Gene Expression Patterns 9 (2009) 586–594

Contents lists available at ScienceDirect

Gene Expression Patterns

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

Expression of cohesin and condensin genes during zebrafish development supports a non-proliferative role for cohesin

Maren Mönnich, Stephen Banks, Michael Eccles, Emma Dickinson, Julia Horsfield *

Department of Pathology, Dunedin School of Medicine, The University of Otago, P.O. Box 913, Dunedin 9054, New Zealand

ARTICLE INFO

Article history:
Received 29 July 2009
Accepted 25 August 2009
Available online 31 August 2009

Keywords:
Sister chromatid cohesion
Cohesin
Condensin
Rad21
Scc1
Mcd1
Pds5
Smc1a
Smc3
Smc2
Smc4
Zebrafish
Cornelia de Lange syndrome

ABSTRACT

Cohesin and condensin are similar, but distinct multi-subunit protein complexes that have well-described roles in sister chromatid cohesion and chromosome condensation, respectively. Recently it has emerged that cohesin, and proteins that regulate cohesin function have additional developmental roles. To further understand the role of cohesin in development, we analyzed the expression of genes encoding cohesin and condensin subunits in developing zebrafish embryos and juvenile brain. We found that cohesin subunits are expressed in a pattern that is similar (but not quite identical) to the expression of condensin subunits. Cohesin genes *smc1a*, *rad21*, *pds5b* and *smc3* were expressed in the forebrain ventricular zone, the tectum, the mid-hindbrain boundary, the fourth ventricle, branchial arches, the otic vesicle, the eye and faintly in the developing pectoral fins. Condensin genes *smc2* and *smc4* were expressed in the forebrain ventricular zone, the tectum, the mid-hindbrain boundary, the fourth ventricle, branchial arches, eye and pectoral fins. Condensin genes were additionally expressed in the hindbrain proliferative zone, an area in which cohesin genes were not detected. A comparison with *pcna* expression and BrdU incorporation revealed that the expression of cohesins and condensins closely overlap with zones of proliferation. Interestingly, cohesin genes were expressed in non-proliferating cells flanking rhombomere boundaries in the developing brain. In mature brain and eye, cohesin was expressed in both proliferating cells and in broad zones of post-mitotic cells. The distribution of cohesin and condensin mRNAs supports existing evidence for a non-cell cycle role for cohesin in the developing brain.

1. Results and discussion

Cohesin is a large multi-subunit protein complex that is best known for its essential role in linking together replicated sister chromatids until cell division. The core subunits of cohesin are two large ATPases known as *s*tructural *m*aintenance of *c*hromosome (SMC) proteins, Smc1 and Smc3. These form a ring-like complex bridged by Rad21/Scc1/Mcd1 (Haering et al., 2002), which is bound by a fourth subunit known as Scc3 in budding yeast or SA1/SA2 in vertebrates (Losada and Hirano, 2005; Nasmyth, 2005; Nasmyth and Haering, 2005). Additional proteins Scc2 (known as Nipped-B in *Drosophila*, NIPBL in human) and Scc4 (also called Mau-2 in *C. elegans*) are required to load cohesin onto chromosomes (Ciosk et al., 2000), where its cohesion activity is further regulated by interaction with Pds5a/b, Eco1/Ctf7/Esco2, Rad61/Wapl and Sororin (Losada, 2008).

Unexpectedly, recent studies have revealed that cohesin and its regulatory proteins also have key roles in tissue-specific developmental processes, such as axon guidance (Seitan et al., 2006; Zhang et al., 2007) and pruning (Dorsett, 2008; Pauli et al., 2008; Schuldiner et al., 2008), hematopoiesis (Horsfield et al., 2007), gut development and skeletal patterning (Zhang et al., 2007). Mutations in *NIPBL*, *SMC1A* and *SMC3* lead to the human developmental disease Cornelia de Lange Syndrome (CdLS) (Deardorff et al., 2007; Krantz et al., 2004; Liu and Krantz, 2008; Musio et al., 2006; Tonkin et al., 2004). Cohesin-dependent regulation of development is poorly understood, but recent evidence suggests that cohesion proteins may control the transcription of developmental genes (Dorsett, 2007, 2009; Dorsett and Krantz, 2009; Horsfield et al., 2007; McNairn and Gerton, 2008). Evidence that cohesin function controls gene expression in humans was demonstrated by a recent microarray analysis of lymphoblastoid cell lines from CdLS patients compared with normal controls (Liu et al., 2009). This study identified misexpressed genes that were clear predictors of CdLS, demonstrating that loss of normal cohesin function leads to consistent changes in the expression of selected genes. However, it is not clear how cohesin functions in embryogenesis at the time when developmental changes that lead to CdLS occur.

Chromosome condensation is mediated by condensin, a protein complex that is structurally related to cohesin. Condensin contains a core dimer comprising the SMC proteins Smc2 and Smc4, plus three associated subunits (CapG/G2, CapD2/D3 and CapH/H2) (Haering and Nasmyth, 2003; Losada and Hirano, 2005; Nasmyth and Haering, 2005). In vertebrates, two forms of condensin exist:

* Corresponding author. Tel.: +64 3 479 7436; fax: +64 3 479 7136.
E-mail address: julia.horsfield@otago.ac.nz (J. Horsfield).

1567-133X/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.gep.2009.08.004

Table 1
Database identity of cohesin and condensin sequences used in phylogenetic study.

	Smc1a	Smc1b	Smc2	Smc3	Smc4	Rad21	Pds5a	Pds5b
Danio rerio	tr:Q6DRM9	ref:XP_001334257	ref:NP_955836	sp:Q803N2	tr:Q8JGS5	sp:Q7ZW30	sp:A1L1F4	ref:XP_693953
Homo sapiens	tr:Q14683	tr:Q8NDV3	sp:O95347	sp:Q9UQE7	sp:Q9NTJ3	sp:O60216	sp:Q29RF7	sp:Q9NTI5
Mus musculus	tr:Q9CU62	tr:Q920F6	sp:Q8CG48	sp:Q9CW03	sp:Q8CG47	sp:Q61550	sp:Q6A026	sp:Q4VA53
Rattus norvegicus	tr:Q9Z1M9	ref: NP_001123970	ref: NP_001102136	ref: NP_113771	ref: NP_001032262	tr:Q4KLH7	sp:A4L9P7	sp:Q6TRW4
Microtus arvalis	–	–	–	–	tr:Q256U4	–	–	–
Gallus gallus	tr:Q8AWB7	ref:XP_416457	sp:Q90988	tr:Q8AWB8	tr:Q8AWB9	ref: NP_001026121	sp:Q5F3V3	sp:Q5F3U9
Xenopus laevis	tr:O93308	–	sp:P50533	tr:Q8AW91	sp:P50532	sp:O93310	sp:Q4QXM3(A) sp:Q4KLU7(B)	sp:Q498H0(A) sp:Q5U241(B)
Takifugu rubripes	gb:AAC15582	tr:Q802S2	tr:Q802S1	tr:Q802S0	tr:Q8AW94	–	–	–
Oryzias latipes	gb:BAC76893	tr:Q765Q4	–	–	–	tr:Q765Q6	–	–
Salmo salar	–	–	–	–	–	tr:B5X3C3	–	–
Drosophila melongaster	tr:Q9VCD8 (SMC1)	–	tr:Q7KK96	ref: NP_727988	tr:Q9V3A7	tr:O96689	ref: NP_610719 (PDS5)	–
Anopheles gambiae	tr:Q8GU56 (SMC1)	–	tr:Q8I953	tr:Q8I952	tr:Q8I951	ref: XP_313862		–
Culex quinquefasciatus	–	–	–	–	–	–	ref: XP_001849932 (PDS5)	–

Key: tr, trEMBL; sp, SwissProt; ref, NCBI refseq; gb, GenBank.

condensin I, and condensin II. Both forms contain the core Smc2–Smc4 heterodimer (Ono et al., 2003). Studies to date indicate that condensin has a prominent role in the organization of higher-order structure within the compacted chromosome (reviewed in Hudson et al., 2009).

Since cohesin and condensin both have well-characterized roles in chromosome architecture from S phase to mitosis, one might expect the expression patterns of cohesin and condensin genes to overlap substantially, and also to coincide with proliferating cells. However, there is strong evidence to suggest that cohesin (and perhaps condensin) also have roles in gene expression and development (Cobbe et al., 2006; Csankovszki, 2009; Dorsett, 2007, 2009). Although a comparative analysis of the gene expression patterns of cohesin and condensin subunits during animal development would provide useful information about potential developmental roles, this has not yet been conducted. Do these proteins have identical distribution in developing tissues? In this study, we compared the expression pattern of cohesin genes *smc1a*, *smc3*, *rad21*, and *pds5b* with the expression of condensin subunits *smc2* and *smc4* in developing zebrafish embryos. We also compared these expression patterns with zones of proliferation in the zebrafish embryo (as determined by *pcna* expression and BrdU labeling), as expression outside of proliferative tissue could inform alternative functions for the protein products of these genes.

1.1. Sequences, synteny and phylogenetic analysis of zebrafish cohesin and condensin genes

Full sequences of the cohesin genes *rad21*, *smc1a*, *smc3*, *pds5b* and the condensin genes *smc2* and *smc4* were available in the public databases NCBI, GenBank, Swissprot and Ensembl. Two zebrafish versions of *smc1a* were recently identified. NCBI assigns reference sequences NM_001161631 to *smc1a* and NM_212810 to *smc1al* (*smc1a*-like); both are on Chromosome 23. The genes are 95% identical to each other at the protein level and 77% identical at the transcript level. Both proteins are around 89% identical to human SMC1A. All the data shown here are for the *smc1a* paralog represented by Ensembl gene ENSDARG00000058203 (SwissProt Q6DRM9). This sequence aligns with NM_212810, *smc1al*. For simplicity we refer to this gene as *smc1a*. We chose to work with this version because our preliminary functional studies are consistent with this gene encoding a *bona fide* cohesin subunit. Knock-down of *smc1a* (ENSDARG00000058203) led to accumulation of cells arrested in M phase with disorganized chromosomes and a phenotype resembling *rad21* mutants and *smc3* knock-down embryos (Dickinson and Horsfield, unpublished data). We used the sequences indicated in Table 1 to construct phylogenetic trees to confirm the evolutionary relationship between the zebrafish genes and those from other species (Fig. 1).

Phylogenetic analysis strongly supports the assigned identity of the zebrafish cohesin and condensin proteins. Zebrafish Rad21, Pds5a, Smc1a, Smc1b, Smc3, Smc2 and Smc4 clearly group with orthologs from other bony fishes, with the next closest neighbor usually being the *Xenopus laevis* ortholog (Fig. 1). The *Danio rerio* Rad21 groups with the medaka and pufferfish Rad21 orthologs on a separate branch from other vertebrates (Fig. 1A). *Pds5* is a duplicated gene in vertebrates, and zebrafish Pds5a and Pds5b group on the appropriate post-duplication branches nearest to the *Xenopus* versions (Fig. 1B). Other fish Pds5 orthologs have not yet appeared in available databases. The zebrafish Smc1a paralog

Fig. 1. Cladograms depicting the evolutionary relationships between zebrafish cohesin and condensin proteins and orthologs from other species. (A) Rad21 tree. (B) Pds5 tree including both Pds5a and Pds5b. (C) Smc1 tree, including both Smc1a and Smc1b. (D) Smc2 tree. (E) Smc3 tree. (F) Smc4 tree. Key to bootstrap value calculation methods: Maximum likelihood (10,000 puzzling steps): Maximum Parsimony (1000 samples): Neighbor-Joining (PAM matrix, 1000 samples). Tree branch lengths were calculated using Neighbor-Joining (PAM matrix); branch length represents phylogenetic distance. Bar = number of amino acid substitutions per site. All sequences used are confirmed at the transcript or protein level, with exception of those marked with an asterisk (predicted proteins). Accession numbers for the protein sequences used can be found in Table 1. hs, human (*Homo sapiens*); rn, rat (*Rattus norvegicus*); mm, mouse (*Mus musculus*); ma, vole (*Microtus arvalis*); gg, chicken (*Gallus gallus*); xl, frog (*Xenopus laevis*); dr, zebrafish (*Danio rerio*); ol, Medaka (*Oryzias latipes*); fr, pufferfish (*Takifugu rubripes*); ss, salmon (*Salmo salar*); dm, fruitfly (*Drosophila melongaster*); ag, mosquito (*Anopheles gambiae*); cq, mosquito (*Culex quinquefasciatus*).