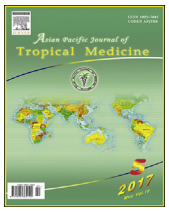
HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Medicine

journal homepage: http://ees.elsevier.com/apjtm



Original research http://dx.doi.org/10.1016/j.apjtm.2017.05.011

Genetic diversity and phylogenetic analysis of EG95 sequences of *Echinococcus granulosus*: Implications for EG95 vaccine application

Wei Pan[1,#], De-Sheng Chen[1,2,#], Yun-Juan Lu[1,2], Hui-Wen Xu[1,2], Wen-Ting Hao[1], Ya-Wen Zhang[1,2], Su-Ping Qin[1], Kui-Yang Zheng[1✉], Ren-Xian Tang[1✉]

[1]*Jiangsu Key Laboratory of Immunity and Metabolism, Department of Pathogenic Biology and Immunology, Laboratory of Infection and Immunity, Xuzhou Medical University, Xuzhou, Jiangsu Province, 221004, PR China*

[2]*Department of Clinical Medicine, Xuzhou Medical University, Xuzhou, Jiangsu Province, 221004, PR China*

ARTICLE INFO

Article history:
Received 6 Mar 2017
Received in revised form 3 Apr 2017
Accepted 1 May 2017
Available online 18 May 2017

Keywords:
Echinococcus granulosus
Cystic echinococcosis
EG95
Vaccine
Genetic diversity

ABSTRACT

Objective: To analyse the genetic variability of EG95 sequences and provide guidance for EG95 vaccine application against *Echinococcus granulosus* (*E. granulosus*).

Methods: We analysed EG95 polymorphism by collecting total 97 different *E. granulosus* isolates from 12 different host species that originated from 10 different countries. Multiple sequence alignments and the homology were performed by Lasergene 1 (DNASTAR Inc., Madison, WI), and the phylogenetic analysis was performed by using MEGA5.1 (CEMI, Tempe, AZ, USA). In addition, linear and conformational epitopes were analysed, including secondary structure, NXT/S glycosylation, fibronectin type III (FnIII) domain and glycosylphosphatidylinositol anchor signal (GPI-anchor). The secondary structure was predicted by PSIPRED method.

Results: Our results indicated that most isolates overall shared 72.6–100% identity in EG95 gene sequence with the published standard EG95 sequence, X90928. However, EG95 gene indeed has polymorphism in different isolates. Phylogenetic analysis showed that different isolates could be divided into three subgroups. Subgroup 1 contained 87 isolates while Subgroup 2 and Subgroup 3 consisted of 3 and 7 isolates, respectively. Four sequences cloned from oncosphere shared a high identity with the parental sequence of the current vaccine, X90928, and they belonged to Subgroup 1. However, in comparison to X90928, several amino acid mutations occurred in most isolates besides oncosphere, which potentially altered the immunodominant linear epitopes, glycosylation sites and secondary structures in EG95 genes. All these variations might change their previous antigenicity and thereby affecting the efficacy of current EG95 vaccine.

Conclusions: This study reveals the genetic variability of EG95 sequences in different *E. granulosus* isolates, and proposed that more vaccination trials would be needed to test the effectiveness of current EG95 vaccine against distinct isolates in different countries.

1. Introduction

The larval stages of *Echinococcus granulosus* (*E. granulosus*) cause cystic echinococcosis (CE) in animals and humans. The parasite has a worldwide distribution [1,2], which is responsible for tremendous economic losses and medical problems. A large amount of variability has long been recognized in *E. granulosus* with several genotypes and designated as G1–G10 [3], and G1 genotype is the most prevalent and causes most human infection [4]. At present, many actions have been taken to treat this disease, of which, surgery in combination with chemotherapy is the first choice. However, this therapy inevitably entails surgical risk and requires considerable labour, material, and financial resource,

First author: Wei Pan, Jiangsu Key Laboratory of Immunity and Metabolism, Department of Pathogenic Biology and Immunology, Laboratory of Infection and Immunity, Xuzhou Medical University, Xuzhou, Jiangsu Province 221004, PR China.
E-mail: panwei525@126.com
✉Corresponding authors: Kui-Yang Zheng, Jiangsu Key Laboratory of Immunity and Metabolism, Department of Pathogenic Biology and Immunology, Laboratory of Infection and Immunity, Xuzhou Medical University, Xuzhou, Jiangsu Province, 221004, PR China. E-mail: ZKY02@163.com
Ren-Xian Tang, Jiangsu Key Laboratory of Immunity and Metabolism, Department of Pathogenic Biology and Immunology, Laboratory of Infection and Immunity, Xuzhou Medical University, Xuzhou, Jiangsu Province, 221004, PR China.
E-mail: Tangrenxian-t@163.com
Peer review under responsibility of Hainan Medical University.
[#] These authors contributed equally to this work.

which limits the practical application. Therefore, it is still urgent to establish novel strategies to control the disease.

Current molecular vaccine presents an alternative and desirable approach. To date, the protective effects of diverse molecules have been investigated on intermediate hosts (sheep, cattle and mice) including AgB [5], EgP29 [6], Eg95 [7] and on definitive hosts (dogs) including EgM and EgA31. Among them, EG95 vaccine, firstly developed against G1 isolates, has shown remarkable success. The vaccine produced high levels of protection (96–100%) in sheep against infection and its efficacy has been demonstrated in independent trials carried out in several countries [7,8]. Moreover, the acquired immunity could be transferred from vaccinated pregnant cattle to their new-born young. These good performances therefore attracted the interest in assessing if the current EG95 vaccine could be a universal vaccine against distinct isolates with varied genotypes other than G1.

It is now well-known that *E. granulosus* exhibits extensive isolate variations that may impact the epidemiology, pathology, and control of CE [9–11], with important implications also for vaccine design. Variability of EG95 gene in different genotypes may directly impact the effectiveness of EG95-based vaccine, which has been revealed in genotype G6 [12]. The situation might also occur in other isolates. In that way, it could limit the practice use of the current EG95 vaccine in the control and prevention of CE for various *E. granulosus* strains. However, little related information is available nowadays. Therefore, fully characterizing the polymorphism of EG95 genes in different *E. granulosus* isolates is desperately needed.

This study collected and analysed 97 EG95 sequences originated from 12 different species and isolated from 10 different countries. The results implied that the isolates could be divided into 3 subgroups. Some genetic changes among these isolates might well influence their antigenicities and thereby cast doubts on the generality of the efficacy of current EG95 vaccine for all the different *E. granulosus* strains. Overall, this study revealed the genetic variability of EG95 genes, which might provide guidance for the application of current EG95 vaccine.

2. Material and methods

2.1. The sources of isolates

The 97 EG95 sequences were collected from *E. granulosus* isolates deposited in GenBank.

2.2. Homology alignment and phylogenetic analysis

Multiple sequence alignments and the homology were performed using the sequence analysis software Lasergene 1 (DNASTAR Inc., Madison, WI). The phylogenetic analysis was performed by the distance-based neighbor-joining method using MEGA5.1 (CEMI, Tempe, AZ, USA). Bootstrap values were calculated on 1000 replicates of the alignment.

2.3. Analysis of linear and conformational epitopes

The linear epitopes of EG95 genes in different isolates were analysed according to Woollard *et al.* [13], which contained four immunodominant regions (T^{29}ETPLRKHFNLTPV^{42}, S^{65}LKAVNPSDPLVYKRQTAKF^{84}, D ^{119}IETPRAGKKESTVMTSGSA^{138}, S^{137}ALTSAIAGFVFSC^{150}).

The characterisitics of conformational epitopes were described based on the analysis of secondary structure, NXT/S glycosylation, fibronectin type III (FnIII) domain and glycosylphosphatidylinositol anchor signal (GPI-anchor) [13]. The secondary structure was predicted by PSIPRED method [14], to compare α-helixs and β-sheets. The prediction server was available at http://bioinf.cs.ucl.ac.uk/psipred/.

3. Results

3.1. The profiles of clinical isolates

Ninety-seven EG95 sequences were analysed. Of these sequences, 97 were isolated from 12 species of hosts including cattle (29), sheep (27), dog (10), Homo sapiens (8), buffalo (6), caprine (5), camel (5), bovine (2), ovine (2), dromedary (1), canine (1) and pig (1). Moreover, they originated from 10 countries including Indian (23), Brazil (15), China (14), Algeria (12), Australia (9), Ethiopia (9), Iran (7), Romania (4), Spain (2) and Argentina (2). Overall, those sequences represented the most prevalent isolates of *E. granulosus* in the world.

3.2. All isolates were relatively conservative

X90928, the parental sequence of EG95 vaccine that cloned from oncosphere, was used as a reference. Those isolates were 72.6–100.0% identical to X90928. In comparison to X90928, the other 4 isolates from oncospheres, including AF503596, AF503597, AF503598 and AY421719, shared 97.7–100% identity while the isolates with G6 (JQ285934, JQ285935, JQ285937, JQ285938, JQ285939) and G7 (AH013644, EU595906) only shared 88.5–99.4% similarity. Overall, the results revealed that the EG95 genes in distinct stages of life cylce of *E. granulosus* isolates were relatively conservative.

3.3. These isolates could be grouped into three subgroups

All isolates were divided into three subgroups according to the phylogenetic analysis. Among the 97 isolates, eighty-seven belonged to Subgroup 1, which consisted of 4 isolates from oncosphere (AF503596, AF503597, AF503598 and AY421719), the current vaccine EG95 (X90928), 2 G6 isolates (JQ285934, JQ285935), 2 G7 isolates (AH013644, EU595906), and others most from G1. This subgroup represented the major branch. In Subgroup 2, it contained 3 G6 isolaltes (JQ285937, JQ285938, JQ285939) while 7 isolates (JF357600, JF829212, AF503599, AF503602, AF503603, AF199347, AF199350) in Subgroup 3 with unknown genotype. The phylogenetic analysis showed the existence of genetic diversity of EG95 genes in different isolates.

3.4. Four linear epitopes were possibly affected by amino acid mutations

The amino acid replacements were observed in the four immunodominant epitopes in EG95 protein identified by Woollard *et al.* [13]. Compared to X90928, the 4 isolates from oncosphere were highly conservative in these epitopes, with

only one mutation in peptide 24 (V^{138} in AF503596). However, a few of isolates exhibited several mutations. Among the test samples, 23, 22, 24, 15 isolates showed the alterations of amino acids in peptide 6, peptide 12/13, peptide 21/22 and peptide 24, respectively. Notably, the insertion mutation of 7 amino acids in the Subgroup 3 altered the previous peptide 21/22, leading to the formation of a new epitope. All those mutations might affect the antigenicity of the four linear epitopes in EG95.

3.5. One additional glycosylation site appeared in some isolates

X90928 and the other 4 isolates from oncosphere were found to have two glycosylation sites at positions 38, 70. The glycosylation site at position 70 was conserved in all analysed isolates. However, other glycosylation sites were changed in the isolates besides oncosphere. Most of isolates in Subgroup 1 and 2 had a glycosylation site at position 38, while the replacements (N to S) occured at this position in Subgroup 3, leading to the disappearance of the glycosylation site. Moreover, 23 isolates in the study were found to have an additional glycosylation site at position 63. The changes might alter the biological function of EG95 in different isolates.

3.6. The secondary structures were partly altered

To compare the potential structural changes after mutation, we performed secondary structure predictions. The reference X90928 protein was predicted to have 2 α-helixs and 11 β-sheets. Among the 97 isolates, 38 shared the similar secondary structure with X90928. However, lots of isolates were exceptional due to variability. Firstly, 22 isolates had 1 α-helixs and 9 β-sheets, and 11 isolates had 2 α-helixs and 12 β-sheets. Moreover, 11 isolates had 2 α-helixs and 10 β-sheets, and 4 isolates had 2 α-helixs and 9 β-sheets. In addition, 3 isolates (HM345596, HM345604, HM345605) contained 7 α-helixs and 1 β-sheets that significantly distinguished from X90928.

In summary, the alternation of amino acids, glycosylation site and secondary structure of EG95, could combinedly affect the function of FnIII domain [13], which has been thought to play an essential role in the protective responses induced by EG95.

4. Discussion

EG95 protein is a 16.6 kDa protein from *E. granulosus* oncosphere that was originally cloned by Lightowlers *et al.* in 1996. EG95-based vaccine was presented as a very promising vaccine against *E. granulosus* isolates and has high protection rate in intermediate hosts [7]. However, the vaccine is not widely distributed worldwide. One potential problem is the uncertainty of this vaccine against distinct isolates in different countries. The polymorphism of EG95 gene remains unclear.

This study found that most of the EG95 genes shared a high similarity except for some isolates with the insertion of 7 amino acids. The result was consistent to the previous studies in G1 isolates, both of which revealed that EG95 genes in different isolates were relatively conservative. Remarkably, 4 sequences derived from oncospheres shared 97.7–100% identity with X90928, which indicated the high conservation of EG95 genes in oncosphere. However, whether EG95 gene is conservative in all *E. granulosus* isolates still needs further investigation.

It is now well recognized that protection effectiveness induced by the EG95 vaccine is strongly associated with humoral immune responses [15]. High levels of EG95-specific IgG antibody are elicited following vaccination and these antibodies are able to effectively kill the parasite in *in vitro* culture [7]. Thus this study mainly analysed the B cell epitope-related characterizations in EG95 genes in different isolates, to further investigate the potential influence of genetic diversity of EG95 on the humoral immune responses mediated by B cells.

Previous finding has suggested that the host protective-epitopes of EG95 vaccine might be linear [15] and 4 immunodominant peptides were then identified [13] to bind the IgG antibodies obtained from the sera of sheep immunized by EG95-GST. Some amino acid mutations were found in the 4 regions of the isolates, which might change the linear epitopes. Actually, the EG95 protein from G6 was found to be unable to bind all the antibodies raised by sheep vaccinated with the current EG95 vaccine [16], which potentially led to the less effectiveness of the vaccine. In this way, the mutated antigen epitopes of isolates in the present study possibly could not be neutralized effectively by the antibodies raised by the current EG95 vaccine, thereby affecting the vaccine efficacy.

However, currently available evidence indicated that the host-protective epitopes of EG95 were conformational and might be related to the FnIII domain [13], which consists of 90–100 amino acids and two β-sheets, since that the specific antibodies elicited by the 4 peptides did not kill the parasite *in vitro* culture assays, nor did the peptides induce protection against challenge infection [15]. This study therefore focused on the sequence characteristics that could potentially affected the conformational epitopes of EG95 proteins in different isolates.

Firstly, several amino acid mutations occurred in the FnIII domain of EG95 genes. The FnIII region appeared to be important for the induction of protective responses since those recombinant truncated forms of EG95 with incomplete FnIII did not induce significant protection against infection. Actually, alternations in the conformation of the domain and its consequences have been previously documented in other literature. For example, new cryptic antigenic sites can be exposed due to a mechanical perturbation or proteolysis [17]. Also, variations in amino acid composition in the strands comprising β-sheets and loops or in specific amino acid residues have demonstrated to result in the change of the conformation in the FnIII domains [18]. It is likely that the differences in the predicted amino acid composition of the FnIII domains in different *E. granulosus* isolates would lead to the conformational change that could be responsible for the emergence of different antigenic epitopes. In addition, the glycosylation sites (N-X-S/T) of EG95 sequences may play an significant role in biological function, which may be changed by some slight changes in glycosylation sites [19]. These differences may compromise the efficacy of vaccine in other isolates than X90928.

In conclusion, this study analysed the 97 EG95 sequences of *E. granulosus*, and found that amino acid substitutions in the genes potentially altered several linear epitopes, glycosylation sites and secondary structures, although that they were relatively conservative in homology. These changes might alter the conformational epitopes of EG95 proteins such that antigenic cross-reactivity with the EG95 vaccine would be reduced or eliminated. This might allow parasites escaping the anti-EG95 immune responses by expressing this protein, which made the

parasite insusceptible to the immune responses induced by the current EG95 vaccine. Although the EG95 gene family has been characterized in G1 and G6 strains [12,16], the validity of their results might be limited due to small number of EG95 isolates. In contrast, this study covered all published EG95 sequences in GenBank and the results might be more comprehensive. This study suggested that more vaccination trials would be needed to test the protective effectiveness of the current EG95 vaccine against distinct isolates in different countries.

Author's contributions

KYZ and RXT designed the research. YJL, HWX, WTH, YWZ and SPQ collected the materials. WP, DSC, KYZ and RXT performed the sequence analysis and wrote the manuscript. KYZ and RXT revised the manuscript critically. All authors read and approved the final manuscript.

Funding

This work was supported by the grants from the National Natural Science Foundation of China (No. 81501762), the Talents Scientific Research Foundation of Xuzhou Medical University (No. D2015004), the Top-notch Academic Programs Project of Jiangsu Higher Education Institutions (PPZY2015B161), the Training Programs of Innovation and Entrepreneurship for College Students in Jiangsu Province (No. 201510313017Z), the Jiangsu Planned Projects for Postdoctoral Research Funds (No. 1501061A), the China Postdoctoral Science Foundation funded project (No. 2015M581864), and the Jiangsu Qing Lan Project. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgements

We thank Professor Yu-gang Wang for revising the manuscript.

References

[1] Song X, Hu D, Yan M, Wang Y, Wang N, Gu X, et al. Molecular characteristics and serodiagnostic potential of dihydrofolate reductase from *Echinococcus granulosus*. *Sci Rep* 2017; **7**(1): 514.

[2] Chebli H, Laamrani El Idrissi A, Benazzouz M, Lmimouni BE, Nhammi H, Elabandouni M, et al. Human cystic echinococcosis in Morocco: ultrasound screening in the Mid Atlas through an Italian-Moroccan partnership. *PLoS Negl Trop Dis* 2017; **11**(3). e0005384.

[3] Maldonado LL, Assis J, Araújo FM, Salim AC, Macchiaroli N, Cucher M, et al. The *Echinococcus canadensis* (G7) genome: a key knowledge of parasitic platyhelminth human diseases. *BMC Genomics* 2017; **18**(1): 204.

[4] Laurimäe T, Kinkar L, Andresiuk V, Haag KL, Ponce-Gordo F, Acosta-Jamett G, et al. Genetic diversity and phylogeography of highly zoonotic *Echinococcus granulosus* genotype G1 in the Americas (Argentina, Brazil, Chile and Mexico) based on 8279bp of mtDNA. *Infect Genet Evol* 2016; **45**: 290-296.

[5] Folle AM, Kitano ES, Lima A, Gil M, Cucher M, Mourglia-Ettlin G, et al. Characterisation of antigen B protein species present in the hydatid cyst fluid of *Echinococcus canadensis* G7 genotype. *PLoS Negl Trop Dis* 2017; **11**(1). e0005250.

[6] Wang H, Li Z, Gao F, Zhao J, Zhu M, He X, et al. Immunoprotection of recombinant Eg. P29 against *Echinococcus granulosus* in sheep. *Vet Res Commun* 2016; **40**(2): 73-79.

[7] Huang F, Dang Z, Suzuki Y, Horiuchi T, Yagi K, Kouguchi H, et al. Analysis on gene expression profile in oncospheres and early stage metacestodes from *Echinococcus multilocularis*. *PLoS Negl Trop Dis* 2016; **10**(4). e0004634.

[8] Larrieu E, Mujica G, Gauci CG, Vizcaychipi K, Seleiman M, Herrero E, et al. Pilot field trial of the EG95 vaccine against ovine cystic echinococcosis in Rio Negro, Argentina: second study of impact. *PLoS Negl Trop Dis* 2015; **9**(10). e0004134.

[9] Ehsan M, Akhter N, Bhutto B, Arijo A, Ali Gadahi J. Prevalence and genotypic characterization of bovine *Echinococcus granulosus* isolates by using cytochrome oxidase 1 (Co1) gene in Hyderabad, Pakistan. *Vet Parasitol* 2017; http://dx.doi.org/10.1016/j.vetpar.2017.04.006.

[10] Kinkar L, Laurimäe T, Simsek S, Balkaya I, Casulli A, Manfredi MT, et al. High-resolution phylogeography of zoonotic tapeworm *Echinococcus granulosus* sensu stricto genotype G1 with an emphasis on its distribution in Turkey, Italy and Spain. *Parasitology* 2016; **143**(13): 1790-1801.

[11] Spotin A, Mahami-Oskouei M, Harandi MF, Baratchian M, Bordbar A, Ahmadpour E, et al. Genetic variability of *Echinococcus granulosus* complex in various geographical populations of Iran inferred by mitochondrial DNA sequences. *Acta Trop* 2017; **165**: 10-16.

[12] Alvarez Rojas CA, Gauci CG, Nolan MJ, Harandi MF, Lightowlers MW. Characterization of the EG95 gene family in the G6 genotype of *Echinococcus granulosus*. *Mol Biochem Parasitol* 2012; **183**(2): 115-121.

[13] Haag KL, Gottstein B, Ayala FJ. The EG95 antigen of *Echinococcus* spp. contains positively selected amino acids, which may influence host specificity and vaccine efficacy. *PLoS One* 2009; **4**(4). e5362.

[14] Chakraborty S, Sharma K, Mukherjee J, Gupta MN, Goyal A. Structural modelling, substrate binding and stability studies of endopectate lyase (PL1B) of family 1 polysaccharide lyase from *Clostridium thermocellum*. *Protein Pept Lett* 2015; **22**(6): 557-568.

[15] Woollard DJ, Heath DD, Lightowlers MW. Assessment of protective immune responses against hydatid disease in sheep by immunization with synthetic peptide antigens. *Parasitology* 2000; **121**(Pt2): 145-153.

[16] Alvarez Rojas CA, Gauci CG, Lightowlers MW. Antigenic differences between the EG95-related proteins from *Echinococcus granulosus* G1 and G6 genotypes: implications for vaccination. *Parasite Immunol* 2013; **35**(2): 99-102.

[17] Vogel V, Sheetz M. Local force and geometry sensing regulate cell functions. *Nat Rev Mol Cell Biol* 2006; **7**(4): 265-275.

[18] Karuri NW, Lin Z, Rye HS, Schwarzbauer JE. Probing the conformation of the fibronectin III1–2 domain by fluorescence resonance energy transfer. *J Biol Chem* 2009; **284**(6): 3445-3452.

[19] Ruiz-Blanco YB, Paz W, Green J, Marrero-Ponce Y. ProtDCal: a program to compute general-purpose-numerical descriptors for sequences and 3D-structures of proteins. *BMC Bioinforma* 2015; **16**: 162.