OF INTEREST

All authors have reviewed the manuscript and read the conflicts of interest notification. The authors declare that there are no conflicts of interest.

FUNDING

This work was supported by an international cooperation project from the Zhejiang Science and Technology Bureau (2012C24029), Science and Technology grants from the Zhejiang Financial Bureau (N20110531 and LY16H160010), and Projects in the National Science & Technology Pillar Program(2015BAI13B05).

REFERENCES

1. Seckl MJ, Sebire NJ, Berkowitz RS. Gestational trophoblastic disease. *Lancet*. 2010; 376: 717-29. <https://doi.org/10.1016/S0140-673660280-2>.
2. Bagshawe KD. Recent observations related to the chemotherapy and immunology of gestational choriocarcinoma. *Adv Cancer Res*. 1973; 18: 231-63.
3. Altieri A, Franceschi S, Ferlay J, Smith J, La Vecchia C. Epidemiology and aetiology of gestational trophoblastic diseases. *Lancet Oncol*. 2003; 4: 670-8.
4. Bagshawe KD, Dent J, Webb J. Hydatidiform mole in England and Wales 1973-83. *Lancet*. 1986; 2: 673-7.
5. Hayashi K, Bracken MB, Freeman DH Jr, Hellenbrand K. Hydatidiform mole in the United States (1970-1977): a statistical and theoretical analysis. *Am J Epidemiol*. 1982; 115: 67-77.
6. Song HZ, Wu PC. Hydatidiform mole in China: a preliminary survey of incidence on more than three million women. *Bull World Health Organ*. 1987; 65: 507-11.
7. Takeuchi S. Incidence of gestational trophoblastic disease by regional registration in Japan. *Hum Reprod*. 1987; 2: 729-34.
8. Slim R, Mehio A. The genetics of hydatidiform moles: new lights on an ancient disease. *Clin Genet*. 2007; 71: 25-34. <https://doi.org/10.1111/j.1399-0004.2006.00697.x>.
9. Vejerslev LO, Dissing J, Hansen HE, Poulsen H. Hydatidiform mole: genetic origin in polyploid conceptuses. *Hum Genet*. 1987; 76: 11-9.
10. Hui P, Buza N, Murphy KM, Ronnett BM. Hydatidiform moles: genetic basis and precision diagnosis. *Annu Rev Pathol*. 2017; 12: 449-85. <https://doi.org/10.1146/annurev-pathol-052016-100237>.
11. Fallahian M, Sebire NJ, Savage PM, Seckl MJ, Fisher RA. Mutations in NLRP7 and KHDC3L confer a complete hydatidiform mole phenotype on digynic triploid conceptions. *Hum Mutat*. 2013; 34: 301-8. <https://doi.org/10.1002/humu.22228>.
12. Murdoch S, Djuric U, Mazhar B, Seoud M, Khan R, Kuick R, Bagga R, Kircheisen R, Ao A, Ratti B, Hanash S, Rouleau GA, Slim R. Mutations in NALP7 cause recurrent hydatidiform moles and reproductive wastage in humans. *Nat Genet*. 2006; 38: 300-2. <https://doi.org/10.1038/ng1740>.
13. Qian J, Deveault C, Bagga R, Xie X, Slim R. Women heterozygous for NALP7/NLRP7 mutations are at risk for reproductive wastage: report of two novel mutations. *Hum Mutat*. 2007; 28: 741. <https://doi.org/10.1002/humu.9498>.
14. Poduri A, Evrony GD, Cai X, Walsh CA. Somatic mutation, genomic variation, and neurological disease. *Science*. 2013; 341: 1237758. <https://doi.org/10.1126/science.1237758>.
15. Watson IR, Takahashi K, Futreal PA, Chin L. Emerging patterns of somatic mutations in cancer. *Nat Rev Genet*. 2013; 14: 703-18. <https://doi.org/10.1038/nrg3539>.
16. Docherty LE, Rezwan FI, Poole RL, Turner CL, Kivuva E, Maher ER, Smithson SF, Hamilton-Shield JP, Patalan M, Gizewska M, Peregud-Pogorzelski J, Beygo J, Buiting K, et al. Mutations in NLRP5 are associated with reproductive wastage and multilocus imprinting disorders in humans. *Nat Commun*. 2015; 6: 8086. <https://doi.org/10.1038/ncomms9086>.
17. Li Y, Vinckenbosch N, Tian G, Huerta-Sanchez E, Jiang T, Jiang H, Albrechtsen A, Andersen G, Cao H, Korneliussen T, Grarup N, Guo Y, Hellman I, et al. Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. *Nat Genet*. 2010; 42: 969-72. <https://doi.org/10.1038/ng.680>.
18. Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng H, Liu T, He W, et al. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*. 2010; 329: 75-8. <https://doi.org/10.1126/science.1190371>.
19. Zang ZJ, Cutcutache I, Poon SL, Zhang SL, McPherson JR, Tao J, Rajasegaran V, Heng HL, Deng N, Gan A, Lim KH, Ong CK, Huang D, et al. Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. *Nat Genet*. 2012; 44: 570-4. <https://doi.org/10.1038/ng.2246>.
20. Johansen P, Andersen JD, Borsting C, Morling N. Evaluation of the iPLEX(R) Sample ID Plus Panel designed for the Sequenom MassARRAY(R) system. A SNP typing assay developed for human identification and sample tracking based on the SNPforID panel. *Forensic Sci Int Genet*. 2013; 7: 482-7. <https://doi.org/10.1016/j.fsigen.2013.04.009>.

21. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc.* 2015; 10: 1556-66. <https://doi.org/10.1038/nprot.2015.105>.
22. Nomura H, Ohtsuka T, Tadokoro S, Tanaka M, Hirashima N. Involvement of ELKS, an active zone protein, in exocytotic release from RBL-2H3 cells. *Cell Immunol.* 2009; 258: 204-11. <https://doi.org/10.1016/j.cellimm.2009.05.005>.
23. Ducut Sigala JL, Bottero V, Young DB, Shevchenko A, Mercurio F, Verma IM. Activation of transcription factor NF-kappaB requires ELKS, an IkappaB kinase regulatory subunit. *Science.* 2004; 304: 1963-7. <https://doi.org/10.1126/science.1098387>.
24. Wu ZH, Shi Y, Tibbetts RS, Miyamoto S. Molecular linkage between the kinase ATM and NF-kappaB signaling in response to genotoxic stimuli. *Science.* 2006; 311: 1141-6. <https://doi.org/10.1126/science.1121513>.
25. Wu ZH, Wong ET, Shi Y, Niu J, Chen Z, Miyamoto S, Tergaonkar V. ATM- and NEMO-dependent ELKS ubiquitination coordinates TAK1-mediated IKK activation in response to genotoxic stress. *Mol Cell.* 2010; 40: 75-86. <https://doi.org/10.1016/j.molcel.2010.09.010>.
26. Airik R, Schueler M, Airik M, Cho J, Ulanowicz KA, Porath JD, Hurd TW, Bekker-Jensen S, Schroder JM, Andersen JS, Hildebrandt F. SDCCAG8 interacts with RAB effector proteins RABEP2 and ERC1 and is required for Hedgehog signaling. *PLoS One.* 2016; 11: e0156081. <https://doi.org/10.1371/journal.pone.0156081>.
27. Nakata T, Kitamura Y, Shimizu K, Tanaka S, Fujimori M, Yokoyama S, Ito K, Emi M. Fusion of a novel gene, ELKS, to RET due to translocation t(10;12)(q11;p13) in a papillary thyroid carcinoma. *Genes Chromosomes Cancer.* 1999; 25: 97-103.
28. Nakata T, Yokota T, Emi M, Minami S. Differential expression of multiple isoforms of the ELKS mRNAs involved in a papillary thyroid carcinoma. *Genes Chromosomes Cancer.* 2002; 35: 30-7. <https://doi.org/10.1002/gcc.10095>.
29. Ottschytch N, Raes A, Van Hoorick D, Snyders DJ. Obligatory heterotetramerization of three previously uncharacterized Kv channel alpha-subunits identified in the human genome. *Proc Natl Acad Sci U S A.* 2002; 99: 7986-91. <https://doi.org/10.1073/pnas.122617999>.
30. Muller D, Cherukuri P, Henningfeld K, Poh CH, Wittler L, Grote P, Schluter O, Schmidt J, Laborda J, Bauer SR, Brownstone RM, Marquardt T. Dlk1 promotes a fast motor neuron biophysical signature required for peak force execution. *Science.* 2014; 343: 1264-6. <https://doi.org/10.1126/science.1246448>.
31. Jurka P, Sacharczuk M, Kawka M, Sobczak-Filipiak M, Golicz M. Partial hydatidiform mole diagnosis in a cat: a case report. *J Feline Med Surg.* 2014; 16: 180-2. <https://doi.org/10.1177/1098612X13496247>.
32. Morris FJ, Kerr SM, Laven RA, Collett MG. Large hydatidiform mole: an unusual finding in a calving cow. *N Z Vet J.* 2008; 56: 243-6. <https://doi.org/10.1080/00480169.2008.36841>.
33. Buetow KH, Edmonson M, MacDonald R, Clifford R, Yip P, Kelley J, Little DP, Strausberg R, Koester H, Cantor CR, Braun A. High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci U S A.* 2001; 98: 581-4. <https://doi.org/10.1073/pnas.021506298>.
34. Hu Z, Wu C, Shi Y, Guo H, Zhao X, Yin Z, Yang L, Dai J, Hu L, Tan W, Li Z, Deng Q, Wang J, et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. *Nat Genet.* 2011; 43: 792-6. <https://doi.org/10.1038/ng.875>.