Evolution of the ErbB gene family and analysis of regulators of *Egfr* expression during development of the rat spinal cord

From the Contents

**Introduction**

**Materials and Methods**

**Results**

**Discussion**

Abstract

**Egfr**, a member of the ErbB gene family, plays a critical role in tissue development and homeostasis, wound healing, and disease. However, expression and regulators of *Egfr* during spinal cord development remain poorly understood. In this study, we investigated ErbB evolution and analyzed co-expression modules, miRNAs, and transcription factors that may regulate *Egfr* expression in rats. We found that ErbB family members formed via *Egfr* duplication in the ancient vertebrates but diverged after speciation of gnathostomes. We identified a module that was co-expressed with *Egfr*, which involved cell proliferation and blood vessel development. We predicted 25 miRNAs and nine transcription factors that may regulate *Egfr* expression. Dual-luciferase reporter assays showed six out of nine transcription factors significantly affected *Egfr* promoter reporter activity. Two of these transcription factors (KLF1 and STAT3) inhibited the *Egfr* promoter reporter, whereas four transcription factors (including FOXA2) activated the *Egfr* promoter reporter. Real-time PCR and immunofluorescence experiments showed high expression of FOXA2 during the embryonic period and FOXA2 was expressed in the floor plate of the spinal cord, suggesting the importance of FOXA2 during embryonic spinal cord development. Considering the importance of *Egfr* in embryonic spinal cord development, wound healing, and disease (specifically in cancer), regulatory elements identified in this study may provide candidate targets for nerve regeneration and disease treatment in the future.

**Key Words:** co-expression; *Egfr*; evolution; FOXA2; gene expression; miRNA; spinal cord; transcription factor

**Introduction**

Epidermal growth factor receptor (EGFR, also known as ErbB1) is a member of the receptor tyrosine kinase (RTK) superfamily. The other three EGFR family members include ErbB2, ErbB3, and ErbB4 (Ayati et al., 2020). EGFR was the first receptor discovered in this family. Until now, at least seven primary ligands can activate EGFR, including epidermal growth factor (EGF), transforming growth factor-α, heparin-binding EGF, β-cellulin, amphiregulin, epiregulin, and epigen (Knudsen et al., 2014). The ErbB family plays a critical role in vital signaling pathways, such as the phosphatidylinositol 3-kinase/Akt, Ras/Raf/MEK/ERK1/2, and phospholipase C pathways, which are essential for cellular proliferation, differentiation, and survival (Schlessinger, 2002). Each EGFR family member contains an EGFr-like domain (a conserved three-loop structure). EGFR consists of an extracellular ligand-binding domain, single transmembrane domain, and cytoplasmic domain. After ligand binding, EGFR can homo- or hetero-dimerize, and the EGFR-ErbB2 dimer forms the most active receptor (Klapper et al., 2000). After dimerization, specific tyrosine residues are phosphorylated and new binding sites are formed for EGFR regulators of downstream signaling pathways.

EGFR plays a crucial role during vertebrate development. In mammals, EGFR improves embryo implantation and placenta development during embryogenesis (Chia et al., 1995). EGFR is involved in the development of many organs, including the brain, spinal cord, heart, bone, and liver (Chen et al., 2016). EGFR can be detected in cells of the central nervous system (CNS) during development, such as neurons, astrocytes, oligodendrocytes, and progenitor cells of the subventricular zone (Seroogy et al., 1994; Kornblum et al., 1997). However, the expression of EGFR is reduced in the adult brain, including in the nigrostriatal system, and the highest levels of expression have been detected during the neonatal stage (Seroogy et al., 1994). In the peripheral nervous system, EGFR is also found in the neurons of dorsal root ganglia (DRG), Schwann cells, and satellite glial cells (Morris et al., 1999). Sibilia et al. (2007) reported that EGFR-null mice displayed serious neural defects. Hence, EGFR participates in the development, differentiation, maintenance, and regeneration of various organs and the nervous system. Understanding the regulation of *Egfr* expression is of great significance for comprehending the development of the nervous system and exploring potential targeted therapies for related diseases and such spinal cord injury.
Gene expression is regulated by several factors, including chromatin modification, transcription factors (TFs), and non-coding RNAs (Kouzarides, 2007; Patel et al., 2014). In addition, accumulating studies showed that activation of G protein-coupled receptors (GPCRs) by agonists (e.g., angiotensin II) could promote EGFR activity via transactivation (Forrester et al., 2016). TFs recognize cis-regulatory regions to correctly initiate gene transcription. MicroRNAs (miRNA, 20-30 NT) regulate gene expression post-translationally by binding to the 3′-untranslated region (3′-UTR) of target miRNAs (Filipovicz et al., 2008). TFs and miRNAs are important regulators of gene expression in eukaryotes. Several TFs (Sp1, EVI, ETF) have been shown to activate Egfr (Kageyama et al., 1988; Mizuguchi et al., 2019; Tsa et al., 2019). Chromatin immunoprecipitation (ChIP) sequences of the discovery of TF binding sites (Mundade et al., 2014), and the majority have been studied in humans and mice. MRNAs, such as miR-34a (Li et al., 2017) and miR-133a (Guo et al., 2018), and long non-coding RNAs, such as lncRNA (Li et al., 2019) and ENCORI (Zhao et al., 2019) have been validated as regulators of Egfr expression in cancers. In our previous sequenced transcription profile of rat spinal cord development (Yang et al., 2017), it was observed that Egfr was highly expressed in the late embryonic period (embryonic day 18, E18d) until postnatal week 1 (P1w) with low expression in adults (postnatal weeks 4–8), suggesting dynamic regulation of Egfr during spinal cord development. However, limited studies have explored potential regulators of Egfr expression in the spinal cord.

In this study, we performed an evolutionary analysis of the ErbB gene family in embryos of pregnant rats (embryo day 11 [E11d], 13 [E13d], and E14d), with Egfr, investigated GPCRs that may be involved in EGFR transactivation, and predicted miRNAs that targeted Egfr and TFs that regulated Egfr expression based on our previous transcription profiling of spinal cord development. We also analyzed the direct binding between a specific TF and Egfr using a cell-based reporter assay in vitro. Finally, we investigated the expression of candidate TFs during spinal cord development at the mRNA and protein level using qRT-PCR and immunohistochemistry, respectively, in rats ranging in age from embryonic to adult.

Materials and Methods

Animals and tissue preparation

The study was approved by the Animal Care and Use Committee of Nantong University and was reported in accordance with the ARRIVE 2.0 guidelines (Animal Research: Reporting of In Vivo Experiments) (Percie du Sert et al., 2020). Specific pathogen-free Sprague-Dawley rats, including pregnant rats (body weights: 300–330 g, n = 2), male pups (postnatal day 1 [P1d] and postnatal week 4 [P4w]; body weights: 300–320 g, respectively), female pups (P4w), group), and adult males (postnatal week 8 [P8w] and P12w; average body weights: 200 g and 320 g, respectively; n = 4/group) were used for sample collection at different developmental stages. All rats were purchased from the Experimental Animal Center of Nantong University, Nantong, China (license No. SCXK [Su] 2019-0001). Four rats were housed per cage, and all rats were allowed access to food and water ad libitum. The animal room temperature was maintained at 23 ± 1°C. Spinal cord tissue samples were collected from embryonic (E11d, E13d, and E14d) and from pups/adults at P1d, P1w, and P8w. The sampling time was 6:30 a.m. The pregnant rats were deeply anesthetized by intraperitoneal injection of 100 mg sodium pentobarbital, 1.5 mM CaCl_2, 6 mM MgSO_4, 0.15 mM KCl, and 2 mM KH_2PO_4. The mice were killed by cervical dislocation, and the muscles, bones, dura mater, arachnoid membranes, pia mater, and large blood vessels were carefully isolated using micro scissors and micro tweezers. For E13d, E18d, P1d, P1w, and P8w rats, total RNA was separately extracted from the spinal cord of E13d, E18d, P1w, and P8w rats using TRIzol (Gibco; Carlsbad, CA, USA) in accordance with the manufacturer’s protocol. Total RNA was used in this study are listed in Table 1.

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) validation

The mRNA expression of Egfr and six candidate TFs (Bhlha15, Ddit3, Foxa2, Klf1, Sox2, and Stat3) in the developing rat spinal cord was examined by qRT-PCR. In total, 12 slices were used for each TF in the qRT-PCR assay. Total RNA was separately extracted from the spinal cord of E13d, E18d, P1w, and P8w rats using TRIzol (Gibco; Carlsbad, CA, USA) in accordance with the manufacturer’s protocol. Egfr expression was analyzed at the additional time points E11d, E14d, E18d, P1d, and P1w. Total RNA (1 μg) was used to synthesize cDNA using the PrimeScript RT reagent kit (Takara Biotechnology CO., Ltd., Dalian, Liaoning Province, China). The qRT-PCR was performed using the “Premix Ex Taq” (Takara). The expression of target genes was normalized to Gadd45a using the ΔCt method (2-ΔΔCt). The primers used in this study are listed in Table 1.
by 4% paraformaldehyde under deep anesthesia via intraperitoneal injections of mixed narcotics as described above. All collected spinal cord tissues were post-fixed in 4% paraformaldehyde at 4°C for 8 hours, dehydrated in sequential 10–20–30% sucrose solutions, and sectioned on a Leica cryostat (Leica, Wetzlar, Germany) into 12-μm-thick sections. Immunofluorescence procedures were performed as follows. The sections were placed in blocking buffer (Beyotime; Shanghai, China), incubated with the primary antibody at 4°C for 12 hours, and then treated with the secondary antibody at room temperature for one hour. Subsequently, the nuclei were counterstained with Hoechst 33342 in PBS (1:2000) for 10 minutes at room temperature. Images were observed using fluorescence microscopy (Leica). To quantify protein expression, the “integrated density” of the target channel (EGFR and FOX2 only) with the same area at each developmental stage was calculated using ImageJ Fiji software (https://imagej.net/software/fiji/). Statistical analysis

No statistical methods were used to predetermine sample sizes; however, our sample sizes were similar to those reported in a previous publication (Zha et al., 2016). Histological analyses were blinded to experimental conditions. Data from the dual-luciferase reporter assays are presented as mean ± SD and were analyzed using one-way analysis of variance (ANOVA) followed by the Dunnett multiple comparison test. Quantitative data from qRT-PCR and immunofluorescence experiments are presented as means ± SEM after analysis using Student’s t-tests. All statistical analyses were performed using GraphPad Prism (v8.3, GraphPad Software Inc., San Diego, CA, USA). P-values < 0.05 were considered statistically significant.

Results

ErbB gene family members are duplicated in vertebrates but diverge in Gnathostomata

When analyzed ErbB family genes in 32 species from Eukaryota (Placozoa, Cnidaria, Protostomia, and Deuterostomia) to understand the evolution of this gene family. The results indicated that ErbB genes occurred in species from Metazoa except for the phyla Placozoa and Cnidaria (Figure 1A). Multiple copies were observed in most species from vertebrate lineages but only in some invertebrate species, such as those in Platyhelminthes and in the earthworm Eisenia andrei (Figure 1A). According to the annotation in the public database, the ErbB gene in invertebrates was Egrf, while Erfgr, Erbb2, Erbb3, and Erbb4 were found in vertebrates except for species from Agnatha which showed the best hit with ErbBb (termed Erfg-like). Phylogenetic analysis of ErbB genes indicated that Erfg was duplicated in the last common ancestor of Gnathostomata and Schmidtea mediterranea and subsequently underwent independent expansion (Figure 1B). Erfg in E. andrei underwent species-specific expansion with a total of 11 copies. Interestingly, although duplication of ErbB genes was observed in vertebrates, divergence of Egrf occurred in Gnathostomata. Five to six ErbB genes in three species belonging to Agnatha were not diverged, but four diverged genes were present in the closest species, i.e., cartilaginous fish (shark), and remaining gnathostomes, such as humans and mice (Figure 1B). ErbB genes in species from Agnatha were clustered together with ErbB3-Erbb4 and divided into two subgroups before speciation (Figure 1B). Six ErbB genes in lampreys were located in six different syntenic blocks. Synteny analysis between ErbB and lampreys showed that blocks containing four ErbB genes in humans had multiple orthologs and cross relationships in lampreys (Figure 1C). These data suggested that the ErbB gene family may have originated via small-scale duplication of ancient Erfg in vertebrates. Gene duplication is an important source of novelty; however, most duplicates are eventually lost, and the remains may be preserved by neofunctionalization and subfunctionalization (Gout and Lynch, 2015). Transcriptomic information provides more detailed information about the potential fate of duplicated genes. Planarian flatworms (Schmidtea mediterranea) have powerful abilities of regeneration after amputation owing to abundant stem cells (neoblasts) throughout the body (Adler et al., 2014). RNA interference (RNAi) knockdown that the expression of Erfg homologues sme-edgr-1–6 in S. mediterranea showed expression divergence in tissues with different copies even in the same group and one copy in each group mainly contributed to the phenotype (Barberan et al., 2016). It was also shown that RNAi specific for sme-edgr-1 and sme-edgr-3 influenced regeneration (Barberan et al., 2016). Nine out of 11 copies of Erfg in E. andrei were consistently up-regulated at 6 hours after injury, suggesting that most expanded Erfg copies were possibly a gene dosage response to injury and regeneration (Shao et al., 2020) (Figure 1D). In vertebrates, time-series transcriptome analysis of sea lamprey embryogenesis showed that most ErbB genes were upregulated in the early (1–2.5 days) period and decreased later (3–5 days) post-fertilization (Bryant et al., 2016) (Figure 1E). Expression divergence of different ErbB genes was observed in our previous rat spinal cord development dataset (Yang et al., 2021) (Figure 1F).
of Egfr and altered expression of Erbb2 (Hynes and Lane, 2005; Tebbutt et al., 2013). In addition to cancer research, Egfr has been reported to be substantially upregulated in proximal nerve segments on days 4 and 7 after sciatic nerve transaction (Gu et al., 2015). In this study, we focused on Egfr and its regulatory elements in the spinal cord. Egfr expression patterns during spinal cord development have not been systematically reported. Our previous RNA-seq data showed that Egfr expression was low in the early embryonic stage (E9d, E11d) and high in the later embryonic and neonatal stages (E18d–P1w). Expression decreased in adolescence and adulthood (P2w–P12w). These data were consistent with the qRT-PCR results for the developmental stages ranging from E11d to P8w (Figure 2A). Next, we investigated Egfr protein expression and cellular localization by immunohistochemistry. Consistent with the observation at the transcriptional level, immunostaining results showed low levels of EGFR protein at E11d and was diffusely distributed within the spinal cord tissue. A relatively strong positive signal was observed in the anterior horn of gray matter (Figure 2B). In the P1d–P8w rats, EGFR protein levels were relatively stable and mostly confined to the ventral horn of the spinal cord with some co-localization with neurons (possibly motoneurons) (Figure 2B).

Figure 2  Spatiotemporal dynamics of EGFR expression during spinal cord development. (A) qRT-PCR analysis of Egfr mRNA expression during spinal cord development. Data are expressed as means ± SEM (n = 3). The relative expression of Egfr from embryonic day 11 (E11d) to postnatal week 12 (P12w). Expression levels were normalized to Gapdh. (B) Immunostaining for EGFR during spinal cord development. Colors depict EGFR-positive cells (red), neural stem cells (Nestin, green), or neuron (NeuN, green), and Hoechst 33342 nuclei (blue). Individual stains (b–d) and merged (e) staining of markers within white boxed regions (a) shown on the left. Staining was performed at different time points during spinal cord development. E11d and E14d: embryo day 11 and 14, respectively; P1d: postnatal day 1; P1w and P8w: postnatal week 1 and 8, respectively. Left (a) and right (b–e) scale bars represent 100 μm. (C) Histogram depicts the relative fluorescence intensity of EGFR. All values are presented as the mean ± SEM (n = 3). *P < 0.05 vs. E11d (Student’s t-test).

Co-expression of Egfr-related genes inferred from rat spinal cord development data

Gene co-expression network analysis aids in identifying gene modules with potential biological functions and has been widely applied in biomedical research. Based on RNA-seq data from our previous rat spinal cord development study, we first constructed co-expression modules using weighted correlation network analysis. We identified a total of 27 modules of which Egfr and Erbb3 were contained in the blue module (Figure 3A).

Integrating co-expression modules and interaction relationships retrieved from the STRING database, we extracted Egfr-neighboring subnetworks and labeled these genes as Egfr-related (Figure 3B) with a total of 67 genes. Gene ontology enrichment analysis of the 67 Egfr-related genes demonstrated that terms related to cell proliferation, blood vessel development, and Erbb signaling pathway were greatly enriched (Figure 3C). Gene expression profiles consisted of 65 genes with fragments per kilobase of exon per million (FPKM) greater than 1.0. The gene profiles showed that the majority of Egfr-related genes were upregulated in later development of the spinal cord (Figure 3D). We also examined the correlation (Pearson’s correlation coefficient r) of these genes with Egfr using our previous time-series transcriptome dataset from post-injury rat spinal cord (Yu et al., 2019). Forty-six out of 65 (71%) genes showed strong negative or positive correlations with Egfr expression (Figure 3D and Additional Table 1; P < 0.05); for example, the correlation for erbB3 was r = 0.73 (P = 2.94 × 10⁻⁶).

GPCR-mediated Egfr transactivation during spinal cord development

GPCRs have been reported to mediate transactivation of Egfr (Elliott et al., 2013; Forrester et al., 2016). GPCR-mediated transactivation of Egfr has been widely studied in cancers; however, data is limited for its role in spinal cord development. We found that five genes (Gpr153, Gpr21, Dear, Crcl, and Gpr2) assigned to G protein-coupled receptor activity (GO:0004930) were clustered together with Egfr in the blue module. Previous studies have reported that some orphan receptors are involved in EGFR transactivation (Wang et al., 2010; Li et al., 2011; Fujiwara et al., 2012; Girgert et al., 2012; Ochiai et al., 2013; Cho-Clark et al., 2014). We determined that the two orphan receptor genes Gpr153 and Gpr21 showed positive correlations with the expression of Egfr (Pearson’s correlations of 0.87 and 0.56, respectively). Gpr153 was mainly expressed in the CNS (Fathi et al., 1998; Ruiz-Opazo et al., 1998). Mechanisms of GPCR-mediated Egfr transactivation include two types: the triple membrane passing signal (TMAPs) and ligand-independent pathways. Our results showed that some important genes involved in the TMAPs pathway were positively related to the expression of Egfr, such as Adam10, Adam41, Mmp16, Mmp17, Mmp23, Akt1, Sos2, Pik3cd, Raf1, and Map2k2 (Additional Table 2). In the ligand-independent pathway, three genes (Src, Fyn, and Lck) belonging to the Src family also showed a positive relationship with the expression of Egfr (Additional Table 2). The roles of these GPCR-related genes that are co-expressed with Egfr in rat spinal cord development require further experimental investigation.

Prediction of miRNAs targeting Egfr

miRNAs are a class of small, regulatory non-coding RNAs that normally bind to the 3′-UTR of target genes and influence gene expression at the post-
transcriptional level (Bushati and Cohen, 2007). There are 59 human and 17 mouse miRNAs recorded in the TarBase database that are experimentally supported. Various miRNAs targeting humans and mouse Egfr have been widely investigated, but only miR-128-3p has been reported in rats. MiRNAs generally negatively regulate gene expression of their targets. Thus, we combined three widely applied prediction tools (PITA, miRanda, and TargetScan) and calculated correlations with Egfr expression to predict candidate miRNAs targeting Egfr in rats. In total, 120 miRNAs were supported by at least two programs of which 15 miRNAs were predicted by three of the programs (Figure 4A). To select miRNAs that may regulate Egfr during spinal cord development, we calculated Pearson's correlation coefficients between 121 miRNAs, including miR-128-3p, and Egfr. Finally, 25 miRNAs showed a significant negative correlation with Egfr expression (Pearson's correlation coefficients were −0.67 to −0.58; Figure 4B and Additional Table 3). Of the 25 candidates, three miRNAs (miR-34a-5p, miR-21-3p, and let-7b-5p) have been reported to target human Egfr and five miRNAs (miR-378a-3p and let-7a, 7b, 7c, and 7e) have been reported to target mouse Egfr according to miRNA-target pairs deposited in Tarbase. Phylogenetic analysis of the let-7 family (three members, n=4) in humans, mice, and rats demonstrated that five members (let-7b, 7c, 7d, 7e, and 7f) are conserved between these three species and, except for let-7e, they are conserved in their 5’ seed regions (nucleotides 2–8) (Figure 4C). These data suggested that let-7 in the rat may also target Egfr. Additionally, we combined multiple publications deposited in PubMed showed that 13 out of 25 miRNAs have been reported to play a role in inhibiting cell proliferation (Wang et al., 2011, 2019b, c; Kim et al., 2012, 2015; Zhao et al., 2013, Xu et al., 2014; Yang et al., 2015, 2019; Lu et al., 2016, 2017; Fang et al., 2017; Wu et al., 2017, Sun et al., 2018) in humans or mice by targeting Egfr or other targets and include miR-330-5p, miR-29c-3p, and miR-34a-5p (Figure 4B).

Potential TFs regulating Egfr in silico and confirmed in vitro
TFs are the key cellular components that control gene expression via cis-regulatory activity. We predicted potential TFs in rats that regulate Egfr expression by combining multiple methods (Figure 5A). A total of nine TFs were identified as candidates (Figure 5B). A dual-lucerase reporter assay was used to determine the targeting relationship between the nine TFs and Egfr in vitro. Six out of nine TFs showed a significant change in Egfr reporter activity in vitro, indicating that they were in the coregulatory network. FOXA2, KLF4, STAT3, and KLF13 were significantly repressed Egfr reporter activity and FOXA2, SOX17, BHLHA15, and DDI3 activated the Egfr reporter (P < 0.05; Figure 5C). These data suggested that these six TFs may be regulators of Egfr in rats. We also used the same methods to predict TFs that regulate mouse Egfr and human EGFR. Five TFs (FOXA2, KLF4, STAT3, SP1, and KLF1) were shared by the three species (Figure 5D). KLF4 and STAT3 were supported by ChIP-seq results for humans deposited in the ChipBase database, and SP1 activation of EGFR has been experimentally validated (Tsic et al., 2019). Finally, 25 miRNAs showed a significant negative correlation with Egfr expression (Pearson's correlation coefficients were −0.67 to −0.58; Figure 4B and Additional Table 3). Of the 25 candidates, three miRNAs (miR-34a-5p, miR-21-3p, and let-7b-5p) have been reported to target human Egfr and five miRNAs (miR-378a-3p and let-7a, 7b, 7c, and 7e) have been reported to target mouse Egfr according to miRNA-target pairs deposited in Tarbase. Phylogenetic analysis of the let-7 family (three members, n=4) in humans, mice, and rats demonstrated that five members (let-7b, 7c, 7d, 7e, and 7f) are conserved between these three species and, except for let-7e, they are conserved in their 5’ seed regions (nucleotides 2–8) (Figure 4C). These data suggested that let-7 in the rat may also target Egfr. Additionally, we combined multiple publications deposited in PubMed showed that 13 out of 25 miRNAs have been reported to play a role in inhibiting cell proliferation (Wang et al., 2011, 2019b, c; Kim et al., 2012, 2015; Zhao et al., 2013, Xu et al., 2014; Yang et al., 2015, 2019; Lu et al., 2016, 2017; Fang et al., 2017; Wu et al., 2017, Sun et al., 2018) in humans or mice by targeting Egfr or other targets and include miR-330-5p, miR-29c-3p, and miR-34a-5p (Figure 4B).

Expression of six candidate TFs during spinal cord development
We investigated whether FOXA2, SOX17, BHLHA15, DDI3, KLF1, and STAT3 were expressed during spinal cord development at key developmental stages using qRT-PCR and immunohistochemistry. According to the expression pattern of Egfr mRNA during spinal cord development (Figure 2A), we selected several time points (E13d, E18d, P1w, and P8w). The qRT-PCR demonstrated that Foxa2, Klf1, Sox17, and Stat3 were relatively highly expressed during embryonic development (Figure 5E), which then gradually decreased, while Ddit3 expression was relatively low at E13d and increased by E18d. Bhlha15 showed relatively high expression in adult rats (P8w). The differences in expression of some TFs and their effects on EGFR promoter-luciferase activity are caused by the detection of pooled mixed miRNA expression levels from different regions of the spinal cord. Immunostaining indicated that FOXA2 was highly expressed at E13d, and expression decreased by E18d and P1w and was undetectable at P8w (Figure 6A and B). Detectable positive signals for the other five TFs are shown. FOXA2 was co-localized within the cell nucleus, and FOXA2-positive cells were found in the floor plate suggesting the importance of FOXA2 in spinal cord development, which deserves further investigation.

Discussion
EGFR belongs to the receptor tyrosine kinase superfamily and plays an essential role in embryonic development and adult tissue homeostasis, including differentiation, growth, cell maintenance, and repair of damaged tissues. Egfr expression has been explored at specific stages of spinal cord development (Romano and Bucci, 2020); however, evaluation of Egfr expression patterns spanning developmental stages from neural tube formation to the adult spine has not been reported. RNA-seq and qPCR results demonstrated a bell-shaped Egfr expression pattern, which peaked near birth. Immunofluorescence staining showed that EGFR protein levels accumulated in and were confined to the ventral horn of the adult spine. EGFR showed inconsistent expression patterns at the mRNA and protein level, but FOXA2 showed consistent expression patterns. Protein abundance is affected by a series of linked processes, including transcription, post-transcriptional modification, translation, localization, and post-translational modification, and protein degradation (Vogel and Marcotte, 2012). FOXA2 is a developmental transcription factor that is specifically expressed in the spine floor plate and is tightly controlled because knockdown or misexpression causes an abnormal floor plate (Bayly et al., 2012). EGFR has a complex structure with an extracellular ligand-binding domain, a single transmembrane domain, and a cytoplasmic domain. It can homo- or hetero-dimerize or form heterotetramers in response to the binding of its ligand.
expression, which contributes to the formation of glial scars (Erschbamer et al., 2017). Understanding the regulators of Egfr expression is important in CNS development, injury, and regeneration. Transactivation of EGFR has been reported to be mediated by an expanding repertoire of GPCRs and promotes cell survival (Grisanti et al., 2017). We identified several factors that may be involved in transactivation. Several TFs that may regulate Egfr expression were identified in this study. FOXA2 opens compacted chromatin for other transcription-related proteins by displacing nucleosomal core histones and participates in embryonic development and regulation of tissue-specific gene expression (Lee et al., 2019). Additional studies showed that FOXA2 was tightly related to spinal cord embryonic development.

This study has limitations. First, functional experiments, such as RNA interference (RNAi) and over-expression, were not conducted in this study to specifically determine the roles of EGFR and other factors in spinal cord injury. Second, the expression of transcription factors screened by dual-luciferase reporter assay was not consistent with EGFR expression during spinal cord development, which may have resulted from pooling expression levels from different regions of spinal cord tissue. Further studies will be required to confirm and determine the regulatory activities of the identified TFs and miRNAs on Egfr expression in vivo during spinal cord development and injury.

In future studies, we plan to determine the expression of these regulators after spinal cord injury and compare expression to that during spinal cord development. Development-specific regulators will be over-expressed in the spinal cord after injury and effects on spinal cord repair will be evaluated.

In conclusion, we elucidated the long-term expression patterns of EGFR during rat spinal cord development and the evolutional complexity and diversity of the Egfr gene family in Metazoa. Initially screened candidate regulators that mediated Egfr expression were identified and require further investigation, including overexpression and RNAi studies as mentioned above. Our study provides new insight into the evolution of the Egfr gene family and regulators in Egfr expression at the transcriptional or post-transcriptional levels based on the long-term profile of spinal cord development. Given the importance of Egfr in the development and regeneration of the nervous system, the regulatory factors identified in this study may provide new candidate targets for CNS regeneration and disease treatment in the future.

Author contributions: Study design and conception: JY; XSG; bioinformatic analyses: JY, YZ, LX; collection of experimental samples: TZ, XDL, WWY, JC, MG; experimental validations: TZ, YZ, LLZ; manuscript drafting: JY, YZ, LX. All authors approved the final version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Editor note: XSG is an Editorial Board member of Neural Regeneration Research. He was blinded from reviewing or making decisions on the manuscript. The article was subject to the journal's standard procedures, with peer review handled independently of this Editorial Board member and their research groups.

Availability of data and materials: RNA-Seq and miRNA-Seq data for different development stages of the rat spinal cord have been deposited in the NCBI database with BioProject accession tag PRJNA505253. All data generated or analyzed during this study are included in this published article and its supplementary information files.

Open access statement: This is an open access journal, and articles are distributed under the terms of the Creative Commons AttributionNonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

Additional files: Additional Table 1: Pearson's correlation with Egfr expression based on spinal cord injury dataset. Additional Table 2: Key genes involved in TMPS pathway and ligand-independent pathway were positively. Additional Table 3: Predicted miRNA targeting Egfr.

References

Adler CE, Seidel CW, Mckinley SA, sanchez Alvarado A (2014) Selective amputation of the pharynx identifies a FoxA-dependent regeneration program in planaria. Elife 3:e02328.

Agarwal V, Bell GW, Nam JW, Bartol DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. Elife 4:e04505.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410.

Ajuot A, Moghim S, Salarinejad S, Safavi M, Pouramin B, Foroumadi A (2020) A review on progression of epidermal growth factor receptor (EGFR) inhibitors as an efficient approach in cancer targeted therapy. Biogeo Chem 99:1033-1039.

Baylor RD, Brown CE, Agarwala S (2012) A novel role for FOXA2 and SHH in organizing midbrain signaling. Dev Biol 369:32-42.

Bryant SA, Herdy JR, Amerycky CT, Smith AJ (2016) Characterization of somatically-eliminated genes during development of the salamander (Polypterus ensis). Mol Biol Evol 33:2337-2344.

Bushkin N, Cohen SM (2007) microRNA functions. Annu Rev Cell Dev Biol 23:175-205.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAlign: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972-1973.

Chen J, Zeng F, Forrester SL, Eguchi S, Zhang MZ, Harris RC (2016) Expression and function of the epidermal growth factor receptor in physiology and disease. Phys Rev 96:1025-1069.

NEURAL REGENERATION RESEARCH | Vol 17 | No. 11 | November 2022 | 2489
## Additional Table 2

Key genes involved in TMPS pathway and ligand-independent pathway were positively correlated with EGFR.

| GeneID | E9d | E11d | E14d | E18d | P1d | P3d | P1w | P2w | P3w | P4w | P8w | P12w | Pearson Correlation Coefficient with EGFR |
|--------|-----|------|------|------|-----|-----|-----|-----|-----|-----|-----|-----|------------------------------------------|
| Adam1a | 4.33| 2.78 | 4.88 | 6.04 | 3.61| 3.38| 3.64 | 2.66 | 2.3  | 1.67| 2.61| 1.04| 0.54                                      |
| Adam4l1| 2.33| 2.26 | 3.14 | 4.61 | 4.1 | 4.42| 3.58 | 3.24 | 3.35 | 3.23| 3.42| 2.74| 0.47                                      |
| Mmp16  | 9.36| 11.77| 13.45| 18.54| 16.72| 14.01| 13.42| 9.78 | 7.46 | 5.11| 6.11| 5.99| 0.78                                      |
| Mmp17  | 1.82| 2.87 | 3.66 | 4.89 | 6.02| 8.43| 6.89 | 5.09 | 3.7  | 4.19| 3.15| 3.02| 0.54                                      |
| Mmp23  | 0.55| 1.47 | 0.57 | 0.26 | 0.35| 0.32| 0.44 | 0.35 | 0.3  | 0.16| 0.13| 0.12| 0.41                                      |
| Akt1   | 48.48| 54.18| 44.42| 48.47| 37.4| 40.96| 33.96| 32.54| 24.56| 24.73| 29.88| 29.26| 0.58                                      |
| Sos2   | 3.41| 4.6  | 4.54 | 5.29 | 4.96| 4.48| 3.83 | 4.02 | 5.33 | 3.81| 1.39| 0.33| 0.33                                      |
| Ptk3cd | 0.7 | 1.58 | 1.75 | 1.6  | 0.81| 0.76| 0.86 | 0.66 | 0.49 | 0.4 | 0.83| 1.44| 0.33                                      |
| Raf1   | 37.19| 39.36| 43.79| 37.06| 35.45| 28.3| 24.82| 23.81| 16.68| 16.12| 16.56| 12.13| 0.48                                      |
| Map2k2 | 33.52| 36.83| 32.8 | 33.97| 33.75| 34.66| 29.06| 22.92| 19.25| 19.79| 29.77| 28.88| 0.5                                       |
| Src    | 7.95| 7.79 | 10.19| 5.49 | 8.19| 8.6 | 7.02 | 5.18 | 4.69 | 3.9 | 4.77| 4.97| 0.27                                      |
| Fyn    | 23.67| 22.22| 49.59| 52.92| 65.71| 69.96| 77.95| 76.94| 59.5 | 39.67| 40.52| 26.04| 0.36                                      |
| Lok    | 0.59| 0.07 | 1.4  | 1.21 | 1.3 | 1.31| 1.31 | 0.39 | 0.23 | 0.08| 0.23| 0.46| 0.43                                      |
| EGFR   | 1.14| 3.18 | 1.29 | 3.67 | 2.07| 2.84| 3   | 1.98 | 1.18 | 0.94| 1.25| 0.79| -                                          |
| miRNA ID     | Pearson's correlation | Expression | Prediction |
|-------------|-----------------------|------------|------------|
| rno-let-7a-5p | -0.62625              | 0.029354   | 6477.333   |
| rno-let-7b-5p | -0.627513             | 0.028932   | 223.667    |
| rno-let-7c-5p | -0.626625             | 0.035705   | 263        |
| rno-let-7d-5p | -0.609466             | 0.035391   | 3816       |
| rno-let-7e-5p | -0.626573             | 0.029246   | 18170.33   |
| rno-let-7f-5p | -0.625673             | 0.029246   | 18170.33   |
| rno-let-7f-5p | -0.590189             | 0.043365   | 10.66667   |
| rno-let-7f-5p | -0.582893             | 0.046689   | 33.66667   |
| rno-let-7f-5p | -0.586028             | 0.038123   | 805.3333   |
| rno-let-7f-5p | -0.584335             | 0.037402   | 76.66667   |
| rno-let-7f-5p | -0.579426             | 0.048353   | 0.666667   |
| rno-let-7f-5p | -0.580258             | 0.047933   | 1          |
| rno-let-7f-5p | -0.666447             | 0.017952   | 5.333333   |
| rno-let-7f-5p | -0.594402             | 0.041523   | 3.666667   |
| rno-let-7f-5p | -0.644676             | 0.023111   | 0          |
| rno-let-7f-5p | -0.580258             | 0.047933   | 1          |
| rno-let-7f-5p | -0.666447             | 0.017952   | 5.333333   |
| rno-let-7f-5p | -0.635433             | 0.037402   | 76.66667   |
| rno-let-7f-5p | -0.665811             | 0.018102   | 16         |
| rno-let-7f-5p | -0.594402             | 0.041523   | 3.666667   |
| rno-let-7f-5p | -0.644676             | 0.023111   | 0          |
| rno-let-7f-5p | -0.580258             | 0.047933   | 1          |
| rno-let-7f-5p | -0.666447             | 0.017952   | 5.333333   |
| rno-let-7f-5p | -0.635433             | 0.037402   | 76.66667   |
| rno-let-7f-5p | -0.665811             | 0.018102   | 16         |
| rno-let-7f-5p | -0.594402             | 0.041523   | 3.666667   |
| rno-let-7f-5p | -0.644676             | 0.023111   | 0          |
| rno-let-7f-5p | -0.580258             | 0.047933   | 1          |
| rno-let-7f-5p | -0.666447             | 0.017952   | 5.333333   |
| rno-let-7f-5p | -0.635433             | 0.037402   | 76.66667   |
| rno-let-7f-5p | -0.665811             | 0.018102   | 16         |