Journal of Pediatric Surgery xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Journal of Pediatric Surgery

journal homepage: www.elsevier.com/locate/jpedsurg

Gene expression profiling coupled with Connectivity Map database mining reveals potential therapeutic drugs for Hirschsprung disease

Shang-jie Xiao [a,1], Xiao-chun Zhu [a,1], Hua Deng [b], Wei-ping Zhou [b], Wen-yi Yang [c], Li-ke Yuan [a], Jiang-yu Zhang [d], Song Tian [a], Lu Xu [a], Liang Zhang [b,*], Hui-min Xia [e,**]

[a] Department of Neonatal Surgery, Guangdong Women and Children Hospital, Guangzhou, China
[b] Translational Medicine Center, Guangdong Women and Children Hospital, Guangzhou, China
[c] Department of Neonatology, Guangdong Women and Children Hospital, Guangzhou, China
[d] Department of Pathology, Guangdong Women and Children Hospital, Guangzhou, China
[e] Department of Pediatric Surgery, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

ARTICLE INFO

Article history:
Received 25 July 2017
Received in revised form 12 February 2018
Accepted 27 February 2018
Available online xxxx

Key words:
Hirschsprung disease
Neuron development
Therapeutic drugs
Microarray
Connectivity Map

ABSTRACT

Background: Hirschsprung disease (HD) is a congenital intestinal anomaly resulting from a failure to form enteric ganglia in the lower bowel. Surgery is the main therapeutic strategy, although neural stem cell transplantation has recently shown promise. However, HD remains a challenging disorder to treat. Our aim was to identify drugs that could counteract the dysregulated pathways in HD and could thus be potential novel therapies.
Methods: We used microarray analysis to identify genes differentially expressed in ganglionic and aganglionic bowel samples from eight children with HD. The signature of differentially expressed genes was then used as a search query to explore the Connectivity Map (cMAP), a transcriptional expression database that catalogs gene signatures elicited by chemical perturbagens.
Results: We uncovered several dysregulated signaling pathways, and in particular regulation of neuron development, in HD. The cMAP search identified some compounds with the potential to counteract the effects of the dysregulated molecular signature in this disease. One of these, pepstatin A, was recently shown to rescue the migration defects observed in a mouse model of HD, providing strong support for our findings.
Conclusions: This study advances our understanding of the molecular changes in HD and identifies several potential pharmacological interventions. Further testing of the identified compounds is warranted.

Hirschsprung disease (HD) is a congenital anomaly characterized by the absence of ganglion cells in the myenteric and submucosal plexuses of the gut, most commonly affecting the colon. HD affects approximately 1 in 4000 live births and results in intestinal obstruction and severe constipation in affected children [1,2]. Aganglionosis and severe hypoganglionosis can be attributed to an abnormality of the enteric nervous system (ENS), which is the largest and most complex part of the autonomic nervous system [3,4]. The ENS descends from enteric neural crest cells (ENCs) that migrate and colonize the gastrointestinal tract during embryonic development. Once situated within the gut, ENCs proliferate and differentiate into neurons and glial cells, organized as interconnected ganglionic plexuses, which innervate the hindgut. Disruption of this process is responsible for all forms of dysganglionosis in children [5–7].

HD is a multifactorial disorder with high heritability, suggesting a genetic basis [8]. Previous studies have identified several genes, including RET, GDNF, EDN3, SOX10, PHOX2B, and TCF4, with apparent associations to HD [9,10]. For example, in an analysis of the diversity and phenotypic consequences of mutations in RET in HD patients of European/Chinese descent, a 4-fold increase in disease susceptibility was linked to a C/T variant allele (rs2435357) of RET [11]. Despite such noteworthy advances in our understanding of the etiology of HD, much remains to be learned.

To identify additional genes linked to HD, a number of studies have performed gene expression profiling of gut samples from animal models of HD or patients with HD [12–16]. However, these studies have focused only on the dysregulated genes and signaling pathways, and little attention has been paid to identifying potential interventional therapeutic strategies.

Surgery is the main treatment for children suffering from HD; however, removal of the aganglionic segment does not always result in a complete recovery, and functional problems often persist. In fact,

* Correspondence to: L. Zhang, Translational Medicine Center, Guangdong Women and Children Hospital, Guangzhou 511400, China. Tel.: +86 20 39151668.
** Correspondence to: H. Xia, Department of Pediatric Surgery, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou 510623, China. Tel.: +86 20 38076001.
E-mail addresses: zhangliang1999@tsinghua.org.cn (L. Zhang), xia-huimin@foxmail.com (H. Xia).
[1] The first two authors contributed equally to this work.

https://doi.org/10.1016/j.jpedsurg.2018.02.060
0022-3468/

the incidence of various complications, such as postoperative enterocolitis, has been reported to be as high as 30% [17,18]. Treatment of HD by neural stem cell transplantation is a promising approach [19–21]. Nonetheless, it is still in the early stages of development and many questions remain to be answered.

The Connectivity Map (cMAP) (build 02) is a repository of gene expression data arising from treatment of cultured human cells, such as the breast cancer line MCF7, with small molecules. To date, this library contains more than 1.5 million gene expression signatures associated with ~5000 small molecule compounds and ~3000 genetic perturbations. Simple pattern-matching algorithms enable the establishment of functional connections between genes, diseases, and bioactive agents [22,23]. Querying a test gene signature against the signatures in the cMAP database returns a "connectivity score" (ranging from −1 to +1) that reflects the relationship between a query gene signature and a perturbagen. In the case of small molecule drugs, a high positive connectivity score (approaching +1) indicates that the corresponding compound in the cMAP database induces expression of the query signature; conversely, a high negative connectivity score suggests that the corresponding drug may counteract the query signature. Thus, compounds with high negative scores may have therapeutic potential. This tool has previously been used to discover the mechanisms of drug action, as well as to identify potentially novel drugs [24–27].

To better understand the pathophysiology of HD and to identify potential novel therapeutics, we performed whole-genome expression profiling of paired aganglionic and ganglionic bowel segments from eight HD patients. The differential molecular signature obtained between the normal and diseased groups was then used as a query to search the cMAP database. Several compounds with the potential to reverse the HD expression signature were identified. Interestingly, several of these small molecule compounds were shown to regulate neuronal differentiation and muscle contraction in previous studies.

1. Materials and methods

1.1. Patients and clinical samples

The study protocol was approved by the Institutional Review Board of Guangdong Women and Children Hospital and written informed consent was obtained from the guardians of all subjects. We recruited eight children diagnosed with HD (7 boys and 1 girl, aged 1 to 6 months) into our study. Biopsies were taken from resected full-thickness bowel tissues during surgery. HD pathology is characterized by the absence of ganglion cells in the myenteric and submucosal plexuses of the large intestine. Paired samples of ganglionic and aganglionic segments of the colon from each patient underwent histological examination to confirm their status.

1.2. Microarray analysis

Total RNA was extracted from tissue samples with TRIzol Reagent (Invitrogen, Carlsbad, CA), following the manufacturer's instructions. The yield and quality of RNA were assessed by spectrophotometry (NanoDrop 2000; Thermo Fisher Scientific) and formaldehyde agarose gel electrophoresis. cDNA synthesis and cRNA amplification were conducted using the Low-Input Quick AmpWT Labeling Kit (Agilent Technologies, TX). The dual-color approach was employed in this experiment. Paired diseased and normal tissues from HD patients were labeled with Cy3 and a common reference consisting of tissues from eight normal subjects was labeled with Cy5, using the Klenow enzyme labeling method. Labeled cDNAs were combined and hybridized to the microarray (Custom Human GE 8 × 60 K: Agilent, 071342). The slides were washed, scanned, imaged, and the data were extracted using Agilent Feature Extraction software.

After global median normalization, probes with signal intensities >100 across at least 70% of all slides were regarded as "expressed" and retained for further analysis. Differentially expressed genes (DEGs) were screened using Significant Analysis of Microarray software (two class unpaired) with cutoff values set at a fold-change of >2.0 and a q value of <0.05. Average-linkage hierarchical clustering was performed and heatmaps were generated. Gene Ontology pathway analysis was performed to classify DEGs into various categories and biological processes using hypergeometric distribution in the R language software package.

1.3. Quantitative RT-PCR

First-strand cDNA was synthesized from total DNase I-treated RNA using PrimerScript II reverse transcriptase and oligo-dT primers (Takara Biomedical Technology Co., Ltd., Beijing, China) at 42 °C for 1 h. The PCR mixture consisted of cDNA template, SYBR Premix Ex Taq II, ROX Reference Dye II, and forward and reverse primers (Table 1), in a total volume of 20 μL. The cycling conditions were 95 °C for 3 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Glyceraldehyde 3-phosphate dehydrogenase (GADPH) mRNA levels were measured as an internal control. The fold difference in expression was calculated using the $2^{-\Delta\Delta Ct}$ method. All analyses were performed on an ABI 7500 Fast Real-Time PCR system (Applied Biosystems) and the validation assays were repeated three times.

1.4. Immunohistochemistry

Paraffin-embedded colon tissues from eight HD patients were dewaxed, rehydrated, and stained with hematoxylin and eosin. Histological sections were analyzed to ensure the presence of normal colon segments containing ganglion cells and diseased, aganglionic segments in the myenteric plexus. Consecutive serial sections of the same blocks were then cut and immunostained for detection of TLX2 protein. Briefly, sections were incubated overnight with a rabbit antibody to TLX2 (HPA040969, Sigma) in a humidified chamber at 4 °C. The sections were then washed and incubated with a secondary antibody, streptavidin-peroxidase complex, and diaminobenzidine. The sections were counterstained with hematoxylin and imaged.

TLX2 staining was evaluated using a semiquantitative method as described elsewhere [28]. In brief, the average staining intensity of a sample was scored as 0 for negative, 1 for weak, 2 for moderate, and 3 for strong immunoreactivity. The percentage of positively stained cells was scored as 0 for negative, 1 for <1%, 2 for 1%–10%, 3 for 10%–33%, 4 for 33%–66%, and 5 for >66%. The final score of each specimen was calculated as the sum of the intensity and proportion scores. Differences between samples were analyzed using Student's t-test, and a p value <0.05 was considered significant.

1.5. Connectivity Map database mining

To explore potential compounds with molecular signatures that might reverse the transcriptional profiles in HD, we compared the observed gene expression profiles with the cMAP reference database (http://portals.broadinstitute.org/cmap/). Genes differentially expressed between HD and control samples were used as query terms to search the cMAP database. Since an Agilent GE 8 × 60 K array was used in our experiments, whereas the cMAP database was compiled

Table 1
Primer sets used for qRT-PCR.

Primer	Sequence (5′-3′)	Amplicon length (bp)
NEFL-F	GCTCCTATCTGATGTCCAC	169
NEFL-R	GCCTCTTCCTTGTCCTTC	
CDH2-F	ACAGATGTGGACAGGATTG	88
CDH2-R	CAAGGATAAGCAGGATGATG	
TNFRSF12A-F	TCTGAGCCTGACCTTCGT	187
TNFRSF12A-R	TGGATGAATGAATGATGAGTGG	

Please cite this article as: Xiao S, et al, Gene expression profiling coupled with Connectivity Map database mining reveals potential therapeutic drugs for Hirschsprung disea..., J Pediatr Surg (2018), https://doi.org/10.1016/j.jpedsurg.2018.02.060