Gene Section

Review

CXCL12 (chemokine (C-X-C motif) ligand 12)

Giulia Gentile, Maria Guarnaccia, Sebastiano Cavallaro

Functional Genomics Unit, Institute of Neurological Sciences, National Research Council, Catania, Italy

Published in Atlas Database: August 2015
Online updated version: http://AtlasGeneticsOncology.org/Genes/CXCL12ID40219ch10q11.html
DOI: 10.4267/2042/62774
This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 2.0 France Licence.
© 2016 Atlas of Genetics and Cytogenetics in Oncology and Haematology

Abstract

Review on CXCL12, with data on DNA, on the protein encoded, and where the gene is implicated.

Keywords
CXCL12; Lung cancer; Breast cancer; Human Immunodeficiency Virus-type 1 (HIV-1) infection; WHIM Syndrome; Autoimmune Diseases

Identity

Other names: SDF-1, PBSF, SCYB12, TLSF, TPAR1, SDF-1A, SDF-1B, SDF-1a, SDF-1b, IRH
HGNC (Hugo): CXCL12
Location: 10q11.21
Local order: Molecular Location: Chromosome 10; base pairs 44,865,601-44,880,545 reverse strand, in the February 2009 hg19 human assembly (GRCh37).

According to UCSC Genome Browser on Human Feb. 2009 assembly (GRCh37/hg19), genes flanking CXCL12 on 10q11.21, in centromere to telomere direction, are ZNF32 (zinc finger protein 32), HNRNPA3P1 (heterogeneous nuclear ribonucleoprotein A3 pseudogene 1), CXCL12, THEM72 (transmembrane protein 72), RASSF4 (Ras association (RatGDS/AF-6) domain family member 4).

Note
Synonyms: Stromal cell-derived factor 1, Pre-B cell growth-stimulating factor, Intercrine reduced in hepatomas.

DNA/RNA

Description
The CXCL12 gene consists of 4 exons spanning 14.94 kb on the chromosome 10 at band q11.21 (reverse strand) (Figure 1).

Transcription
Alternative splicing results in eight transcript variants (Table 1), as shown in Ensembl database. Among them, six correspond to as many protein isoforms (Yu et al., 2006).

Pseudogene
No known pseudogenes.

Protein

<table>
<thead>
<tr>
<th>UniProtKB Isoform ID</th>
<th>Isoform Name</th>
<th>Length AA</th>
</tr>
</thead>
<tbody>
<tr>
<td>P48065-1</td>
<td>Beta (full-length)</td>
<td>90</td>
</tr>
<tr>
<td>P48065-1-2</td>
<td>Alpha (full-length)</td>
<td>89</td>
</tr>
<tr>
<td>P48065-1-3</td>
<td>Gamma</td>
<td>119</td>
</tr>
<tr>
<td>P48065-1-4</td>
<td>Delta</td>
<td>140</td>
</tr>
<tr>
<td>P48065-1-5</td>
<td>Epsilon</td>
<td>90</td>
</tr>
<tr>
<td>P48065-1-6</td>
<td>Theta</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 1: Data concerning the six protein isoforms of human CXCL12 have been extracted from UniProtKB database.

Description
CXCL12 gene encodes a stromal cell-derived alpha chemokine, also known as SDF1, a member of the intercrine family.
CXCL12 (chemokine (C-X-C motif) ligand 12)


Figure 1
Schematic structure of the four exons of human CXCL12 (exon1: 153bp; exon2: 118bp; exon3: 87bp; exon4: 3191bp).
Two of them are made up not only of coding regions (in green color) but also of non-coding ones (in red color).

<table>
<thead>
<tr>
<th>N°</th>
<th>Ensembl Transcript ID</th>
<th>Length (bp)</th>
<th>Ensembl Protein ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ENST00000374429</td>
<td>3532</td>
<td>ENSP00000363551</td>
</tr>
<tr>
<td>2</td>
<td>ENST00000343575</td>
<td>1857</td>
<td>ENSP00000363913</td>
</tr>
<tr>
<td>3</td>
<td>ENST00000374426</td>
<td>470</td>
<td>ENSP00000363548</td>
</tr>
<tr>
<td>4</td>
<td>ENST000003395794</td>
<td>1052</td>
<td>ENSP00000379140</td>
</tr>
<tr>
<td>5</td>
<td>ENST000003395795</td>
<td>618</td>
<td>ENSP00000379141</td>
</tr>
<tr>
<td>6</td>
<td>ENST000003395793</td>
<td>665</td>
<td>ENSP00000379139</td>
</tr>
<tr>
<td>7</td>
<td>ENST00000496375</td>
<td>803</td>
<td>No protein product</td>
</tr>
<tr>
<td>8</td>
<td>ENST00000486591</td>
<td>556</td>
<td>No protein product</td>
</tr>
</tbody>
</table>

Table 1 Data concerning the eight transcript variants of human CXCL12 have been extracted from Ensembl database.

Figure 2
The figure shows (A) the molecule processing of CXCL12 beta isoform sequence, in which the potential signal peptide is marked in red color while the polypeptide chain in blue color (data obtained from UniProtKB database); (B) the multiple sequence alignment of the six CXCL12 isoforms. For each isoform, after the 88 identical positions there are a certain number of different residues, highlighted in light blue color. (Alignment produced by CLC Sequence Viewer 7.0 from CLCbio - QIAGEN Company, using Clustal Omega multiple alignment program). SDF-1-alpha(3-67) and SDF-1-beta(3-72), processed by post-translational proteolytic cleavage, are marked by yellow and orange boxes respectively (data on both sequences belong to UniProtKB database).
This family is defined by the location of the first two cysteine residues in the sequence, which are separated by one amino acid (C-X-C chemokine) (Hromas, 1997). Six protein isoforms have been identified in human: Alpha, Beta, Gamma, Delta, Epsilon and Theta (Table 2; Figure 2), produced by alternative splicing events (Yu et al., 2006).

In particular Alpha and Beta isoforms, secreted as full-length molecules, undergo post-translational modifications by a proteolytic cleavage, becoming respectively processed forms SDF-1-alpha(3-67) and SDF-1-beta(3-72) (De la Luz Sierra et al., 2004). The Beta isoform has been chosen as the canonical sequence and, together with Alpha isoform, is ubiquitously expressed in liver, pancreas and spleen. The Gamma Isoform is mainly expressed in heart, while the isoforms Delta, Epsilon and Theta are mainly expressed in pancreas. In the developmental stage, the isoform Alpha is ubiquitously expressed in fetal tissues, Beta and Delta isoforms in fetal spleen and liver, while Gamma and Theta isoforms are weakly detected in fetal kidney (Yu et al., 2006).

**Expression**

CXCL12 is widely expressed in a variety of tissue types, such as heart, liver, spleen, kidney, brain, skeletal muscle, endothelium, epithelium, lymphoid organs, stem cells as well as overexpressed in cancer cells (Kryczek et al., 2007; Teicher and Fricker, 2010).

**Localisation**

Extracellular region.

**Function**

The CXCL12 protein functions as a ligand for two seven-transmembrane receptors (7-TMRs). The first one is the chemokine (C-X-C motif) receptor 4 (CXCR4), a monogamous receptor that signals through heterotrimeric G proteins and beta-arrestin; the second one is the chemokine (C-X-C motif) receptor 7 (CXCR7), a non-monogamous receptor that does not activate G-protein-mediated signal transduction but signals only through beta-arrestin (Oberlin et al., 1996; Rajagopal et al., 2010; Sun et al., 2010; Zhu et al., 2012; Sanchez-Martin et al., 2013). In particular, CXCL12 has a higher affinity of binding to CXCR7 than to CXCR4 (Zhu et al., 2012), even if its affinity to CXCR7 seems to be reduced by the expression of CXCR4 at the membrane (Sanchez-Martin et al., 2013). CXCL12 is secreted in the extracellular space as monomeric and dimeric forms, which can trigger different effects on cell signaling (Ray et al., 2012). In fact, whereas CXCR4 binds both monomeric and dimeric forms, CXCR7 binds preferentially the dimeric one (Sanchez-Martin et al., 2013).

For example, dimeric CXCL12 form induces calcium mobilization but fails to promote chemotaxis, which is induced by the monomer-based interactions (Ray et al., 2012; Sanchez-Martin et al., 2013).

Many others cellular functions depend on CXCL12 activity, including embryogenesis, apoptosis and survival, immune response, tissue homeostasis, angiogenesis, calcium ion homeostasis, clathrin-mediated endocytosis, cytoskeletal rearrangement, cell proliferation and migration, tumor growth and metastasis (Vlahakis et al., 2002; Goda et al., 2006; Petit et al., 2007; Khan et al., 2008; Agle et al., 2010; Drury et al., 2010; Karin, 2010; Kremer, 2010; Sun et al., 2010; Zhu et al., 2012).

**Homology**

The CXCL12 Gene Tree shows a great evolutionary conservation across species (Figure 3). The internal nodes of the phylogenetic tree are annotated for duplication (red boxes) and speciation (blue boxes) events, which correspond to paralogs and orthologs homologous genes respectively.

**Mutations**

**Note**

It has been suggested that a single nucleotide polymorphism (SNP) in the 3’ untranslated region of SDF-1-beta transcript is associated not only with a delayed onset (Winkler et al., 1998) and a modest protective effect against infection and progression of AIDS (Modi et al., 2005), but also with the early onset of type 1 diabetes (Dubois-Laforge et al., 2001) and with an increased likelihood of developing several cancers (as lung, breast, colorectal and prostate) (Ma et al., 2012; Shi et al., 2013).

This could be explained by an increased level and stability of SDF-1-beta transcript as effect of this mutation (Garcia-Moruja et al., 2009).

**Somatic**

A SNP consisting of a G-to-A transition at position 801, counting from the ATG start codon in the 3’ UTR of the Genbank reference sequence L36033, was represented in the SDF-1-beta transcript but not in the SDF-1-alpha transcript (Winkler et al., 1998). The mutation CXCL12-G801A-3 Prime UTR, also known as SDF1-3-prime-A, located at chromosome 10 position 44873550 on Assembly GRCh37, has been classified as pathogenic variant (dbSNP: rs387906400) related to the phenotype Human immunodeficiency virus type 1, resistance to (OMIM: 600835.0001).
Figure 3 The CXCL12 Gene Tree shows the maximum likelihood phylogenetic tree representing the evolutionary history of the CXCL12 gene, constructed using the alignment of a CXCL12 representative protein for each species (green bars). The Gene tree has been generated by Ensembl (GeneTree ENSGT00390000014056 - August 2015) using the Gene Orthology/Paralogy prediction method pipeline (Vilella et al., 2009).
Implicated in

Lung cancer

The CXCL12 expression showed an increase in lung cancer cell lines (small cell lung cancers and non-small cell lung cancers) compared to non-malignant human bronchial epithelial cell lines. This overexpression was positively but weakly correlated with those of CXCR4 or CXCR7, suggesting that CXCL12 may differentially interact with its receptors depending on the cellular context (Imai et al., 2010). Furthermore, different evidences support the involvement of the CXCL12/CXCR4 axis, but not CXCL12/CXCR7, in the metastatic behavior of non-small cell lung cancer, suggesting their potential use as prognostic markers and drug targets (Paratore et al., 2011; Cavallaro, 2013; Choi et al., 2014).

Breast cancer

The CXCL12 expression has been shown to stimulate breast cancer cells proliferation and promote tumor growth (Allinen et al., 2004; Duda et al., 2011). Moreover, a CXCL12 gene variant CXCL12-A (CXCL12-G801A, a single nucleotide polymorphism in the 3’ untranslated region) was associated with an increased susceptibility to breast cancer (Dimberg et al., 2007). The interactions of CXCL12/CXCR4 seems to have a critical role in determining the metastatic destination of breast cancer metastasis (Muller et al., 2001; Hinton et al., 2011), while the CXCR7 expression has been linked to the ability of tumor cells to produce lung and brain metastasis (Sun et al., 2010). In a recent study, the expression profiles of the six CXCL12 isoforms and both receptors have been investigated in a large clinical cohort and common breast cancer cell lines. As result, isoform-specific differences in expression and breast cancer outcomes have been established, while CXCR4 and CXCR7 showed an opposite pattern in cancer as compared with normal and further differences between hormone receptor status and molecular subtypes (Zhao et al., 2014).

Other malignancies

The importance of CXCL12 pathways in the proliferation, growth and metastasis processes has been assessed for many other types of tumors, among which prostate cancer (Vaday et. al., 2004; Zhang et al., 2008; Sun et al., 2010; Duda et al., 2011), pancreatic adenocarcinoma (Shen et al., 2013; Wu et al., 2013), neuroblastoma (Zagozdzon et al., 2008; Liberman et al., 2012), glioblastoma (Gatti et al., 2013; Wurth et al., 2014; Yao et al., 2015), colorectal cancer (Brand et al., 2005; Akishima-Fukasawa et al., 2009; Drury et al., 2010), melanoma (Scala et al., 2005; Toyozawa et al., 2012; Mitchell et al., 2014), bladder cancer (Retz et al., 2005; Shen et al., 2013), esophageal cancer (Sasaki et al., 2008; Wang et al., 2009; Tachezy et al., 2013), renal cancer (Pan et al., 2006; Ieran et al., 2014) and ovarian cancer (Popple et al., 2012).

Human Immunodeficiency Virus-type 1 (HIV-1) infection

Since the discovery of the leukocyte-derived seven-transmembrane domain receptor (LESTR) twenty years ago (Loetscher et al., 1994) and of its function of co-receptor (termed fusin) for lymphocyte-tropic HIV-1 strains, the role of this chemokine receptor in modulating cell permissiveness to the infection was delineated in few years (Feng et al., 1996). At the same time, was reported the identification of the SDF-1 human chemokine as the natural ligand for LESTR/fusin protein, so named CXCR4, and its function of infection inhibitor by lymphocyte-tropic HIV-1 strains (Bleul et al., 1996; Oberlin et al., 1996).

It was clarified that the HIV-1 infection requires expression of CD4, as primary receptor, and the CXCR4 as entry co-factor at the target cell surface; the engagement of these receptors by the HIV-1 envelope glycoprotein is essential for membrane fusion and, HIV infection can be prevented by HIV co-receptor antagonists (Davis et al., 1997). The HIV suppressive activity of SDF-1, by inhibition of HIV replication, was explained by a SDF-1 alpha-dependent internalization of the CXCR4 (Amara et al., 1997). Furthermore, as previously described in the mutations section, the SDF-1-beta chemokine gene variant in the homozygous state (SDF1-3’A/3’A), has been correlate to a delay of the AIDS onset (Winkler et al., 1998).

During these years, different CXCR4 antagonists have reached later stage clinical trials but no one is currently underway (Henrich and Kuritzkes, 2013).

WHIM Syndrome

WHIM is an acronym for an immunodeficiency disorder characterized by (w)arts, (h)yogammaglobulinemia, (i)nfections and (m)ylomatosis symptoms. In most cases, it is caused by an inherited mutation affecting the CXCR4 gene (Hernandez et al., 2003; Liu et al., 2012).

Autoimmune Diseases

CXCL12 functions as an anti-inflammatory chemokine during autoimmune inflammatory responses suggesting the use of CXCL12-based therapies for autoimmune inflammatory diseases (Karín, 2010; Villalvilla et al., 2014).

References

Agle KA, Vongsa RA, Dwinell MB. Calcium mobilization triggered by the chemokine CXCL12 regulates migration in
CXCL12 (chemokine (C-X-C motif) ligand 12)


Teicher BA, Fricker SP, CXCL12 (SDF-1)/CXCR4 pathway in cancer. Clin Cancer Res. 2010 Jun 1;16(11):2927-31


Gomperts E, Vlahov D, Phair J, O'Brien SJ. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC) Science. 1998 Jan 16;279(5349):389-93


This article should be referenced as such: