Janyakhantikul et al. Genome Biology 2010, 11(Suppl 1):P20
http://genomebiology.com/2010/11/S1/P20

Genome Biology

POSTER PRESENTATION Open Access

Evolution of haplotypes at *CCL3L1/CCL4L1*

Somwang Janyakhantikul*, Danielle Carpenter, John AL Armour

From Beyond the Genome: The true gene count, human evolution and disease genomics
Boston, MA, USA. 11-13 October 2010

Background

CCL3L1 and *CCL4L1* are chemokine genes, located on chromosome 17q12 (Figure 1). They are copy number variable genes that share 95% sequence identity with their non-copy number variable paralogues *CCL3* and *CCL4* [1]. The copy number (CN) of these genes varies between populations [2] and has been shown to be associated with phenotypes, such as susceptibility to HIV infection [2] and SLE [3]. A *CCL3L1* pseudogene, also known as *CCL3L2,* is present in the *CCL3L1* region. This pseudogene has sequences similar to *CCL3L1* gene, but lacks exon 1 of *CCL3L1*[1]. As a result, its presence might affect copy number (CN) measurement and subsequent interpretations in association studies between *CCL3L1* CN and diseases [4]. The copy number of *CCL3L1/CCL4L1* was measured using paralogue ratio test (PRT) in 270 HapMap samples 192 UK samples and 157 Basques samples [5]. Firstly, we examined the association between the presence of the *CCL3L1* pseudogene and *CCL3L1* CN in the UK samples and HapMap samples by PCR. The pseudogene was found in 52 out of 192 (27.08%) of UK samples. The presence of this pseudogene is strongly associated with higher copy number of *CCL3L1/CCL4L1* ($P<1 \times 10^{-10}$) (Figure 2). The presence of the pseudogene was tested in all HapMap populations (Table 1). SNP genotyping and *CCL3* microsatellite assays were then carried out to define a set of flanking markers that may predict *CCL3L1/CCL4L1* CN in UK and Basque samples. The best combination of 2 flanking SNPs, rs16972085 and rs8064426, can be used - to predict the *CCL3L1/CCL4L1* CN in UK and Basque samples with only 70% accuracy. Although the *CCL3* microsatellite alleles are not associated with *CCL3L1/CCL4L1* copy number, there is extensive allelic diversity in the microsatellite. Finally, to improve the accuracy of *CCL3L1/CCL4L1* CN prediction, the *CCL3L1/CCL4L1* genes were sequenced in 90 CEU samples to identify sequence variants within the copy-variable genes themselves. Analysis of *CCL3L1/CCL41* haplotypes in CEU samples is underway to provide information on evolution of the *CCL3L1/CCL4L1* haplotypes and the relationship between these haplotypes, flanking SNPs and the presence of the *CCL3L1* pseudogene.

Conclusion

The *CCL3L1* pseudogene and the combination of SNPs rs16972085 and rs8064426 are associated with the *CCL3L1/CCL4L1* copy number, but the association is not absolute. However, data on evolution of *CCL3L1/*

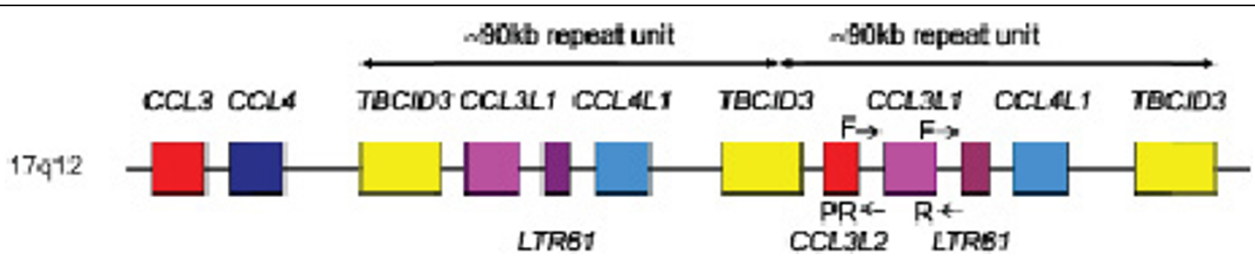

Figure 1 Diagram of *CCL3L1* and *CCL4L1* at Chr.17q12 and the primers used in the *CCL3L1* pseudogene assay (not to scale).

Institute of Genetics, School of Biology, University of Nottingham, NG7 2UH, UK

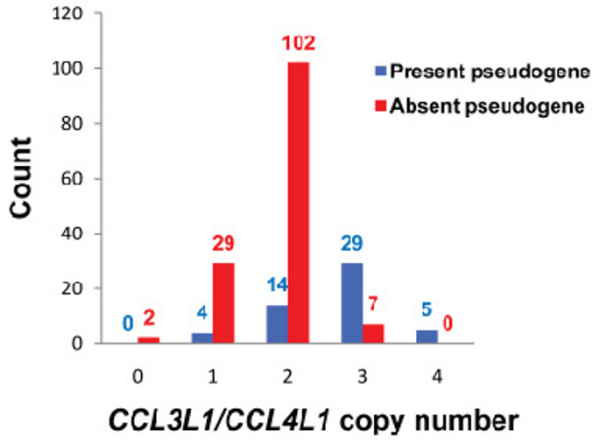

Figure 2 Relationship between presence of the pseudogene and *CCL3L1/CCL4L1* copy number in 192 UK samples.

Table 1 Copy number of *CCL3L1/CCL4L1* and presence of *CCL3L1* pseudogene in HapMap populations

Samples	Population	Number of samples	Range of *CCL3L1/CCL4L1* CN	Mean *CCL3L1/CCL4L1* CN	Frequency of *CCL3L1* pseudogene positive (%)
Yoruba	African	90	2-10	5	58 (64.44%)
CHB-JPT	Asian	90	1-8	4	48 (53.33%)
CEU	European	90	0-3	2	17 (18.89%)

CCL4L1 haplotypes and the relationship between these haplotypes, flanking SNPs and the *CCL3L1* pseudogene in CEU samples could provide variable insight to this region.

Published: 11 October 2010

References

1. Modi WS: **CCL3L1 and CCL4L1 chemokine genes are located in a segmental duplication at chromosome 17q12.** *Genomics* 2004, **83**:735-738.
2. Gonzalez E, *et al*: **The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility.** *Science* 2005, **307**:1434-1440.
3. Mamtani M, *et al*: **CCL3L1 gene-containing segmental duplications and polymorphisms in CCR5 affect risk of systemic lupus erythaematosus.** *Ann Rheum Dis* 2008, **67**:1076-1083.
4. Shrestha S, Tang J, Kaslow RA: **Gene copy number: learning to count past two.** *Nature Med* 2009, **15**:1127-1129.
5. Walker S, Janyakhantikul S, Armour JAL: **Multiplex Paralogue Ratio Tests for accurate measurement of multiallelic CNVs.** *Genomics* 2009, **93**:98-103.

doi:10.1186/gb-2010-11-S1-P20
Cite this article as: Janyakhantikul *et al.*: Evolution of haplotypes at *CCL3L1/CCL4L1*. *Genome Biology* 2010 **11**(Suppl 1):P20.