Identification of biological pathways and genes associated with neurogenic heterotopic ossification by text mining

Yichong Zhang1,*, Yuanbo Zhan2,*, Yuhui Kou1, Xiaofeng Yin1, Yanhua Wang1 and Dianying Zhang1

1 Department of Trauma and Orthopaedic Surgery, Peking University People’s Hospital, Beijing, China
2 Department of Periodontology and Oral Mucosa, the Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China
* These authors contributed equally to this work.

ABSTRACT

Background: Neurogenic heterotopic ossification is a disorder of aberrant bone formation affecting one in five patients sustaining a spinal cord injury or traumatic brain injury (SCI-TBI-HO). However, the underlying mechanisms of SCI-TBI-HO have proven difficult to elucidate. The aim of the present study is to identify the most promising candidate genes and biological pathways for SCI-TBI-HO.

Methods: In this study, we used text mining to generate potential explanations for SCI-TBI-HO. Moreover, we employed several additional datasets, including gene expression profile data, drug data and tissue-specific gene expression data, to explore promising genes that associated with SCI-TBI-HO.

Results: We identified four SCI-TBI-HO-associated genes, including GDF15, LDLR, CCL2, and CLU. Finally, using enrichment analysis, we identified several pathways, including integrin signaling, insulin pathway, internalization of ErbB1, urokinase-type plasminogen activator and uPAR-mediated signaling, PDGFR-beta signaling pathway, EGF receptor (ErbB1) signaling pathway, and class I PI3K signaling events, which may be associated with SCI-TBI-HO.

Conclusions: These results enhance our understanding of the molecular mechanisms of SCI-TBI-HO and offer new leads for researchers and innovative therapeutic strategies.

Subjects Bioinformatics, Orthopedics, Surgery and Surgical Specialties
Keywords Heterotopic ossification, Spinal cord injury, Traumatic brain injury adipogenesis, Text mining

INTRODUCTION

Heterotopic ossification (HO), also known as heterotopic bone formation, is the process of benign bone formation and growth outside of normal skeletal locations (Sullivan et al., 2013). Neurogenic heterotopic ossification (NHO) was first described at the end of the nineteenth century and was predominantly observed among soldiers who had experienced spinal cord trauma during combat in World War I (Shehab, Elgazzar & Collier, 2002). At present, three subtypes of NHO have been described: traumatic, neurologic, and genetic. The latter subtype derives from a rare hereditary disorder associated with NHO.
The first two subtypes, by contrast, are more common and are acquired following various traumas, fractures, dislocations, elective orthopedic surgeries (e.g., hip surgery), electrocution and burn injuries, and neurological damage. Specifically, neurologic NHO refers primarily to the spinal cord and traumatic brain injury. In spinal cord-injured patients, the incidence of ectopic bone formation is between 10% and 25% (Sakellariou et al., 2012; Vanden Bossche & Vanderstraeten, 2005), with most formation occurring in the hip region. Ectopic bone formation occurs in 40–50% of patients with traumatic brain injury, with the hip (Sakellariou et al., 2012), shoulder, and elbow being common sites of formation. Clinically, NHO caused by traumatic brain injury always affects the joints of both upper and lower extremities, but NHO caused by spinal cord injury always occurs predominantly below the level of the lesion. Ossification is also generally seen anteriorly involving the iliopsoas and femoral neurovascular structures, laterally within gluteus minimus, and posteriorly extending from the ilium to the posterior femur encasing the sciatic nerve (Sullivan et al., 2013). The occurrence of NHO is problematic for patients with brain and spinal cord injuries, as it can cause pain, inflammation, and loss of range of motion, as the joint gradually becomes ankylosed, leading to greater functional disability (Genêt et al., 2011). Although NHO has long been reported, and the relevant literature is continuously growing, the etiology and mechanisms of neurologic NHO have not yet been recognized. There are currently no efficacious means of prevention and few treatment options for neurologic NHO. Once acquired NHO develops, surgical removal is the only effective treatment, and is usually followed by local radiation or nonsteroidal anti-inflammatory drugs (NSAIDs) to prevent recurrence (Sullivan et al., 2013; Cullen & Perera, 2009). Indomethacin is now the gold standard for pharmacological prevention of NHO (Sullivan et al., 2013). Another drug, rofecoxib, significantly decreases the rate of NHO and has been used as a primary prevention strategy for NHO, although it was taken off the market by the United States Food and Drug Administration in September 2004 (Sullivan et al., 2013). Moreover, the efficacy of bisphosphonates (specifically etidronate) as a primary prevention strategy for neurogenic NHO in patients with spinal cord and traumatic brain injury has been reported and reviewed (Sullivan et al., 2013; Sakellariou et al., 2012). The lack of preventative and treatment options may be attributed to a lack of accepted animal models (Kan & Kessler, 2011), and many researchers only investigate and report the epidemiological features (incidence, recurrence, and localization in patients) (Genêt et al., 2011), although still in the experimental stages Reznik et al. (2017a, 2017b) report upon the novel intervention of extracorporeal shock wave therapy for the management/treatment of NHO in TBI.

Over the course of time, the biomedical literature has rapidly expanded. When clinicians and researchers search the literature for relevant information (e.g., the genes or pathways associated with HNO), they are often faced with a common problem: there is such a large quantity of literature that it is difficult to collate all the relevant information in their field, and even more so outside it (Hristovski, Rindflesch & Peterlin, 2013). Fortunately, text mining technology can be used as an approach to gain a broad understanding of an entire dataset and the dynamics of the study issues. More importantly,
text mining can pose a hypothetical question or find potential relationships (that may otherwise be ignored) between entities of interest. One relevant area of research is literature-based discovery (LBD); some researchers have successfully used this strategy to promote new discoveries for diseases (Hristovski, Rindflesch & Peterlin, 2013; Karic et al., 2011; Swanson, 1986; Karić & Karić, 2011; Zhan et al., 2017; Swanson, 1988).

One of the most well-known examples of the successful use of LBD is the discovery that patients with Raynaud’s disease could benefit from the consumption of fish oils (Swanson, 1986).

Nearly all LBD systems are inspired by or derived from Swanson’s ABC co-occurrence model (Swanson, 1988; Henry & McInnes, 2017). The methods and evolution of LBD have been introduced and reviewed elsewhere (Henry & McInnes, 2017; Hristovski et al., 2005). Briefly, “A implies B” and “B implies C” relationships are generated by the explicit knowledge that already exists in the literature, then implicit knowledge is discovered by drawing a “therefore A implies C” conclusion (Henry & McInnes, 2017). There are two main approaches to LBD: open discovery and closed discovery (Fig. 1) (Hristovski et al., 2006). These two methodologies have different focuses. Open discovery requires a single term/concept (“A,” e.g., a disease), and then first-order terms that relate to the single term/concept are categorized (“B,” e.g., drug). Then, those first-order concepts are further categorized (“C,” e.g., gene). Thus, information flow can be linked from “A” (disease) to “B” (drug), then to “B” (drug) to “C” (gene). If these “C” (gene) terms have not been mentioned in the literature, then they are potentially related to “A” and “B,” but are
untested. Closed discovery requires a start term/concept and a target term/concept, and then seeks intermediate links to generate potential explanations for the relationship between two terms/concepts (start and target) (Hristovski et al., 2006). Hristovski, Rindflesch & Peterlin (2013) and Weeber, Kors & Mons (2005) have provided a detailed review of the different functions of mining tools available in the literature. Some researchers have successfully used text mining tools to identify candidate genes for diseases (Hristovski, Rindflesch & Peterlin, 2013; Karić & Karić, 2011).

In the present study, we integrate the text mining methodologies and other dataset (gene expression and drug target) to discover candidate genes or molecules that may play important roles in NHO caused by spinal cord and traumatic brain injury (SCI-TBI-HO). To the best of our knowledge, this is the first time this strategy has been applied in this area.

**MATERIALS AND METHODS**

**Discovery of potential SCI-TBI-HO mechanisms**

BITOLA (Hristovski et al., 2006) (http://ibmi.mf.uni-lj.si/bitola) is an interactive literature-based biomedical discovery support system (Hristovski et al., 2006). This system features two types of methodology: closed and open discovery. The closed discovery methodology allows the input of two concepts (e.g., two disorders) and generates terms for potential explanations of their relationship (such as “Genes,” “Cell Function,” “Cell or Molecular Dysfunction,” and “Molecular Function”) (Hristovski et al., 2006). By contrast, the major role of the open discovery methodology is to mine new content (genes, drugs, or neuroscience, etc.) associated with the disease. Hence, we first deemed HO and spinal cord or traumatic brain injury as two concepts and attempted to explain the relationship between them by using the closed discovery methodology (Fig. 2). Secondly, we explored the candidate genes or molecules associated with NHO caused by spinal cord and traumatic brain injury (SCI-TBI-HO) using the open discovery methodology (Fig. 3).

The set of concepts available in the BITOLA system are currently Medical Subject Headings (MeSH), which are utilized to index Medline, and human genes from the Human Genome Organization (Hristovski et al., 2006). After searching for the keywords “Heterotopic ossification” and “Ectopic ossification,” we found only one concept, namely, “Ossification, heterotopic” (C0029396). We further verified this keyword by searching the MeSH database (https://www.ncbi.nlm.nih.gov/mesh/), and confirmed that the standard term for HO is “Ossification, heterotopic” (MeSH Unique ID: D009999). Therefore, “Ossification, heterotopic” was used for subsequent analysis.

Equivalent terms for the “Spinal cord and traumatic brain injury,” such as “Brain Injury (C0270611)” and “Compression of Spinal Cord,” indeed exist and are used in the literature. In order to incorporate other concepts that have a close relationship with, or are aliases for, spinal cord and traumatic brain injury, we searched for related disorders in BITOLA, as recommended. Subsequently, we included “Craniocerebral Trauma (C0018674)” and “Compression of Spinal Cord (C0037926),” which have a close relationship with spinal cord and traumatic brain injury, in the search (Table S1).
Specifically, in this study, in order to enhance understanding of the molecular mechanisms for SCI-TBI-HO, we applied the “Semantic Type” filter for “Gene or Gene Product.” In order to find the cell and molecular functions for HO caused by spinal cord and traumatic brain injury, we applied the “Semantic Type” filter for “Cell Function” (defined as “mining molecular function”). Lastly, the intermediate “Y” concepts were retrieved.

It is generally recognized that pathway-based analysis may provide more meaningful information than gene-based analysis. Therefore, the relevant pathways for the “Gene or Gene Product” filtered terms were analyzed by functional enrichment using FunRich software (http://www.funrich.org) (Pathan et al., 2015). For ease of description, we define the “Brain Injuries (C0270611),” and “Cranioencebral Trauma (C0018674) as TBI-HO (caused by brain injury), and define “Spinal Cord Injuries (C0037929)” and “Compression of spinal cord (C0037926)” as SCI-HO (caused by spinal cord injury).
Discovery of candidate genes or molecules associated with SCI-TBI-HO

As above, “Heterotopic ossification” was used as a starting concept (X), and “Spinal Cord Injuries (C0037929),” “Brain Injuries (C0270611),” “Craniocerebral Trauma (C0018674),” and “Compression of spinal cord (C0037926)” were used as the second concepts. Finally, the related “Z” terms (which here refer primarily to the “Gene or Gene Product” semantic type) were explored by defining the semantic type as “genes or molecules” and the semantic group as “Any.”

Via above steps, the “Gene or Gene Product” (“Z”) candidate genes/molecules were constructed, and this candidate genes/molecules set was used for the next analysis.

Integration of gene expression profiles to filter the candidate genes or molecules

If the above candidate genes/molecules are expressed abnormally in the course of the NHO, it indicates that they may have important role in the progress of disease. Hence, we
tentatively filtered and evaluated the “candidate genes/molecules” through the overview of gene expression levels (mRNA, Messenger Ribonucleic Acid) under different conditions (disease vs. control). Finally, we detected whether the candidate genes/molecules extracted by text mining matched with those different expression genes (DEGs). If they did, the candidate genes/molecules would be retained for next analysis, and if they did not, it would be excluded. We defined those genes that passed the filter criterion as “promising genes.”

**Microarray data**

Gene expression datasets were obtained from Gene Expression Omnibus (GEO) database. For the NHO, there was only one datasets (GSE94683) associated with the NHO. The more detail information about GSE94683 can refer to Torossian et al. (2017). Briefly, the platform used in GSE94683 was GPL10630 Agilent-021531 Whole Human Genome Oligo Microarray 4×44K (Feature Number version). The Whole Human Genome Oligo Microarrays were used to compare expression profiling of mesenchymal stromal cells (MSCs) from NHO patients and from bone marrow of healthy donors (control). Total RNA extracted from NHO-MSCs (7 samples) and Health-MSCs (9 samples) was used for cDNA array hybridization.

**Statistical analysis of microarray data**

Gene chip data of GSE94683 were analyzed by using BRB Array Tools (version 4.5.1 and R version 3.2.5; http://linus.nci.nih.gov/BRB-ArrayTools.html). The raw expression data were converted to log2 values and then normalized by using the quantiles normalization. In terms of spot filters, spots with intensity <10 were removed. The replicate spots within an array were averaged. Moreover, the genes under any of the following conditions were excluded: less than 20% of expression data have at least a 1.5-fold change in either direction from gene’s median value percent of data missing or filtered out exceeds 50%.

**Analysis of differentially expressed genes**

To identify the DEGs between NHO-MSCs and Health-MSCs, a random-variance t-test was used. Genes were considered highly significant if their P value was less than 0.001 and the false discovery rate (FDR) less than 0.05.

**Filtering of the intermediate concepts by manually checking**

False-positive may existed due to the defects of literature mining itself. For example, the ambiguity of gene symbols (the symbol may refer to something other than the gene), is a known problem in text mining. Thus, we manually checked the “promising genes” to exclude the ambiguous term by reading the co-occurrence literature.

**Explore the relationship between the promising genes and drug targets**

Clinically, multiple randomized controlled trials support the use of various NSAIDs for primary prevention of NHO following CNS injury (Banovac et al., 2004, 2001), indomethacin, rofecoxib, and etidronate being the most commonly used drugs.
We identified the Anatomical Therapeutic Chemical (ATC) classes for indomethacin, rofecoxib, and etidronate using iATC-mHyb (Cheng et al., 2017) (http://www.jci-bioinfo.cn/iATC-mHyb). The target genes were extracted for these three drugs. Firstly, we used the Comparative Toxicogenomics Database (Davis et al., 2017) (CTD; http://ctdbase.org) to explore which genes/proteins interact with indomethacin, rofecoxib, and etidronate. Finally, we detected whether the “promising genes” were also matched with those target genes.

**Explore the expression pattern of the promising genes by tissue-specific gene data**

The “promising genes” were further analyzed by tissue-specific gene data using the Human eFP Browser (Patel, Hamanishi & Provart, 2016). Clinically, the HO caused by traumatic brain injury always affects the joints of both upper and lower extremities, and HO caused by spinal cord injury always occurs below the level of the lesion. This ossification is also seen anteriorly, involving the iliopsoas and femoral neurovascular structures, laterally within gluteus minimus, and posteriorly extending from the ilium to the posterior femur encasing the sciatic nerve (Sullivan et al., 2013). The Human “electronic Fluorescent Pictograph” (eFP) Browser (Patel, Hamanishi & Provart, 2016) was used to interrogate whether the “promising genes” were most strongly expressed in these tissues.

**RESULTS**

**Discovery of potential SCI-TBI-HO mechanisms**

Neurogenic heterotopic ossification is the ectopic formation of lamellar bone in non-osseous tissues following traumatic brain or medullar injury (Debaud et al., 2017). NHO is usually observed after insults such as spinal cord injury and traumatic brain injury (Genêt et al., 2011). Despite its known associations, etiology, and risk factors, the pathogenesis of HO remains poorly understood (Sakellariou et al., 2012).

Hence, we explored the potential mechanisms for SCI-TBI-HO using LBD.

In the closed discovery methodology, “Heterotopic ossification” was used as starting concept (similar to “A” in Swanson’s ABC model), and “Spinal Cord Injuries (C0037929),” “Compression of spinal cord (C0037926),” “Brain Injuries (C0270611),” and “Craniocerebral Trauma (C0018674)” used as end concepts (similar to “B” in Swanson’s ABC model).

As results, 51 and 47 genes were “Gene or Gene Products” were suggested for TBI-HO and SCI-HO, respectively (Tables S2–S7) In total, 33 (50.8%) genes were overlapped between TBI-HO and SCI-HO. By functional enrichment analysis of these “Gene or Gene Product,” we found that 50 (89.3%) pathways were overlapped between TBI-HO and SCI-HO, including BMP receptor signaling, ALK1 pathway, Calcineurin-regulated NFAT-dependent transcription in lymphocytes, Integrin family cell surface interactions, Alpha9 beta1 integrin signaling events, Beta1 integrin cell surface interactions, Insulin Pathway, Internalization of ErbB1, Urokinase-type plasminogen activator (uPA) and uPAR-mediated signaling, PDGFR-beta signaling pathway, EGF receptor (ErbB1) signaling pathway, and Class I PI3K signaling events, and so on (Fig. 4A). However, some
pathways appear only in one state, such as "ALK1 signaling events and Polyamines are oxidized to amines," "aldehydes and H2O2 by PAOs" were enriched only for TBI-HO. Glucocorticoid receptor regulatory network and Glucocorticoid receptor signaling were enriched only for SCI-HO.

In addition, we mined relative cell function for associations with SCI-TBI-HO. We showed that 22, 9, 22, and 13 terms related to cell function are associated with "Spinal Cord Injuries," "Compression of spinal cord," "Brain injuries," and "Craniocerebral Trauma," respectively (Tables S8–S10). Predictably, cell function was mainly involved in Nerve Regeneration, Cell Differentiation process, and Neural Conduction for Spinal Cord Injuries; Neural Conduction, Nerve Regeneration, and Cell division for Compression of spinal cord; Cell Differentiation process, Nerve Regeneration, Cell division, and Apoptosis for Brain Injuries; Metabolic Process, Cellular, Neural Conduction, Neural Conduction, Cell Differentiation process, and Nerve Regeneration for Craniocerebral Trauma (Tables S8–S10). An overlap of the four concepts, Spinal Cord Injuries, "Compression of spinal cord," "Brain injuries," and "Craniocerebral Trauma," yielded 7 terms: Nerve Regeneration, Cell Differentiation process, Neural Conduction, Cell division, Cell survival, Apoptosis, and Signal Transduction (Fig. 5). However, the Embryonic induction were only for TBI-HO (caused by brain injury), and Cell Aggregation were only for SCI-HO (caused by spinal cord injury).

**Candidate genes or molecules derived from text mining**

In the open discovery methodology, "Heterotopic ossification" was used as a starting concept, and "Spinal Cord Injuries (C0037929)," "Brain Injuries (C0270611)," "Craniocerebral Trauma (C0018674)," and "Compression of spinal cord (C0037926)"
were defined as second concepts (similar to “B” in Swanson’s A to B to C model). In total, 526 “Gene or Gene Product” terms (candidate genes) for TBI-HO and 428 candidate genes for SCI-HO were found by the BITOLA system (Table S11).

Differentially expressed genes

After data pre-processing and class comparison between different groups (NHO-MSCs vs. Health-MSCs), 187 DEGs were identified (Table S12). And 7 genes (PTS, KCNN4, BAMBI, GDF15, LDLR, CCL2, and EBP) were overlapped between the DEGs and the “Candidate genes or molecules” of TBI-HO, 3 genes (CCL2, PTS and CLU) were overlapped between the DEGs and the “Candidate genes or molecules” of SCI-HO.

Manual checking

After manual checking, PTS were removed since the abbreviation or the alternative names of those genes used in the articles were not the original idea for the “gene or gene product.” PTS in the text is for “Posttraumatic seizures” (Zimmermann, Martin & Girgis, 2017) or “Physical therapists” (Gómar-Toldrà, Sliwinski & Dijkers, 2014) but not for the gene PTS (6-pyruvoyltetrahydropterin synthase) in the co-occurrence literature. Hence, PTS
was removed and no further analysis. Therefore, the left 7 genes were seemed as promising genes.

**The promising genes were also the drug targets for indomethacin, rofecoxib, and etidronate**

The consultation of the Comparative Toxicogenomics Database (Davis et al., 2017) revealed 2,028, 115, and 17 genes that interacted with indomethacin, rofecoxib, and etidronate, respectively (Tables S13–S15). In total, combining them, we obtained 2,115 target genes (Table S16). By comparing the 7 promising genes with the drug target genes, we found 4 genes (GDF15, LDLR, CCL2, and CLU) seemed as most promising genes) that were present in both sets. Hence, our next step is to more deeply analyze the expression of these four genes in specific tissues.

**Expression of the most promising genes in specific tissues**

The analysis of tissue-specific gene expression may provide better explanations for each gene’s importance in the pathogenesis of HO caused by traumatic brain injury. We examined the gene expression patterns of the 4 most promising genes in specific tissues by using the Human eFP Browser (Patel, Hamanishi & Provart, 2016). These genes showed different expression patterns. In particular, we found that GDF15, LDLR, CCL2, and CLU had higher expression in the brain and spinal cord, they were also expression in smooth muscle and skeletal muscle. The first example output shown in Fig. 6 is GDF15. Detailed information for other genes can be found in the Supplemental Files (Tables S17–S24).

**DISCUSSION**

NHO is a disorder of aberrant bone formation affecting one in five patients sustaining a spinal cord injury or traumatic brain injury (SCI-TBI-HO) (Sullivan et al., 2013). Once formed, the heterotopic bone typically needs to be excised surgically, which may result in wound healing complications, in addition to a risk of recurrence. Moreover, prevention remains challenging, and the cellular mechanisms of NHO remain unclear.

Disease gene identification is an important step in understanding the pathogenesis of the disease (Zhan et al., 2014). The expanding rate of knowledge of gene-disease associations can hardly keep pace with the rapidly growing body of biomedical literature. A considerable amount of useful biological or biomedical knowledge is buried in the form of free text, waiting to be discovered and transformed by researchers into more useful information (Liu, Liang & Wishart, 2015). Fortunately, text-mining technology and biomedical text extraction can largely overcome the issue of buried information.

In this study, text mining was used to generate potential explanations for the occurrence of SCI-TBI-HO. As results, 51 and 47 genes were “Gene or Gene Products” were suggested for TBI-HO and SCI-HO, respectively (Tables S2–S5). In total, 33 (50.8%) genes were overlapped between TBI-HO and SCI-HO. By functional enrichment analysis of these “Gene or Gene Product,” we found that 50 (89.3%) pathways were overlapped between TBI-HO and SCI-HO, including BMP receptor signaling, ALK1 pathway,
Calcineurin-regulated NFAT-dependent transcription in lymphocytes, Integrin family cell surface interactions, Alpha9 beta1 integrin signaling events, Beta1 integrin cell surface interactions, Insulin Pathway, Internalization of ErbB1, uPA and uPAR-mediated signaling, PDGFR-beta signaling pathway, EGF receptor (ErbB1) signaling pathway, and Class I PI3K signaling events, and so on (Fig. 4A).

By using the closed discovery methodology and enrichment analysis, many pathways were found to be associated with TBI-HO and SCI-HO. In total, 50 (89.3%) pathways were overlapped between TBI-HO and SCI-HO. Among them, many pathways were related to integrin events, such as Integrin family cell surface interactions, Alpha9 beta1 integrin signaling events, and Beta1 integrin cell surface interactions. Generally, cells adhere to the substrate surface via integrins, a family of cell-surface receptors. Integrin-based attachment to the matrix plays a key role in physiological and pathological responses, such as development, tissue morphogenesis, adult tissue homeostasis, remodeling, and repair.
Particularly, the disturbance of the extracellular matrix-integrin-cytoskeleton signaling axis often results in disease and tissue dysfunction (Docheva et al., 2014; Loeser, 2014). Moreover, normal adult articular chondrocytes express many subtypes of integrins, including α1β1, α3β1, α5β1, α10β1, αVβ1, αVβ3, and αVβ5 (Loeser, 2014); these integrins modify their expression pattern and regulate chondrocyte differentiation in the articular cartilage (Garcia-Diegocázares et al., 2015). Importantly, integrin signals can also play important roles as regulators of osteogenesis and mineralization (Shekaran et al., 2014; Di Maggio et al., 2017). On the other hand, joint and relative osteogenesis processes are always implicated in SCI-TBI-HO. Therefore, the integrin-based pathways suggested by our study may provide a further understanding of SCI-TBI-HO. Other pathways, such as the Insulin Pathway (Zhang et al., 2014), Internalization of ErbB1 (Zhang et al., 2011), uPA and uPAR-mediated signaling (Kalbasi Anaraki et al., 2015), PDGFR-beta signaling pathway (Tokunaga et al., 2008), EGF receptor (ErbB1) signaling pathway (Zhang et al., 2011), BMP receptor signaling and Class I PI3K signaling events (Kalbasi Anaraki et al., 2015), have been associated with osteogenesis and mineralization or HO. More interestingly, the PDGFR-beta signaling pathway has been implicated in familial idiopathic basal ganglia calcification, a brain disease characterized by anatomically localized calcifications in or near the blood microvessels (Betsholtz & Keller, 2014); hence, this pathway deserves further research, to comprehensively investigate its association with SCI-TBI-HO.

Reportedly, the pathophysiology of NHO in these two conditions is not the same (Yildiz & Ardiç, 2010, Reznik et al., 2014), after enrichment analysis, we found that some pathways appear only in one state, such as “Polyamines are oxidized to amines and aldehydes and H2O2 by PAOs” and “Interconversion of polyamines” were enriched only for TBI-HO. Polyamines have been implicated in a large number of cellular processes, including functioning of ion channels, nucleic acid packaging, DNA replication, apoptosis, transcription (Childs, Mehta & Gerner, 2003). More importantly, the polyamines, spermine and spermidine, both were significantly elevated during HO (Davis et al., 2016). We have also found that “Glucocorticoid receptor regulatory network” and “Glucocorticoid receptor signaling” were enriched only for SCI-HO. It has been proved that glucocorticoid receptor expression was slightly increased at 15 min after spinal cord injury (Yan et al., 1999). However, glucocorticoid receptors expression was reported to decreased during 1–3 days after traumatic brain injury and increased 4–7 days (Hwang et al., 2006). Although not yet know whether there is different expression pattern of glucocorticoid receptors in HO caused by different reasons, those pathways provide new clues for detecting the different pathogenesis between TBI-HO and SCI-HO.

Previously, we have used a similar technique to that used in this study to detect the relative genes/candidate molecules in the interaction between oral lichen planus and depression (Zhan et al., 2017). We have innovatively used a text mining methodology alongside other dataset, including gene expression profiles data, drug target data and gene expression data of specific tissues, to explore promising genes for TBI-HO and SCI-HO.
We found three most promising genes (GDF15, LDLR, and CCL2) for TBI-HO, 2 most promising genes (CCL2 and CLU) for SCI-HO. These genes have been reported to play important roles in nerve injury, osteogenesis, or ossification. Among them, GDF15 (growth differentiation factor 15), is noteworthy, mainly owing to its special expression pattern and function (Fig. 6). GDF15 is a distant member of the TGFβ superfamily and is described as a novel trophic factor for motor and sensory neurons (Ruschke et al., 2012). Reportedly, elevated GDF15 expression is a hallmark of many tissue injuries, GDF15 mRNA is upregulated in lesioned neurons and leasoned brain (Strelau et al., 2003; Schindowski et al., 2011). In the other hand, GDF15 can also induce BMSCs toward osteogenic lineage (Uchiyama et al., 2015). These findings imply that GDF15 may have a direct or indirect relationship in SCI-TBI-HO, which warrants further study. Low density lipoprotein receptor (LDLR) had been correlation with the calcification in neck vessels (Awan et al., 2011). The expression of LDLR was also changed during mouse brain development and after ischemic stroke (Rousselet et al., 2011). CCL2, C-C motif chemokine ligand 2, is one of several cytokine genes. The important role of CCL2 in the brain injury (mainly about traumatic brain injury) have been broadly reported (Semple et al., 2010; Liu et al., 2013; Luo et al., 2016). Moreover, CCL2 has been shown to be induced and involved in various neurodegenerative disorders including Alzheimer’s disease, multiple sclerosis, and ischemic brain injury (Semple, Kossmann & Morganti-Kossmann, 2010). Previously, macrophage contribution to HO has also been noted in vivo (Mosser & Edwards, 2008), macrophages are capable of secreting numerous pro-inflammatory cytokines and chemokines, including CCL2 (Convente et al., 2015). Combined with the topic of our study (NHO), it is reasonable to believe that this gene requires further study for NHO. CLU, also known as clusterin, is a complement inhibitor that appears to have a neuroprotective effect in response to tissue damage, and its expression is upregulated following acute head injury (Troakes et al., 2017). In addition, it has previously been observed that secreted clusterin protein inhibits osteoblast differentiation of bone marrow mesenchymal stem cells (Abdallah, Alzahrani & Kassem, 2018). CLU should also be given future focus in the study of SCI-TBI-HO.

CONCLUSIONS

In conclusion, the underlying mechanisms of SCI-TBI-HO have been difficult to elucidate due to multiple factors. Using text mining, some pathways, such as the BMP receptor signaling, ALK1 pathway, Calcineurin-regulated NFAT-dependent transcription in lymphocytes, Integrin family cell surface interactions, Alpha9 beta1 integrin signaling events, Beta1 integrin cell surface interactions, Insulin Pathway, Internalization of ErbB1, uPA and uPAR-mediated signaling, PDGFR-beta signaling pathway, EGF receptor (ErbB1) signaling pathway, and Class I PI3K signaling events may serve as potential explanations for SCI-HO and TBI-HO, and enhance our understanding of its molecular mechanisms. The glucocorticoid receptors signal pathways provide new insights for detecting the different pathogenesis between TBI-HO and SCI-HO. Moreover, the four most promising genes, GDF15, LDLR, CCL2 and CLU identified in this study, may offer new clues for researchers to plan targeted experiment for SCI-TBI-HO.
Funding
This work was funded by the National Natural Science Fund (Nos. 31640045 and 31671246), and the National Key Research and Development Program (2016YFC1101604). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors:
National Natural Science Fund: 31640045 and 31671246.
National Key Research and Development Program: 2016YFC1101604.

Competing Interests
The authors declare that they have no competing interests.

Author Contributions
• Yichong Zhang analyzed the data, conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
• Yuanbo Zhan analyzed the data, conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
• Yuhui Kou performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
• Xiaofeng Yin performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
• Yanhua Wang performed the experiments, prepared figures and/or tables, and approved the final draft.
• Dianying Zhang conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability
The following information was supplied regarding data availability:
We text mined MEDLINE®, BITOLA System, GEO (for gene expression data), Comparative Toxicogenomics Database (drug target genes), and the Human eFP Browser (tissue-specific gene data). The complete text mining results are available in the Supploental Tables.

Supplemental Information
Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.8276#supplemental-information.

REFERENCES
Abdallah BM, Alzahrani AM, Kassem M. 2018. Secreted Clusterin protein inhibits osteoblast differentiation of bone marrow mesenchymal stem cells by suppressing ERK1/2 signaling pathway. Bone 110:221–229 DOI 10.1016/j.bone.2018.02.018.
Awan Z, Denis M, Bailey D, Giaid A, Prat A, Goltzman D, Seidah NG, Genest J. 2011. The LDLR deficient mouse as a model for aortic calcification and quantification by micro-computed tomography. *Atherosclerosis* 219(2):455–462. DOI 10.1016/j.atherosclerosis.2011.08.035.

Banovac K, Williams JM, Patrick LD, Levi A. 2004. Prevention of heterotopic ossification after spinal cord injury with COX-2 selective inhibitor (rofecoxib). *Spinal Cord* 42(12):707–710. DOI 10.1038/sj.sc.3101662.

Banovac K, Williams JM, Patrick LD, Haniff YM. 2001. Prevention of heterotopic ossification after spinal cord injury with indomethacin. *Spinal Cord* 39(7):370–374. DOI 10.1038/sj.sc.3101166.

Betsholtz C, Keller A. 2014. PDGF, pericytes and the pathogenesis of idiopathic basal ganglia calcification (IBGC). *Brain Pathology* 24(4):387–395. DOI 10.1111/bpa.12158.

Convente MR, Wang H, Pignolo RJ, Kaplan FS, Shore EM. 2015. The immunological contribution to heterotopic ossification disorders. *Current Osteoporosis Reports* 13(2):116–124. DOI 10.1007/s11914-015-0258-z.

Cheng X, Zhao SG, Xiao X, Chou KC. 2017. iATC-mHyb: a hybrid multi-label classifier for predicting the classification of anatomical therapeutic chemicals. *Oncotarget* 8(35):58494–58503. DOI 10.18632/oncotarget.17028.

Childs AC, Mehta DJ, Gerner EW. 2003. Polyamine-dependent gene expression. *Cellular and Molecular Life Sciences* 60(7):1394–1406. DOI 10.1007/s00018-003-2332-4.

Cullen N, Perera J. 2009. Heterotopic ossification: pharmacologic options. *Journal of Head Trauma Rehabilitation* 24(1):69–71. DOI 10.1097/HTR.0b013e31819a8fcc.

Davis EL, Salisbury EA, Olmsted-Davis E, Davis AR. 2016. Anaplerotic accumulation of tricarboxylic acid cycle intermediates as well as changes in other key metabolites during heterotopic ossification. *Journal of Cellular Biochemistry* 117(4):1044–1053. DOI 10.1002/jcb.25454.

Di Maggio N, Martella E, Frisantiene A, Resink TJ, Schreiner S, Lucarelli E, Jaquiery C, Schaefer DJ, Martin I, Scherberich A. 2017. Extracellular matrix and α5β1 integrin signaling control the maintenance of bone formation capacity by human adipose-derived stromal cells. *Scientific Reports* 7(1):44398. DOI 10.1038/srep44398.

Docheva D, Popov C, Alberton P, Aszodi A. 2014. Integrin signaling in skeletal development and function. *Birth Defects Research Part C: Embryo Today: Reviews* 102(1):13–36. DOI 10.1002/bdrc.21059.

Debaud C, Salga M, Begot L, Holy X, Chedik M, De l’Escalopier N, Torossian F, Levesque JP, Lataillade JJ, Le Bousse-Kerdilès MC, Genêt F. 2017. Peripheral denervation participates in heterotopic ossification in a spinal cord injury model. *PLOS ONE* 12(8):e0182454. DOI 10.1371/journal.pone.0182454.

Davis AP, Grondin CJ, Johnson RJ, Sciaky D, King BL, McMorran R, Wiegers J, Wiegers TC, Mattingly CJ. 2017. The comparative toxicogenomics database: update 2017. *Nucleic Acids Research* 45(D1):D972–D978. DOI 10.1093/nar/gkw838.

Genêt F, Jourdan C, Schnitzler A, Lautridou C, Guillemot D, Judet T, Poiraudou S, Denormandie P. 2011. Troublesome heterotopic ossification after central nervous system damage: a survey of 570 surgeries. *PLOS ONE* 6(1):e16632. DOI 10.1371/journal.pone.0016632.

Gómar-Toldrá N, Sliwinski M, Dijkers MP. 2014. Physical therapy after spinal cord injury: a systematic review of treatments focused on participation. *Journal of Spinal Cord Medicine* 37(4):371–379. DOI 10.1179/2045772314Y.0000000194.

GarciaDiegocázares D, Aguirresánchez HI, Abarcabuis RF, Kouri JB, Velasquillo C, Ibarra C. 2015. Regulation of α5 and αV integrin expression by GDF-5 and BMP-7 in chondrocyte
differentiation and osteoarthritis. *PLOS ONE* **10**(5):e0127166
DOI 10.1371/journal.pone.0127166.

Hristovski D, Rindflesch T, Peterlin B. 2013. Using literature-based discovery to identify novel therapeutic approaches. *Cardiovascular & Hematological Agents in Medicinal Chemistry* **11**(1):14–24 DOI 10.2174/1871525713110100005.

Henry S, McInnes BT. 2017. Literature based discovery: models, methods, and trends. *Journal of Biomedical Informatics* **74**:20–32 DOI 10.1016/j.jbi.2017.08.011.

Hristovski D, Peterlin B, Mitchell JA, Humphrey SM. 2005. Using literature-based discovery to identify disease candidate genes. *International Journal of Medical Informatics* **74**(2–4):289–298 DOI 10.1016/j.ijmedinf.2004.04.024.

Hristovski D, Friedman C, Rindflesch TC, Peterlin B. 2006. Exploiting semantic relations for literature-based discovery. *AMIA Annual Symposium Proceedings Archive* **2006**:349–353 DOI 10.1016/j.jbi.2017.08.011.

Hwang IK, Yoo KY, Nam YS, Choi JH, Lee IS, Kwon YG, Kang TC, Kim YS, Won MH. 2006. Mineralocorticoid and glucocorticoid receptor expressions in astrocytes and microglia in the gerbil hippocampal CA1 region after ischemic insult. *Neuroscience Research* **54**(4):319–327 DOI 10.1016/j.neures.2005.12.012.

Karč A, Karč A. 2011. Using the BITOLA system to identify candidate genes for Parkinson’s disease. *Bosnian Journal of Basic Medical Sciences* **11**(3):185–189 DOI 10.17305/bjbms.2011.2572.

Kalbasi Anaraki P, Patecki M, Tkachuk S, Kiyan Y, Haller H, Dumler I. 2015. Urokinase receptor mediates osteoclastogenesis via M-CSF release from osteoblasts and the c-Fms/PI3K/Akt/NF-κB pathway in osteoclasts. *Journal of Bone and Mineral Research* **30**(2):379–388 DOI 10.1002/jbmr.2350.

Kan L, Kessler JA. 2011. Animal models of typical heterotopic ossification. *Journal of Biomedicine and Biotechnology* **2011**(263):309287 DOI 10.1155/2011/309287.

Karic A, Terzic R, Karic A, Peterlin B. 2011. Identifying candidate genes for Parkinson’s disease by integrative genomics method. *Biochimia Medica* **21**:174–181 DOI 10.11613/BM.2011.027.

Liu S, Zhang L, Wu Q, Wu Q, Wang T. 2013. Chemokine CCL2 induces apoptosis in cortex following traumatic brain injury. *Journal of Molecular Neuroscience* **51**(3):1021–1029 DOI 10.1007/s12031-013-0091-8.

Luo YJ, Li RB, Ma SY, Lu MY. 2016. Progress on hypoxic-ischemic brain damage associated with CCR2 and CCL2. *Fa Yi Xue Za Zhi* **32**:54–57 DOI 10.3969/j.issn.1004-5619.2016.01.012.

Mosser DM, Edwards JP. 2008. Exploring the full spectrum of macrophage activation. *Nature Reviews Immunology* **8**(12):958–969 DOI 10.1038/nri2448.

Patel RV, Hamanishi ET, Provart NJ. 2016. A human “eFP” browser for generating gene expression anatograms. *PLOS ONE* **11**(3):e0150982 DOI 10.1371/journal.pone.0150982.

Pathan M, Keerthikumar S, Ang CS, Gangoda L, Quek CY, Williamson NA, Mouradov D, Sieber OM, Simpson RJ, Salim A, Bacic A, Hill AF, Stroud DA, Ryan MT, Agbinya JI, Mariadason JM, Burgess AW, Mathivanan S. 2015. FunRich: an open access standalone functional enrichment and interaction network analysis tool. *Proteomics* **15**(15):2597–2601 DOI 10.1002/pmic.201400515.
Reznik JE, Biros E, Lamont AC, Sacher Y, Kibrik O, Milanese S, Gordon S, Galea MP. 2017a. A preliminary investigation on the effect of extracorporeal shock wave therapy as a treatment for neurogenic heterotopic ossification following traumatic brain injury. Part I: effects on pain. *Brain Injury* 31:526–532 DOI 10.1080/02699052.2017.1283059.

Reznik JE, Biros E, Sacher Y, Kibrik O, Milanese S, Gordon S, Galea MP. 2017b. A preliminary investigation on the effect of extracorporeal shock wave therapy as a treatment for neurogenic heterotopic ossification following traumatic brain injury. Part II: effects on function. *Brain Injury* 31(4):533–541 DOI 10.1080/02699052.2017.1283060.

Reznik JE, Biros E, Marshall R, Jelbart M, Milanese S, Gordon S, Galea MP. 2014. Prevalence and risk-factors of neurogenic heterotopic ossification in traumatic spinal cord and traumatic brain injured patients admitted to specialised units in Australia. *Journal of Musculoskeletal and Neuronal Interactions* 14:19–28.

Ruschke K, Hiepen C, Becker J, Knaus P. 2012. BMPs are mediators in tissue crosstalk of the regenerating musculoskeletal system. *Cell and Tissue Research* 347(3):521–544 DOI 10.1007/s00441-011-1283-6.

Rousselet E, Marcinkiewicz J, Kriz J, Zhou A, Hatten ME, Prat A, Seidah NG. 2011. PCSK9 reduces the protein levels of the LDL receptor in mouse brain during development and after ischemic stroke. *Journal of Lipid Research* 52(7):1383–1391 DOI 10.1194/jlr.M014118.

Semple BD, Bye N, Rancan M, Ziebell JM, Morganti-Kossmann MC. 2014. Role of CCL2 (MCP-1) in traumatic brain injury (TBI): evidence from severe TBI patients and CCL2−/− mice. *Journal of Cerebral Blood Flow & Metabolism* 30(4):769–782 DOI 10.1038/jcbfm.2009.240.

Semple BD, Kossmann T, Morganti-Kossmann MC. 2010. Role of chemokines in CNS health and pathology: a focus on the CCL2/CCR2 and CXCL8/CXCR2 networks. *Journal of Cerebral Blood Flow & Metabolism* 30(3):459–473 DOI 10.1038/jcbfm.2009.240.

Shehab D, Elgazzar AH, Collier BD. 2002. Heterotopic ossification. *Journal of Nuclear Medicine* 43:346–353 DOI 10.1007/978-3-319-56782-2_244-2.

Sakellariou VI, Grigoriou E, Mavrogenis AF, Soucacos PN, Papagelopoulos PJ. 2012. Heterotopic ossification following traumatic brain injury and spinal cord injury: insight into the etiology and pathophysiology. *Journal of Musculoskeletal and Neuronal Interactions* 12:230 DOI 10.5435/00124635-200911000-00003.

Sullivan MP, Torres SJ, Mehta S, Ahn J. 2013. Heterotopic ossification after central nervous system trauma: a current review. *Bone & Joint Research* 2(3):51–57 DOI 10.1302/2046-3758.23.2000152.

Swanson DR. 1986. Fish oil, Raynaud’s syndrome, and undiscovered public knowledge. *Perspectives in Biology and Medicine* 30(1):7–18 DOI 10.1353/pbm.1986.0087.

Swanson DR. 1988. Migraine and magnesium: eleven neglected connections. *Perspectives in Biology and Medicine* 31(4):526–557 DOI 10.1353/pbm.1988.0009.

Shekaran A, Shoemaker JT, Kavanaugh TE, Lin AS, LaPlaca MC, Fan Y, Guldberg RE, García AJ. 2014. The effect of conditional inactivation of beta 1 integrins using twist 2 Cre, Osterix Cre and osteocalcin Cre lines on skeletal phenotype. *Bone* 68:131–141 DOI 10.1016/j.bone.2014.08.008.

Strelau J, Schober A, Sullivan A, Schilling L, Unsicker K. 2003. Growth/differentiation factor-15 (GDF-15), a novel member of the TGF-beta superfamily, promotes survival of lesioned mesencephalic dopaminergic neurons in vitro and in vivo and is induced in neurons following cortical lesioning. *Journal of Neural Transmission Supplementa* 65:197–203 DOI 10.1007/978-3-7091-0643-3.
Schindowski K, Von Bohlen, Halbach O, Strelau J, Ridder DA, Herrmann O, Schober A, Schwaninger M, Unsicker K. 2011. Regulation of GDF-15, a distant TGF-β superfamily member, in a mouse model of cerebral ischemia. *Cell and Tissue Research* **343**(2):399–409 DOI 10.1007/s00441-010-1090-5.

Torossian F, Guerton B, Anginot A, Alexander KA, Desterke C, Soave S, Tseng HW, Arouche N, Boutin L, Kulina I, Salga M, Jose B, Pettit AR, Clay D, Rochet N, Vlachos E, Genet G, Debaud C, Denormandie P, Genet F, Sims NA, Banzet S, Levesque JP, Lataillade JJ, Le Bousse-Kerdiles MC. 2017. Macrophage-derived oncostatin M contributes to human and mouse neurogenic heterotopic ossifications. *JCI Insight* **2**(21): doi: 10.1172/jci.insight.96034.

Tokunaga A, Oya T, Ishii Y, Motomura H, Nakamura C, Ishizawa S, Fujimori T, Nabeshima Y, Umezawa A, Kanamori M, Kimura T, Sasahara M. 2008. PDGF receptor beta is a potent regulator of mesenchymal stromal cell function. *Journal of Bone and Mineral Research* **23**(9):1519–1528 DOI 10.1359/jbmr.080409.

Troakes C, Smyth R, Noor F, Maekawa S, Killick R, King A, Al-Sarraj S. 2017. Clusterin expression is upregulated following acute head injury and localizes to astrocytes in old head injury. *Neuropathology* **37**(1):12–24 DOI 10.1111/neup.12320.

Uchiyama T, Kawabata H, Miura Y, Yoshioka S, Iwasa M, Yao H, Sakamoto S, Fujimoto M, Haga H, Kadowaki N, Maekawa T, Takaori-Kondo A. 2015. The role of growth differentiation factor 15 in the pathogenesis of primary myelofibrosis. *Cancer Medicine* **4**(10):1558–1572 DOI 10.1002/cam4.502.

Vanden Bossche L, Vanderstraeten G. 2005. Heterotopic ossification: a review. *Journal of Rehabilitation Medicine* **37**(3):129–136 DOI 10.1080/16501970510027628.

Weeber M, Kors JA, Mons B. 2005. Online tools to support literature-based discovery in the life sciences. *Briefings in Bioinformatics* **6**(3):277–286 DOI 10.1093/bib/6.3.277.

Yan P, Xu J, Li Q, Chen S, Kim GM, Hsu CY, Xu XM. 1999. Glucocorticoid receptor expression in the spinal cord after traumatic injury in adult rats. *Journal of Neuroscience* **19**(21):9355–9363 DOI 10.1523/JNEUROSCI.19-21-09355.1999.

Yildiz N, Ardiç F. 2010. Pathophysiology and etiology of neurogenic heterotopic ossification. *Turkish Journal of Physical Medicine & Rehabilitation/Turkiye Fiziksel Tip ve Rehabilitasyon Dergisi* **56**(2):81–87.

Zhang X, Tamasi J, Lu X, Zhu J, Chen H, Tian X, Lee TC, Threadgill DW, Kream BE, Kang Y, Partridge NC, Qin L. 2011. Epidermal growth factor receptor plays an anabolic role in bone metabolism in vivo. *Journal of Bone and Mineral Research* **26**(5):1022–1034 DOI 10.1002/jbm.b.295.

Zhan Y, Zhang R, Lv H, Song X, Xu X, Chai L, Lv W, Shang Z, Jiang Y, Zhang R. 2014. Prioritization of candidate genes for periodontitis using multiple computational tools. *Journal of Periodontology* **85**(8):1059–1069 DOI 10.1902/jop.2014.130523.

Zhan Y, Zhou S, Li Y, Mu S, Zhang R, Song X, Lin F, Zhang R, Zhang B. 2017. Using the BITOLOA system to identify candidate molecules in the interaction between oral lichen planus and depression. *Behavioural Brain Research* **320**:136–142 DOI 10.1016/j.bbr.2016.11.047.

Zhang J, Zhao Y, Hou X, Chen B, Xiao Z, Han J, Shi C, Liu J, Miao Q, Dai J. 2014. The inhibition effects of insulin on BMP2-induced muscle heterotopic ossification. *Biomaterials* **35**(34):9322–9331 DOI 10.1016/j.biomaterials.2014.07.056.

Zimmermann LL, Martin RM, Girgis F. 2017. Treatment options for posttraumatic epilepsy. *Current Opinion in Neurology* **30**(6):580–586 DOI 10.1097/WCO.0000000000000505.