Journal of Animal Research: v.3 n.2 p.223-230. December, 2013

Identification and *In-silico* Annotation of Functional Single Nucleotide Polymorphisms (SNPs) of the Candidate Gene Association with the Canine Transmissible Venereal Tumor Disease

Neelesh Sharma[#], Mrinmoy Ghosh[#], Simrinder Singh Sodhi, Jong-hyun Kim, Nameun Kim, Raj Kumar Mongre, Pradeep Adhikari and Dong Kee Jeong*

Laboratory of Animal Genetic Engineering and Stem Cell Biology
Department of Animal Biotechnology, Faculty of Biotechnology
Jeju National University, Jeju, REPUBLIC OF KOREA
[#]Both Authors Contributred equally

**Corresponding Author: Dong Kee Jeong; Email: dkjeong@jejunu.ac.kr*

Received: 15 October 2013; Accepted: 27 November 2013

ABSTRACT

Canine transmissible venereal tumor (CTVT) is a histiocytic tumor of the dog that mainly affects the external genitalia, commonly found at tropical and subtropical zones. In the present investigation, we undertook this work mainly to perform a computational analysis of snSNP in the BTNL2 to identify the possible mutations and proposed the model structure of the mutant protein. Four deleterious mutations were identified in BTNL2 in 109 and 319 residual positions. Moreover, we constructed the homolgus structure of native and mutant proteins to predicate the stability. I-Mutant was used for routine analysis of protein stability and for the single site mutation analysis. It was found that mutation of L to S at residual position 109 and A to T at 319 residue position has shown maximum negative effect on the protein stability and considered for further analysis. The mutational effect on the protein function was analyzed by project HOPE. It was found that the wild-type residue is very conserved, but a few other residue types have been observed at this position too. Based on conservation scores this mutation is probably damaging to the protein. The present investigation was further used for molecular expect of the CTVT infection which might be useful in diagnosis and prevention of CTVT in canine.

Keywords: BTNL2; canine transmissible venereal tumor; immunoglobulin V; non synonymous mutation; SNP.

Sticker's sarcoma also known as Canine transmissible venereal tumor (CTVT), is a very unusual form of cancer affecting canines. CTVT is a histiocytic tumor of the

dog and other canids that mainly affects the external genitalia, which transmitted during copulation (Cohen *et al.*, 1985). Furthermore, the effect of the CTVT have also been reported in the inguinal skin, lymph nodes, lips, and buccal and nasal mucosa whereas, less effect have reported in the tonsils, liver, pancreas, spleen, and mesenteric lymph lung, kidney (McLeod *et al.*, 1972; Placke *et al.*, 1987; Mozos *et al.*, 1996). Several lines of evidence provided indirect support for this hypothesis (Das and Das., 2000). It has a worldwide distribution, but is detected mainly in tropical and subtropical zones (Varaschin *et al.*, 2001).

Previously, it was suspected that viruses were the cause of spreading CTVT, like in the same way that the human papilloma virus spreads cervical cancer to women through sex. However, the new genetic analysis was shown that the dog cancer cells are direct descendents of tumour cells from the long-dead animal in which the disease originated. Murgia *et al.* (2006) proved the allograft hypothesis for CTVT. According to this allograft hypothesis, tumour cells from different animals should be genetically clustered and different from normal cells of the host animal.

The sequencing of the microsatellite of the DNA regions showed that CTVT from different animals had less variability than what is observed within the most inbred breed of dogs. Therefore, the transmition of tumor is occurr from one dog to another which confirming the spread of this tumour from an ancestral clone. Indeed, CTVT could be quoted as the "older cancer known to science". The CTVT is one of only three known transmissible cancers; another is Devil facial tumor disease (DFTD), a cancer which occurs in Tasmanian devils. It is quite interesting that karyotype anomalies of DFTD derevitated cells have shown the similar to those of cancer cells isolated from CTVT (Bostanci, 2005). Murchison *et al.* (2012) reported that mutated or deleted genes such as RET, FANCD2, MAST3 and BTNL family gene might be associated with DFTD.

It has been proposed that BTNL proteins family might be particularly important in mediating immune regulation in tissues. However, the polymorphisms of BTNL family could leads to sarcoidosis, inflammatory bowel disease and myositis in human (Valentonyte *et al.*, 2005; Mochida *et al.*, 2007). The present work was focused on buytrophilin type (BTNL) family gene from dog, which is conspicuously cause for DFTD. We selected BTNL2 gene from dog for present annotation. At the time of the manuscript preparation, no report has been found on the *in-silico* analysis of functional nsSNPs of the BTNL2 gene from dog. The aim of this present investigation to understand about the mutational effect on the BTNL2 gene from dog might be the cause for CTVT disease. Therefore, we undertook this work mainly to perform a computational analysis of snSNP in the BTNL2 to identify the possible mutations and proposed the model structure of the mutant protein.

MATERIALS AND METHODS

Retrieval of the information of BTNL 2 gene from the database

The data for BTNL family gene from dog was retrieved from Entrez Gene on NCBI (http://www.ncbi.nlm.nih.gov/gene). The single nucleotide polymorphisms (SNPs) and their related information were obtained from the NCBI dbSNP (http://www.ncbi.nlm.nih.gov/snp/) and SWISSProt database (http://expasy.org/) for computational analysis.

Predication of non-synonymous SNP of BTNL 2 using Sorting intolerant from tolerant

SIFT (http://www.blocks.fhcrc.org/sift/SIFT.html) was used to predict whether an amino acid substitution affects protein function. Ng and Henikoff (2011) reported that SIFT was discrimates between the functionally neutral and deleterious polymorphisms in the mutagenesis studies. The threshold levels for intolerance score were fixed to be 0.005 or less. SIFT scores were classified as damaging (0.00-0.05), potentially damaging (0.051-0.10), borderline (0.101-0.20), or tolerant (0.201- 1.00) (Hussain *et al.*, 2012).

Protein stability predication using I-Mutant

The I-Mutant was used for routine analysis of protein stability and for the single site mutation analysis. The FASTA sequence of protein was used as an input to predict the mutational effect on protein stability. It was used for the automatic prediction of protein stability changes upon single point mutations (http://folding.biofold.org/cgi-bin/i-mutant2.0.cgi).

Homology modeling of BTNL2and validation of the model

The three dimensional structure of BTNL2 from dog was generated by homology modelling using SWISS-MODEL workspace (http://swissmodel.expasy.org/). The Psi/Phi Ramachandran plot was generated through the PROCHECK analysis. The packing quality of the refined structure was investigated by the PROCHECK Quality Control Check.

Annotation of the effect of mutation on the protein by Project Hope

The effect of the mutation on the protein function was analyzed by project HOPE (Have yOur Protein Explained). The protein sequence was used as the input for selection of the mutant variants. The output is given in the form of the structural variation between mutant and wild types residues.

RESULTS AND DISCUSSION

In the present investigation, total four missense variants, 37 intron variant, and one variant from the each 5' and 3' UTR prime was found in dog BTNL2. We selected nonsynonymous coding nsSNPs for our investigation (Table 1). The consequences of four nsSNPs were submitted to SIFT. It was quite interesting that at 109 residues

position, leucine mutate with serine and threonine could show the deleterious effect on the protein function. Furthermore, at 319 position, the change of alanine with serine and threonine might be responsible for nonsynonimus mutation. However, all of those four nsSNP could share the common SNP ID. The SNPs are hypothesized to play an important role in the functional and structural analysis of the protein. SNPs are the common form of genetic variations among individuals, and are thought to be accountable for the majority of heritable traits, including a large portion of the inherited disease vulnerability (Alanazi *et al.*, 2011). Thus, those SNPs are hypothesized to play an important role in the functional and structural analysis of the protein.

Table1: List of the nsSNPs that were predicted to have functional significance by SIFT

Transcript	SNP ID	Amino acid change	Amino acid Position	Tolerance index in SIFT
ENSCAFT00000001240	rs8795693	L/S	109	0
ENSCAFT00000001247	rs8795693	L/T	109	0
ENSCAFT00000001240	rs8795693	A/S	319	0
ENSCAFT00000001247	rs8795693	A/T	319	0

I-Mutant2.0 predictions were performed starting either from the protein structure or, more importantly, from the protein sequence. The effect of the mutation on the protein stability was evaluated with the free energy change value (Table 2). It was observed that mutation at the 109 reidues position with S and T have shown maximum effect on the protein stability. However, the less effect on the stability was estimated on the mutation of A à S at 319 residue position. Therefore, in the present investigation, we predicated the mutational effect on the stability index of the protein. Khan and Vihinen (2010) reported that the accuracy of the SIFT was further validated by the most reliable predictor I-Mutant. Furthermore, based on the I-Mutant result, we considered the mutation point of L à S at residual position 109 and A à T at 319 residue position for further analysis.

Table 2: Determination of the protein stability by I-Mutant

Position	WT	New	DDG	Stability
109	L	S	-2.46	Decrease
109	L	T	-2.56	Decrease
319	A	S	-0.65	Decrease
319	A	T	-1.01	Decrease

The 3D structures of the native and mutant proteins were modeled using SWISS-MODEL expasy. The suitable template for the homology modeling was searched by BLASTp for the target protein. The template 4hh8A was selected for the homology model (Figure 1). The constructed model was validated through PROCHECK. The PROCHECK program was applied to check out the overall stereochemical quality

of the model. The darkest areas correspond to the "core" regions representing the most favorable combinations of phi-psi values. The validation by PROCHECK certifies that 87.9% of the residues in these "core" regions, while, 11 % residues were found in the additional allowed regions (Figure 2). The average Z score, RMS of Z score, and distribution of the atomic Z score are represented in figure 3. The OMEAN Z score was found -0.33 (Figure 3).

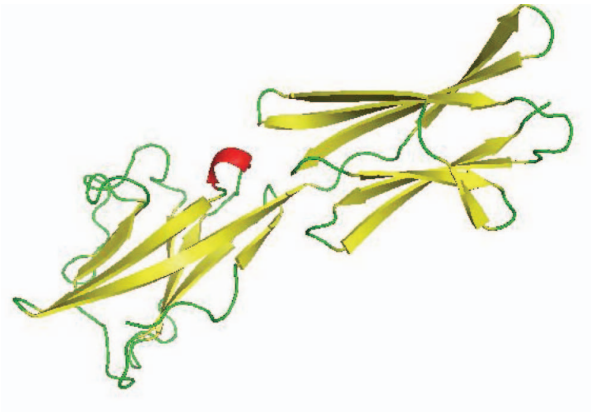

Fig. 1: The Predicated homology 3D structure of BTNL2 from dog. The red color indicated the helix, yellow for sheet and green indicate the loop of the structure.

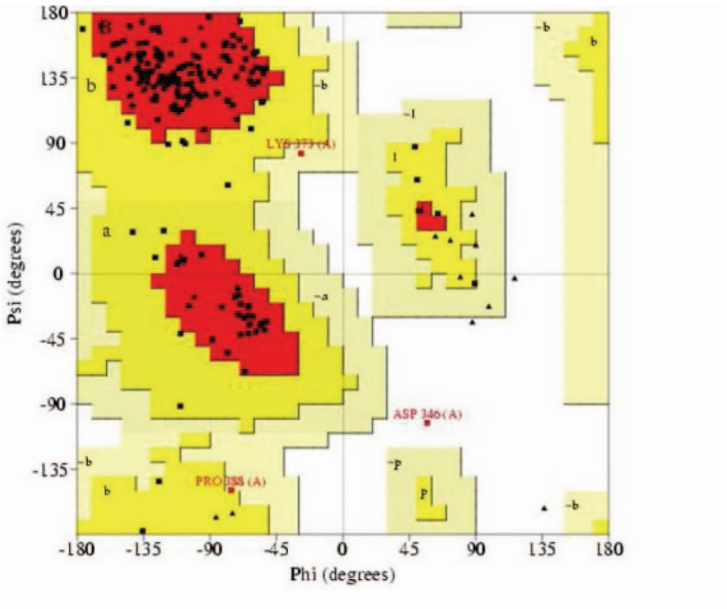

Figure 2: The Ramachandran Plot for quality check of BTNL2 from dog. Regions A, B, L as most favored (87.9%), additional allowed regions are a, b, l, p (20%), residues in generously allowed regions 0.6%% and residues in the disallowed regions 0.6% were found.

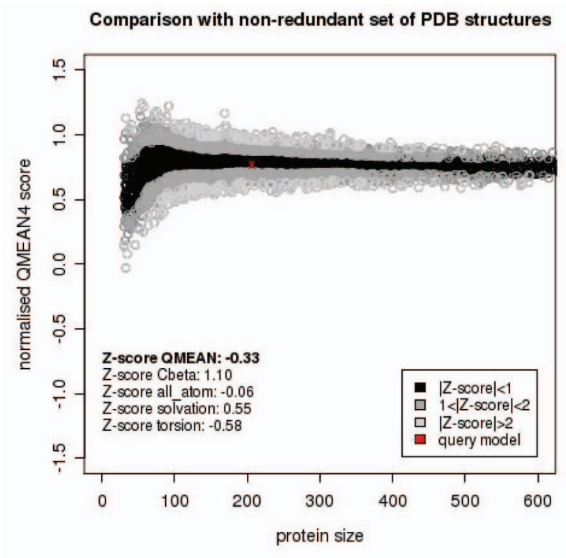

Figure 3: QMEAN Z score of the model. All the statistical Z score term was depicted on the left corner of the figure and the reliable range of the Z score for the model was given in the right corner.

The residues L and A are the part of an interpro domain named "Immunoglobulin V-set, subgroup" reinvolved in a wide variety of functions which usually requires interaction of the intact domain with another protein/molecule. A mutation in such a domain could disturb this interaction. Moreover, The mutation is located within a domain, annotated in Uniprot as: "Ig-like V-type 1". The mutation introduces an amino acid with different properties, which can disturb this domain and abolish its function. The wild-type residue is very conserved, but a few other residue types have been observed at this position too. Based on conservation scores this mutation is probably damaging to the protein. It was found that the mutation is located in a region with known splice variants, isoform 6. The wild-type residue is not conserved at this position. Another residue type was observed more often at this position in other homologous sequences. This means that other homologous proteins exist with that other residue type than with the wild-type residue in your protein sequence. The other residue type is not similar to your mutant residue. Therefore, the mutation is possibly damaging.

B-lymphocytes were considered the primary source of serum Ig. However, it was reported that the cells from epithelial cancer and hyperplasia could also express Ig (Nzula *et al.*, 2003). Zheng *et al.*, (2009) demonstrated that functional Ig gene recombination and transcription occurred in a variety of cancer types. Lindley and Steele (2012) reported that inappropriately activation of Ig somatic hypermutation in many non-lymphoid tissues via hormonal and/or inflammation-related processes may be leading to cancer. Therefore, from the present investigation, we have predicted

that the mutation causes in the 109 and 319 residues could be led to the tumor development in dogs.

CONCLUSION

A major goal of this paper was to provide an explanation about the mutation effect in BTNL2 from dog. It is well reported that CTVT in dog and DFTD in Tasmanian devils has shown similar pattern of infection and karyotypic analysis. Therefore, we selected that gene, which is responsible for DFTD in Tasmanian devils and analysis the mutational effect in dog. However, it is the first report on the nsSNP analysis of BTNL2 from dog to understand the effect of the mutation in this gene which cause to lead proffering cancer in dog. From the present investigation, it was coming to clear that mutations in the residual position 109 and 319 might be the cause for the tumor development in dog breeds. In this paper, we try to identify the non synonymous residues having effect on the tumor development, however, further studies has needed for better understand at molecular expect to prevent the spreading CTVT in dog.

ACKNOWLEDGEMENT

Present study was supported by a grant from Golden Seed Project (No.12567), South Korea, hence the authors are thankful to this organization. Authors are Equally thanks full to the Indian Council of Agricultural Research, New Delhi, India for providing internastional fellowhsip.

REFERENCES

Alanazi, M., Abduljaleel, Z., Khan, W., Warsy, A. S., Elrobh, M., Khan, Z., Al, Amri Abdullah. and Bazzi, M. D. 2011 In Silico Analysis of Single Nucleotide Polymorphism (SNPs) in Human â-Globin Gene. *PLOS One*, **6** (10): e25876.

Bostanci, A. 2005. "A Devil of a Disease". *Science*, **307** (5712): 1035.

Cohen, D. 1985 The canine transmissible venereal tumor: a unique result of tumor progression. *Advance Cancer Research*, **43**: 75-112.

Das, U. and Das, A.K. 2000 Review of canine transmissible venereal sarcoma. *Veterinary Research Communication*, **24**: 545-56.

Hussain, M.R.M., Shaik, N.A., Yousuf, Al-Aama, J., Asfour, H.Z., Khan, F.S., Masoodi, T. A., Khan, M.A. and Shaik, N.S. 2012 In silico analysis of Single Nucleotide Polymorphisms (SNPs) in human BRAF gene. Gene **508**: 188-196.

Khan, S. and Vihinen, M. 2010 Performance of protein stability predictors. Human Mutation **31**: 675-678.

Lindley, R.A. and Steele, E.J. 2012. Critical Analysis of strand-biased somatic mutation signatures in TP53 versus Ig genes, in genome-wide data and the etiology of cancer. ISRN genomics. http://dx.doi.org/10.1155/2013/921418

McLeod, C.G. and Lewis, J.E. 1972 Transmissible venereal tumor with metastases in three dogs. *Journal of American Veterinary Medicine Association*, **161**: 198-200.

Mochida, A., Kinouchi, Y., Negoro, K., Takahashi, S., Takagi, S., Nomura, E., Kakuta, Y., Tosa, M. and Shimosegawa, T. 2007. Butyrophilin-like 2 gene is associated with ulcerative colitis in the Japanese under strong linkage disequilibrium with HLA-DRB1*1502.

Tissue Antigens, **70**:128–135.

Murgia, C., Pritchard, J. K., Kim, S. Y., Fassati, A. and Weiss, R.A. 2006 Clonal origin and age of a naturally transmissible cancer. *Cell*, **1126** (3): 477-487.

Murchison, P.E., Schulz-Trieglaff O.B. , Ning, Z., Alexandrov, L.B., Bauer, M.J., Fu, B., Hims, M., Ding, Z., Ivakhno, S., Stewart, C. and Ng, B.L. *et al.*, 2012 Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer.". Cell **148** (4): 780–791.

Mozos, E., Mendez, A., Gomez-Villamandos J. C, Martín de las Mulas, J. and Perez, J. 1996. Immunohistochemical characterization of canine transmissible venereal tumor. *Veterinary Pathology*. **33:** 257.

Ng, P.C. and Henikoff, S. 2001. Predicting detelerious amino acid substitutions. *Genome Research*. **11**: 863-874.

Nzula, S., Going, J.J. and Stott, D.I. 2003 Antigen-driven clonal proliferation, somatic hypermutation, and selection of B lymphocytes infiltrating human ductal breast carcinomas. *Cancer Research*, **63**(12):3275-80.

Placke, M.E., Hill, D.L. and Yang, T.J. 1987. Cranial metastasis of canine transmissible venereal sarcoma. *Journal of Veterinary Medicine Association*, **34**: 125-132.

Valentonyte, R., Hampe, J., Huse, K., Rosenstiel, P., Albrecht, M., Stenzel, A., Nagy, M., Gaede, K.I., Franke, A., Haesler, R., Koch, A., Lengauer, T., Seegert, D., Reiling, N., Ehlers, S., Schwinger E., Platzer, M., Krawczak, M., Müller-Quernheim, J., Schürmann, M. and Schreiber, S. 2005. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nature Genetics*, **37**:357–364.

Varaschin, M. S., Wouters, F., Bernins, V. M. O, Soares, T. M. P., Tokura, V. N. nd Dias, M. P. L. L. 2001 Tumor venereo transmissivel canino na regiao de Alfenas, Minas Gerais; formas de apresentacao clinico-patologicas. *Clinica Veterinaria*, **6**: 332-338.

Zheng, J., Huang, J., Mao, Y., Liu, S., Sun, X., Zhu, X., Ma, T., Zhang, L., Ji, J., Zhang, Y., Yin, C.C. and Qiu, X. 2009 Immunoglobulin Gene Transcripts Have Distinct VHDJH Recombination Characteristics in Human Epithelial Cancer Cells. *The Journal of Biological Chemistry*, **284**:13610-13619.