REVIEW ARTICLE

A review on the Variable copies of C-C motif Chemokine Ligand 3 Like 1 (*CCL3L1*) Gene among Different Populations and Current Methods for Quantification

Jalilah Jamaluddin, Suhaili Abu Bakar

Molecular Biology and Bioinformatics Unit, Department of Biomedical Science, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang Selangor, Malaysia

ABSTRACT

C-C motif Chemokine Ligand 3 Like 1 (*CCL3L1*) is characterized as a gene with copy number variable (CNV) and clustered at the segmental duplication on chromosome 17q12. *CCL3L1* is responsible for the production of macrophage inflammatory protein (MIP) 1α that plays an important function in the immune system and host defense. Various copies range of *CCL3L1* have been reported and associated with the diseases in different populations. Thus, this review aimed to summarise the distribution of *CCL3L1* copy number from different populations according to the geographical region and highlighted the lacking of data from Malaysian population, which is one of the multi-ethnic countries due to the impacts of *CCL3L1* copies on various diseases. Besides, we also outlined the methodologies available for the copy number typing. In overall, this review could provide significant information on the role of *CCL3L1* copies in disease association and as well as providing evidence on the population diversity.

Keywords: C-C motif Chemokine Ligand 3 Like 1; *CCL3L1*; Macrophage inflammatory protein (MIP)-1α; Copy Number Variation; CNV

Corresponding Author:
Suhaili Abu Bakar , PhD
Email: suhaili_ab@upm.edu.my
Tel: +603-89472306

INTRODUCTION

Genome is a complete set of DNA which responsible for all the instructions and information within a living organism (1). The differences of DNA between individuals are termed as the genomic variation (2). Genomic variation includes single nucleotide polymorphism (SNP) known as the most common genetic variation and copy number variation (CNV) formally known as gene duplication (3). Generally, CNV refers to an alteration of genetic sequence including deletion and duplication of DNA segments larger than 1 kb. Genes with CNV includes among others beta-defensin (*DEFB*), alpha-synuclein (*SNCA*), Fc fragment of IgG receptor III (*FCGR3*) and salivary amylase gene (*AMY1*) (4).

CCL3L1 is gene with CNV located on the chromosome 17q12 and encodes for macrophage inflammatory protein (MIP) 1α, a chemokine that involves in a variety of immune cells and plays a part in anti-tumour immunity (5). Different range of *CCL3L1* copy numbers has been observed and seem to be unique to the population and geographical region. Online searches through journal websites, research articles, and database (PubMed, ScienceDirect, SAGE and Database of Genomic Variants websites), shows no reports on the Malaysian population or in disease association studies with *CCL3L1*. Thus, more studies are encouraged to be conducted in order to discover the copy number range in Malaysia so that this information could contribute in minimising the gap in the attempt of mapping the genomic CNVs.

COPY NUMBER VARIATION (CNV)

CNV is defined as gaining or losing of DNA segments when it is compared with the reference genome. Gene dosage imbalance can be considered as the most outcome of CNV by changing the DNA sequences (6). Until November 2018, there are 72 entries of research related to CNV displayed in the online database on the website of Database of Genomic Variants (http://dgv.tcag.ca/dgv/app/home) which represents the growing study of CNVs in current research field.

COPY NUMBER VARIABLE OF *CCL3L1*

Distribution of *CCL3L1* Copy Number (CN) among Populations

Many studies on *CCL3L1* copy number were carried out in several populations, however the number of

investigation among the Asian populations is still minimum. As shown in Fig. 1, a number of *CCL3L1* studies have been carried out in different countries; United Kingdom (7,8), Norway (9), Estonia (10), Germany (11), French (12), Tunisia (12), Tanzania (13), Ethiopia (13), Zimbabwe (14), Peru (15), Botswana (15), South African (15), China (16), Korea (17), Japan (18), and Ohio, Columbia and San Antonio (19). The pattern of *CCL3L1* copy number range could be further analysed based on the geographical regions. CN between zero and seven was reported in the European regions which include United Kingdom, Norway, Estonia, France and Tunisia while Germany records up to 12 copies. Meanwhile, in the African regions comprising Ethiopia, Tanzania, Zimbabwe, South Africa and Botswana, *CCL3L1* copy number ranged from 0 to 14. Studies in the Asian region that includes the Indian, Korean, Chinese and Japanese populations reported copy number ranged between 0 and 10.

CCL3L1 copy number has been proven to be varied in different populations but such information is yet to be reported in Malaysia. From Fig. 1, it could be proposed that the copy number range of *CCL3L1* among Malaysian population of different origins might be from 0 to 10 copies based on the geographical factors of Malaysia which is a part of the Asian region. However, this is only an assumption and requires valid genetic and scientific evidences. Indeed, more investigations on genetic diversity among Malaysians should be performed to delineate our ancestral history. Thus, documentation on the *CCL3L1* copy number among Malaysians will provide valuable fundamental information in the genetic database. Therefore, future investigation on the association of gene copy number towards the susceptibility of gene-related diseases could be determined when compared to the reference distribution.

A recent study discovered that various frequencies of *CCL3L1* were observed among four native Peninsular Malaysians; Malay, Negrito, Senoi and Proto-Malay, with a higher frequency of duplication in Negrito as compared to the other groups (20). However, the exact

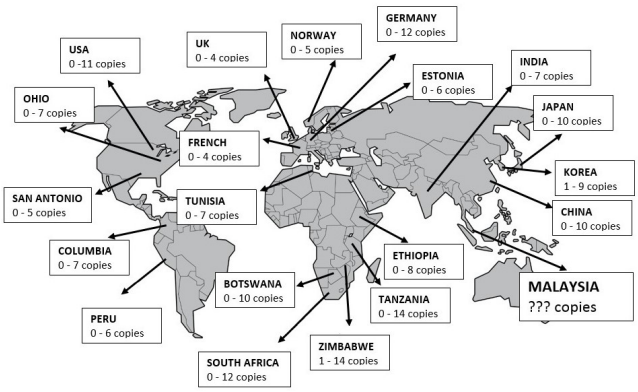

Figure 1: Distribution of *CCL3L1* copy number in different populations.

number of *CCL3L1* copies has not been fully determined in this report. This is mainly due to SNP array that was performed in this study is inadequate to determine the exact copy number of multi-allelic CNVs. Uniquely, Malaysia is a multi-ethnic country which comprises a majority of the Malay, Chinese and Indian ethnics, therefore it is necessary to investigate the possibility of significant differences between ethnic groups. In fact, the findings will aid in further understanding of human genetic variation in a similar population with the multi-ethnic background.

Numerous of studies were performed on the association of *CCL3L1* with disease susceptibilities (Table I). Table I summarized the studies of *CCL3L1* copy number among the diverse population (control group) from the year 2010 until 2016. The range of *CCL3L1* copy numbers varies from one population to another. Discrepancies showed in the disease association of *CCL3L1* copy number suggested further investigations need to be carried out. Different techniques and methodologies were used in the quantification of *CCL3L1* copy numbers whereby qPCR was the most preferred method (11). However, the qPCR method showed declining usage since the year 2011 as other methods such as PRT gives high accuracy and efficiency in the quantification of *CCL3L1* copy number (7).

Functional Impacts of *CCL3L1* Copy Number Variable Gene

CCL3L1 was previously known as *LD78β* and paralogue to *CCL3* (invariant copy number) (21). The term paralogue is defined as homologues derives from the same ancestral gene with similar sequence but reside at different locations within the same genome and have a slightly different function (22). Both *CCL3* and *CCL3L1* are 96% identical in nucleotides and proteins (23). Two isoforms of MIP-1α; LD78α and LD78β which are encoded by these highly related non-allelic genes *CCL3* and *CCL3L1*, respectively (24). MIP-1α is classified under the MIP-1 family which is involved in inflammatory ailments such as asthma, wound healing, and arthritis (25). Thus, it plays a crucial role in pathogen response and cell damages. MIP-1α protein is also a critical macrophage chemoattractant vital in the repair process of cutaneous wound that helps the healing process (26).

LD78β is two-fold more efficient at chemo-attracting human lymphocytes than the LD78α isoform (27). However both LD78α and LD78β products were able to bind to specific chemokine receptors such as CC motif chemokine receptor (CCR) 1, 3 and 5 to activate their functional effects (28). Each binding will mediate a different function. Conversely, LD78β produced by *CCL3L1* was found to be in a truncated-2 form which leads to the higher binding affinity to CCR5 by approximately 6-folds (29).

Molecularly, *CCL3L1* is located at base pairs 250,577

Table I: *CCL3L1* studies in the different populations from the year 2010.

Population/ Cohort	Common range of *CCL3L1*	Method used	Diseases studied	Association of *CCL3L1* copy number with diseases	Number of samples involved	Year of published paper
Europe (Estonian)	0 to 6 copies	qPCR	HIV-1	Fewer copies than two associated with risk of HIV seropositive	374 samples	2010
Europe (Germans)	0 to 12 copies	qPCR	Hepatitis C	Two or fewer copies associated with the risk of Hepatitis C	254 cases, 210 controls	2010
Asia (Japanese)	1 to 9 copies	qPCR	Kawasaki disease	Higher than four copies associated with increased risk of Kawasaki disease	133 cases, 312 controls	2010
South America (Colombian)	0 to 7 copies	qPCR	Tuberculosis	Higher copy number of two associated with risk of active tuberculosis	114 cases, 184 controls	2011
Europe (British)	0 to 4 copies	PRT	Rheumatoid arthritis, Psoriasis and Chron's disease	No association	1581 samples	2011
Asia (Korean)	1 to 5 copies	qPCR	Asthma	Low copy number associated with risk of asthma	161 cases, 372 controls	2011
Asia (Korean)	1 to 5 copies	qPCR	Kawasaki disease	No association	459 cases, 496 controls	2012
Africa (Tanzanian)	1 to 10 copies	PRT	Malaria	Weak evidence of low copy number associated with increased malaria risk	922 samples	2012
Europe (Norwegian)	0 to 5 copies	PRT	Rheumatoid arthritis	No association	905 cases, 905 controls	2012
Sub-Saharan Africa (Ethiopian and Tanzanian)	0 to 8 copies (Ethiopian) 0 to 14 copies (Tanzanian)	PRT	HIV load and immune reconstitution	No association towards viral load, however, low copy number associated with stronger response towards treatment	1134 samples	2013
South America (Peruvian)	0 to 6 copies	PRT	Tuberculosis	No association	621 cases, 511 controls	2014
South Africa (Botswanan)	0 to 10 copies	PRT	Tuberculosis	No association	141 cases, 464 controls	2014
South Africa (Mixed ancestry)	0 to 12 copies	PRT	Tuberculosis	No association	56 cases, 165 controls	2014
Europe (Tunisian)	0 to 7 copies	ddPCR	Rheumatoid arthritis	Higher than two copy number associated with rheumatoid arthritis risk	166 cases, 102 controls	2016
Europe (French)	0 to 4 copies	ddPCR	Rheumatoid arthritis	Higher than two copy number associated with rheumatoid arthritis risk	100 trio families (one patient with two unaffected parents)	2016

to 252,466 on chromosome 17q21.1 (Homo sapiens Annotation Release 108, GRCh38.p7) (30). This gene lies within the copy variable repeat unit (approximately 90 kb in size) together with *CCL4L1, TBC1D3*, and *LTR61* on chromosome 17. Both *CCL3L1* and *CCL4L1* are neighbouring their paralogues, *CCL3* and *CCL4*, respectively (7). The position of *CCL3L1* and *CCL4L1* on the chromosome with their paralogues is illustrated in Fig. 2.

Variable copies of *CCL3L1* may affect the susceptibility of individuals towards immune disease since the protein encoded by this gene, participates in the immunoregulatory and inflammatory process (31). Gonzalez et al. (2005) have demonstrated that low *CCL3L1* copy number may increase the risk of the HIV-1 susceptibility among African populations (32). Different study suggested that an individual is positively susceptible to HIV-1 infection in the case of CN loss

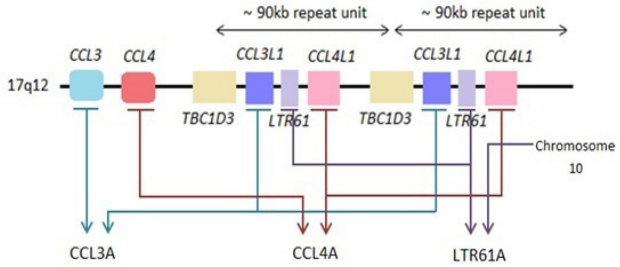

Figure 2: Three systems of PRT in *CCL3L1* copy number quantification; *CCL3A* amplifies *CCL3* and *CCL3L1, CCL4A* amplifies *CCL4* and *CCL4L1* and *LTR61A* amplifies *LTR61* and unlinked reference on chromosome 10.

(<2) in *CCL3L1* (10). Nevertheless, high copy number of CCL3L1 may produce abundant chemokine and suppress the infection activity of HIV-1 as it acts as the competitive inhibitor of the HIV-1 virus (33).

Additionally, Lee and friends have shown the association of one copy or less of *CCL3L1* in promoting asthma attack among Korean populations (34). Protein produced by *CCL3L1* was found to hinder the allergic response of asthma in a dosage depending on the expression level of IL-10. IL-10 acts as an anti-inflammatory cytokine and reduces the level of protein, which may trigger positive responses to the asthma attack. Therefore, a low copy number of *CCL3L1* may increase the risk of asthma in an individual (34).

Nevertheless, Kawasaki disease is linked to the high copy number of *CCL3L1*. Kawasaki disease is an inflammatory disorder involving systemic vasculitis and it is common among young children (35). A study in Japan reported that the risk of Kawasaki disease is correlated higher than four copies of *CCL3L1* due to the elevated inflammatory chemokine production by a high amount of *CCL3L1* copy number, which eventually triggers the antigenic stimulus. Consequently, high *CCL3L1* copy number may increase the susceptibility of developing the disease (35). However, another finding in Korea failed to replicate the results due to the different methodology used in the study (17).

Meanwhile among the Colombians, individuals with more than two copies of *CCL3L1* have an increased risk of tuberculosis activation. A higher copy number of *CCL3L1* will increase the chemokine production that might alter the immune response, resulting the increase activation of tuberculosis (36). Conversely, studies in Tunisia and France showed that more than two copies of *CCL3L1* resulted in significantly higher risk of rheumatoid arthritis (12). When the *CCL3L1* copy number is greater than two, it will increase the chemokines level and irritate the synovium. This irritated synovium may trigger a massive response by leukocytes and chemotactic activity which eventually contribute to rheumatoid arthritis (37). Measuring of *CCL3L1* copy number range towards an association with certain disease requires careful strategy particularly on the methodology approaches and data analysis.

MOLECULAR TECHNIQUES FOR COPY NUMBER TYPING

Discrepancy in reporting the copy number might result from different technologies applied in the studies. A reproducible, reliable and high throughput methods is important in quantification of CNV genes. Thus, several technologies for measuring the CNV genes with their own strengths and limitation has been described here.

Southern Blotting

Southern blotting was named after Sir Edwin Southern, a British biologist who developed this method (38). The procedure of Southern blotting is crucial during the DNA fragmentation as it involves the use of restriction endonuclease enzyme. Gel electrophoresis is then performed to transfer DNA fragments to a nylon or nitrocellulose membrane through the blotting process. For the identification, the labelled probe is hybridized to the target site of DNA fragments and the results are then analysed with an autoradiograph (39). Intensities obtained between the bands are compared to a reference band. High or low intensities ratio indicate the increase or decrease of CNV respectively (40).

There were few studies that applied the Southern blotting analysis and densitometry to quantify the copy number. The radiolabelled probe was specific for *CCL3L1* cDNA to allow hybridization by autoradiographs. The cDNA probe was used in this analysis to ensure the exact breakpoint (41). The analysis of the copy number will show two bands of EcoRI with its own molecular size whereby 4.2 kb reflects CCL3 and 4.8 kb for *CCL3L1*. The ratio is calculated based on the intensities between these two bands and considered as the copy number of *CCL3L1* (37). Unfortunately, Southern blot showed less accuracy in obtaining copy number when compared to the qPCR as the analysis of Southern blot strongly relies on the exposure time, transferring of DNA, and efficiency of hybridization. These factors may lead to inconsistency of the result produced by Southern blot analysis and densitometry. Besides, the huge amount of genomic DNA required to perform a good result makes it less desirable for researchers to use this technique (37).

Fluorescence in situ Hybridization (FISH)

FISH is one of the methods used to quantify CNV which was developed in the late 1980s for human genes mapping (42). Other variants of FISH were developed such as fibre-FISH, spectral karyotyping (SKY) and Multicolour-FISH (M-FISH). Conventional FISH either on metaphase or interphase chromosomes may identify duplication (large >10Mb and small of 40kb) and deletion (more than 40kb) and have been performed as a validation tool (43).

Fiber-FISH was used as a validation tool for quantifying

copy number of *CCL3L1* among the Yoruba from Ibadan (YRI), Nigeria population (13). Stretched DNA from lymphoblastoid cell lines is prepared on the slides. Labelled fosmid probes used in this method are G248P85689G4 (white), G248P84883A8 (green) and G248P8961D8 (red) to determine the absolute copy number of *CCL3L1*. Once the detection succeeds, the slides are mounted with mounting solution containing 4′,6-diamidino-2-phenylindole and fluorescent microscope is used to visualize their chromosomal target site. These probes emit light when viewed under a fluorescent microscope which displays *CCL3* in green, *CCL3L1* in red, and a DNA segment between these two genes in white (13). Indeed, the FISH-based technique was primarily chosen due to its high resolution, accuracy and precision in abnormality detection. However, this technique requires labour-intensive workflow and a high-quality sample. Moreover, the presence of overlapping signals in highly variable regions makes it challenging to interpret the results (40).

Complex Genomic Hybridization (CGH)

CGH is developed to map DNA sequence copy number within the genome in molecular cytogenetic analysis. This technique simultaneously compares two genomic DNA samples labelled with different fluorescent dye. The mixture of tumour labelled with green and normal DNA with red fluorochrome in a 1:1 ratio is hybridised to normal human metaphase spread of chromosomes. It is a fast screening process with no cell culture requirement which offers researchers to analyse the chromosomal changes in solid tumours (44). Later, array CGH (aCGH) was introduced to overcome the limitation of low resolution by conventional CGH. In aCGH, small DNA segments were used as a target instead of targeting metaphase of reference chromosomes (45). Human Whole-Genome TilePath (WGTP) array is an aCGH platform where the ratio of green-to-red fluorescence for each DNA segment arrayed on the microscope slide is determine as the distinction between losses, gains or balanced status. (46). This method is used in the detection of CNV and presented as an additional diagnostic tool. However, there are limitations especially when it comes to detecting a balance alteration of DNA sequence such as translocation and inversion (47).

Quantitative Polymerase Chain Reaction (qPCR)

Polymerase chain reaction (PCR) was firstly discovered by Kary Mullis (1984) (48) and real-time PCR, also known as qPCR was then developed. It is able to measure the accumulation of PCR products by dual-labelled fluorogenic probes (49). qPCR comes with some advantages, such as lesser chance of contamination in the laboratory since it is performed in a closed system (50). Besides, qPCR gives a high yield of samples throughout due to the absence of post-PCR process (51). This method has been extensively used in quantifying the copy number loss or gain of *CCL3L1*. During the amplification process, qPCR detects fluorescence emitted from the probe of *CCL3L1* and the reference gene, β globin. Later, the value of Ct is identified and determined as the initial target amplification. The Ct value is then converted into a standard curve to calculate the ratio of *CCL3L1* to the β globin and is multiplied by two. The final copy number is determined by rounding off the value to the nearest copy number integer (40). However, some qPCR findings on *CCL3L1* copy number were not consistent due to shift of the results to the right when compared to the Paralogue Ratio Test caused by the amplification of *CCL3L1* pseudogene. This indicates the probability in the false-positive calling of *CCL3L1* copy number (52). In addition, the chemistry of qPCR amplification limits the ability to genotype multi-allelic CNV as it gives relatively broader distributions. It is also challenging to design a primer of small size CNV or CNV within a repetitive region and with the necessity to run each sample in triplicate leading to a higher cost (52).

Multiplex Ligation-dependent Probe Amplification (MLPA)

MLPA is a PCR-based method which offers simultaneous detection of duplication and deletions up to 40 target sequences (53) where the amplification takes place on the MLPA probes instead of DNA samples (54). This method was used in screening a large number of genes with the alteration in the copy number and could potentially becomes a gold standard in the molecular diagnosis of genetic-related diseases due to the deletion or duplication (55). Determination of *CCL3L1* copy number by MLPA was successfully performed previously by combining two sets of probes (two colours). Different fluorophore was labelled for each probe, targeting the *CCL3L1* at exon 1 and exon 3 (56). Two probes with known copy number (diploid) were used in the normalization process as reference probes. The targeted genes for the reference probes were EP300 and CREBBP located on chromosome 22 and 16 respectively. The MLPA probes of *CCL3L1* and two additional reference loci were firmly held by hybridization and ligation on the target sequence of DNA (56). Relative peak heights were extracted and the ratios were generated to determine the losses or gains of the target DNA sequence (56). However, it is suggested to perform independent multiple measurements for copy number quantification to accurately distinguish the copy number assigned to the respective samples (56).

Paralogue Ratio Test (PRT)

PRT is a multiplex PCR-based technology where the primers are specifically designed to amplify one from the repeat unit (test) and another from the constant unit (reference) loci simultaneously. PRT has been considered as a much simplified and robust methodology with a high throughput and accurate in measuring the copy number of CNV with minimum of 10ng DNA only (57). Two duplex PCRs were used in quantifying the *CCL3L1* copy number. The first PCR had combined two systems; CCL3A and LTR61A while

second PCR combined *CCL4A* and *LTR61A*. *CCL3A* is used to amplify both *CCL3L1* (test locus) and *CCL3* (reference locus) while, *CCL4A* amplifies *CCL4L1* (test locus) and *CCL4* (reference locus). *LTR61A* serves as an additional system that measures the long terminal repeat (LTR) in between *CCL3L1* and *CCL4L1* against unlinked reference locus on chromosome 10 as illustrated in Fig. 2. Ratios obtained from test to reference loci represent the integer of *CCL3L1* copy number (7).

Droplet Digital Polymerase Chain Reaction

Development of Digital Polymerase Chain Reaction (dPCR) was to improve the conventional PCR by partitioning the DNA samples and dividing the PCR reaction into smaller and multiple reactions (58). Droplet Digital PCR (ddPCR) is the new format of dPCR which has been commercialized in 2011 (59). Fig. 3 shows the workflow of ddPCR where the droplet generation oil and sample are partitioned into thousands of droplets prior PCR. Droplet containing targeted DNA is indicated as a positive droplet. The fluorescence emitted by positive droplets is compared to the reference or negative droplets per fluorophore per sample by using the data analysis software which provides the concentration of target DNA (59).

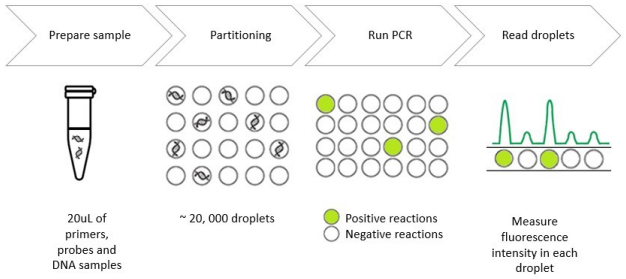

Figure 3: Schematic showing the workflow of ddPCR. Firstly, prepare the sample then followed by partitioning to make approximately 20,000 droplets containing generation oil. Next, run the PCR and read all droplets by measuring the fluorescence intensity. The different in fluorescence intensity is calculated in order to determine the copy number.

Several studies had used this method to determine the *CCL3L1* copy number. Specific probes and primers were designed to amplify the exon 1 of *CCL3L1* at the 120 bp sequence. A fluorescent probe and specific primers were also prepared for RPP30, the reference gene with a known copy number of two for normalization. A software (QuantaSoft) was used to quantify the *CCL3L1* by counting and comparing the signal events of positive and negative fluorescence. The positive fluorescent correlates to droplets with at least one copy of *CCL3L1* and the calculation of copy number relies on the frequency of negative fluorescence. The final copy number values was obtained from the ratio between the concentration (copies/uL) of *CCL3L1* and RPP30 (12). Reliability of this method in obtaining a high-level precision of copy number measurements relies on the multiple reactions of more than 10 000 (approximately 20 000 reactions in ddPCR). ddPCR is applicable in demonstrating a more accurate measurement of CNV than qPCR and PRT (60). Besides, the analyses without the requirement of the standard curve unlike qPCR and Southern blot offers an absolute quantification of copy number. It also reduces error rate in PCR as well as lower equipment cost needed to perform the procedure (59).

Single Nucleotide Polymorphism (SNP) Array

SNP array is a technique developed to genotype SNPs by using microarray platform (61) and have been extended to detect the CNV (62). Affymetrix and Illumina are the primary suppliers of SNP array commercially. Both rely on the same principle of fragmented single-stranded DNA hybridization to arrays even though they use different chemistries (63). The principle of this array is to capture DNA fragments by using short base-pair sequences. In order to infer the copy number, it depends on the hybridization intensities of the test DNA sample that is co-hybridized to a target sample (63). The SNP probe was tagged within or close to the site of the CNV and then the total intensity of normalized allelic was determined to identify the regions of deletion and duplication (64).

Affymetrix Genome-Wide Human SNP 6.0, has been developed with the features of many SNP and CNV probes. These probes efficiently capture SNP and CNV data simultaneously in a single array with just a simple preparation of DNA on the microarray and the data is analysed using Genotyping Console Software. This array manages to differentiate the frequency of *CCL3L1* between Tibetans with other populations (65). Meanwhile, Illumina HumanHap 550v1 and Illumina WG-6 v3 array were used in the determination of *CCL3L1* and *CCL3* mRNA expression (73). Both contain specific probe sets designed for *CCL3L1* (Illumina probe IDs: 1747355, 1773245, and 2218856). The raw data of expression must be first normalized to avoid any bias in RNA loading and subsequently, linear regression is performed to assess the expression data. Later, the effect of copy number on expression could be determined (66). Detection of CNV by using SNP array has become more alluring as there are several new array designs which include monomorphic probes to improve the CNV analysis (67). Besides, the use of a high probe density may increase the sensitivity of SNP array due to more detection power (68). Although SNP array can be used for copy number detection, the array-based approaches like SNP array tend to be inaccurate when involving high copy number counts (more than 4 copies) (40). Thus, SNP array is not applicable for multi-allelic analysis.

CONCLUSION

As the conclusion, *CCL3L1* copy number has been reported in various range among different populations; European with 0 to 7 copies, African with 0 to 14 copies while Asian with 0 to 10 copies. Most of the report also associated the copy number with certain

diseases particularly HIV. However, limited report found on Malaysia which is known as a multi-ethnic country that could enhance our knowledge in human genetic diversity as well as gene association towards ethnic-specific diseases. Moreover, this review gives an insight on the appropriate techniques for typing *CCL3L1* copy number. PRT and ddPCR methods are highly recommended to detect copy number for multi-allelic CNVs as both methods yield good results with high accuracy of copy number reported.

REFERENCES

1. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular biology of the cell. 4th ed. New York: Garland Science; 2002.
2. Griffiths A, Miller J, Suzuk D, Al. E. An Introduction to Genetic Analysis. 7th editio. National Center for Biotechnology Information. New York: W. H. Freeman; 2000.
3. Ku CS, Loy EY, Salim A, Pawitan Y, Chia KS. The discovery of human genetic variations and their use as disease markers: past, present and future. J Hum Genet. 2010;55(7):403–15.
4. Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet. 2009;10:451–81.
5. Kouno J, Nagai H, Nagahata T, Onda M, Yamaguchi H, Adachi K, et al. Up-regulation of CC chemokine, CCL3L1, and receptors, CCR3, CCR5 in human glioblastoma that promotes cell growth. J Neurooncol. 2004;70(3):301–7.
6. Henrichsen CN, Chaignat E, Reymond a. Copy number variants, diseases and gene expression. Hum Mol Genet. 2009;18(R1):R1–8.
7. Walker S, Janyakhantikul S, Armour J a L. Multiplex Paralogue Ratio Tests for accurate measurement of multiallelic CNVs. Genomics. 2009;93(1):98–103.
8. Carpenter D, Walker S, Prescott N, Schalkwijk J, Armour J Al. Accuracy and differential bias in copy number measurement of CCL3L1 in association studies with three auto-immune disorders. BMC Genomics. 2011;12(1):418.
9. Nordang G, Carpenter D, Viken M, Kvien T, Armour J, Lie B. Association analysis of the CCL3L1 copy number locus by paralogue ratio test in Norwegian rheumatoid arthritis patients and healthy controls. Genes Immun. 2012;13(10):579–82.
10. Huik K, Sadam M, Karki T, Avi R, Krispin T, Paap P, et al. CCL3L1 Copy Number Is a Strong Genetic Determinant of HIV Seropositivity in Caucasian Intravenous Drug Users. J Infect Dis. 2010;201(5):730–9.
11. Grunhage F, Nattermann J, Gressner OA, Wasmuth HE, Hellerbrand C, Sauerbruch T, et al. Lower copy numbers of the chemokine CCL3L1 gene in patients with chronic hepatitis C. J Hepatol. 2010;52(2):153–9.
12. Ben Kilani MS, Achour Y, Perea J, Cornelis F, Bardin T, Chaudru V, et al. Characterization of copy number variants for CCL3L1 gene in rheumatoid arthritis for French trio families and Tunisian cases and controls. Clinical Rheumatology. 2016;1–6.
13. Aklillu E, Odenthal-Hesse L, Bowdrey J, Habtewold A, Ngaimisi E, Yimer G, et al. CCL3L1 copy number, HIV load, and immune reconstitution in sub-Saharan Africans. BMC Infect Dis. 2013;13(1):536.
14. Larsen MH, Thɯrner LW, Zinyama R, Amstrup J, Kallestrup P, Gerstoft J, et al. CCL3L gene copy number and survival in an HIV-1 infected Zimbabwean population. Infect Genet Evol. 2012;12(5):1087–93.
15. Carpenter D, Taype C, Goulding J, Levin M, Eley B, Anderson S, et al. CCL3L1 copy number, CCR5 genotype and susceptibility to tuberculosis. BMC Med Genet. 2014;15(1):5.
16. Li H, Xie HY, Zhou L, Feng XW, Wang WL, Liang TB, et al. Copy number variation in CCL3L1 gene is associated with susceptibility to acute rejection in patients after liver transplantation. Clin Transpl. 2012;26(2):314–21.
17. Kim HE, Kim JJ, Han MK, Lee KY, Song MS, Lee HD, et al. Variations in the number of CCL3L1 gene copies and Kawasaki disease in Korean children. Pediatr Cardiol. 2012;33(8):1259–63.
18. Nakajima T, Ohtani H, Naruse T, Shibata H, Mimaya JI, Terunuma H, et al. Copy number variations of CCL3L1 and long-term prognosis of HIV-1 infection in asymptomatic HIV-infected Japanese with hemophilia. Immunogenetics. 2007;59(10):793–8.
19. Mamtani M, Rovin B, Brey R, Camargo JF, Kulkarni H, Herrera M, et al. CCL3L1 gene-containing segmental duplications and polymorphisms in CCR5 affect risk of systemic lupus erythaematosus. Ann Rheum Dis. 2008;67(8):1076–83.
20. Fu R, Mokhtar SS, Phipps ME, Hoh BP, Xu S. A genome-wide characterization of copy number variations in native populations of Peninsular Malaysia. Eur J Hum Genet. 2018;1–12.
21. HUGO Gene Nomenclature Committee. CCL3L1 C-C motif chemokine ligand 3 like 1 NCBI [Internet] 2017 [cited 2017 Apr 24]. Available from: https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=6349
22. Jensen RA. Orthologs and paralogs -we need to get it right. Genome Biol. 2001;2(8):1002–1.
23. Wolpe SD, Davatelis G, Sherry B, Beutler B, Hesse DG, Nguyen HT, et al. Macrophages secrete a novel heparin-binding protein with inflammatory and neutrophil chemokinetic properties. J Exp Med. 1988;167(2):570–81.
24. Carpenter D, McIntosh RS, Pleass RJ, Armour JAL. Functional effects of CCL3L1 copy number. Genes Immun. 2012;13(5):374–9.
25. Murdoch C, Finn A. Chemokine receptors and their role in inflammation and infectious diseases.

Blood. 2000;95(10):3032–43.
26. DiPietro LA, Burdick M, Low QE, Kunkel SL, Strieter RM. Mip-1α as a critical macrophage chemoattractant in murine wound repair. J Clin Invest. 1998;101(8):1693–8.
27. Aquaro S, Menten P, Struyf S, Proost P, Damme J Van, Clercq E De, et al. The LD78beta Isoform of MIP-1alpha Is the Most Potent CC- Chemokine in Inhibiting CCR5-Dependent Human Immunodeficiency Virus Type 1 Replication in Human Macrophages. J Virol. 2001;75(9):4402–6.
28. Struyf S, Menten P, Lenaerts JP, Put W, D'Haese A, De Clercq E, et al. Diverging binding capacities of natural LD78beta isoforms of macrophage inflammatory protein-1alpha to the CC chemokine receptors 1, 3 and 5 affect their anti-HIV-1 activity and chemotactic potencies for neutrophils and eosinophils. Eur J Immunol. 2001;31(7):2170–8.
29. Nibbs RJB, Yang J, Landau NR, Mao JH, Graham GJ. LD78B, A Non-allelic Variant of Human MIP-1a (LD78a), Has Enhanced Receptor Interactions and Potent HIV Suppressive Activity. J Biol Chem. 1999;274(25):17478–83.
30. Modi WS. CCL3L1 and CCL4L1 chemokine genes are located in a segmental duplication at chromosome 17q12. Genomics. 2004;83(4):735–8.
31. Colobran R, Comas D, Faner R, Pedrosa E, Anglada R, Pujol-Borrell R, et al. Population structure in copy number variation and SNPs in the CCL4L chemokine gene. Genes Immun. 2008;9(4):279–88.
32. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. Science. 2005;307(5714):1434–40.
33. Jones KL, Maguire JJ, Davenport AP. Chemokine receptor CCR5: From AIDS to atherosclerosis. British Journal of Pharm. 2011; p. 1453–69.
34. Lee H, Bae S, Choi BW, Choi JC, Yoon Y. Copy number variation of CCL3L1 influences asthma risk by modulating IL-10 expression. Clin Chim Acta. 2011;412(23–24):2100–4.
35. Mamtani M, Matsubara T, Shimizu C, Furukawa S, Akagi T, YOnouchi Y, et al. Association of CCR2-CCR5 haplotypes and CCL3L7 copy number with kawasaki disease, coronary artery lesions, and IVIG responses in Japanese children. PLoS One. 2010;5(7):1–7.
36. Mamtani M, Mummidi S, Ramsuran V, Pham MH, Maldonado R, Begum K, et al. Influence of Variations in CCL3L1 and CCR5 on Tuberculosis in a Northwestern Colombian Population. J Infect Dis. 2011;203(11):1590–4.
37. Townson JR, Barcellos LF, Nibbs RJB. Gene copy number regulates the production of the human chemokine CCL3-L1. Eur J Immunol. 2002;32(10):3016–26.
38. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol. 1975;98(3):503–17.
39. Mellars G, Gomez K. Mutation detection by Southern blotting. Methods Mol Biol. 2011;688:281–91.
40. Cantsilieris S, Baird PN, White SJ. Molecular methods for genotyping complex copy number polymorphisms. Genomics. 2013;101:86–93.
41. White SJ, den Dunnen JT. Copy number variation in the genome; the human DMD gene as an example. Cytogenet Genome Res. 2006;115(3–4):240–6.
42. Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J, et al. Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci U S A. 1988;85(23):9138–42.
43. Lee C, Iafrate AJ, Brothman AR. Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. Nat Genet. 2007;39(7s):S48–54.
44. Weiss MM, Hermsen MAJA, Meijer GA, Van Grieken NCT, Baak JPA, Kuipers EJ, et al. Comparative genomic hybridisation [Internet]. Journal of Clinical Path. 1999;p. 243–51.
45. Hester SD, Reid L, Nowak N, Jones WD, Parker JS, Knudtson K, et al. Comparison of comparative genomic hybridization technologies across microarray platforms. J Biomol Tech. 2009;20(2):135–51.
46. Perry GH, Yang F, Marques-Bonet T, Murphy C, Fitzgerald T, Lee AS, et al. Copy number variation and evolution in humans and chimpanzees. Genome Res. 2008;18(11):1698–710.
47. Shao L, Kang S-HL, Li J, Hixson P, Taylor J, Yatsenko SA, et al. Array comparative genomic hybridization detects chromosomal abnormalities in hematological cancers that are not detected by conventional cytogenetics. J Mol Diagn. 2010;12(5):670–9.
48. Deepak S, Kottapalli K, Rakwal R, Oros G, Rangappa K, Iwahashi H, et al. Real-Time PCR: Revolutionizing Detection and Expression Analysis of Genes. Curr Genomics. 2007;8(4):234–51.
49. Heid C a, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. Genome Res. 1996;6(10):986–94.
50. Ponchel F, Toomes C, Bransfield K, Leong FT, Douglas SH, Field SL, et al. Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. BMC Biotechnol. 2003;3:18.
51. Weaver S, Dube S, Mir A, Qin J, Sun G, Ramakrishnan R, et al. Taking qPCR to a higher level: Analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. Methods. 2010;50(4):271–6.

52. Haridan US, Mokhtar U, Machado LR, Aziz ATA, Shueb RH, Zaid M, et al. A comparison of assays for accurate copy number measurement of the low-affinity FC gamma receptor genes FCGR3A and FCGR3B. PLoS One. 2015;10(1):1–13.
53. Schouten JP, Mcelgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res. 2002;30(12).
54. Os PGCE- Van, Schouten JP. Multiplex Ligation-Dependent Probe Amplification (MLPA®) for the Detection of Copy Number Variation in Genomic Sequences. In Humana Press. 2011;p. 97–126.
55. Stuppia L, Antonucci I, Palka G, Gatta V. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. Int J Mol Sci. 2012;13(3):3245–76.
56. Cantsilieris S, Western PS, Baird PN, White SJ. Technical considerations for genotyping multi-allelic copy number variation (CNV), in regions of segmental duplication. BMC Genomics. 2014;15(1):329.
57. Armour JAL, Palla R, Zeeuwen PLJM, Heijer M Den, Schalkwijk J, Hollox EJ. Accurate, high-throughput typing of copy number variation using paralogue ratios from dispersed repeats. Nucleic Acids Res. 2007;35(3).
58. Vogelstein B, Kinzler KW. Digital PCR. Proc Natl Acad Sci U S A. 1999;96(16):9236–41.
59. Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, et al. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. Anal Chem. 2012;84(2):1003–11.
60. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. Anal Chem. 2011;83(22):8604–10.
61. Wang DG. Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome. Science. 1998;280(5366):1077–82.
62. LaFramboise T. Single nucleotide polymorphism arrays: A decade of biological, computational and technological advances. Nucleic Acids Res. 2009;p. 4181–93.
63. Oldridge DA, Banerjee S, Setlur SR, Sboner A, Demichelis F. Optimizing copy number variation analysis using genome-wide short sequence oligonucleotide arrays. Nucleic Acids Res. 2010;38(10):3275–86.
64. Cooper GM, Zerr T, Kidd JM, Eichler EE, Nickerson DA. Systematic assessment of copy number variant detection via genome-wide SNP genotyping. Nat Genet. 2008;40(10):1199–203.
65. Lou H, Li S, Yang Y, Kang L, Zhang X, Jin W, et al. A Map of Copy Number Variations in Chinese Populations. PLoS One. 2011;6(11):e27341.
66. Urban TJ, Weintrob AC, Fellay J, Colombo S, Shianna K V, Gumbs C, et al. CCL3L1 and HIV/AIDS susceptibility. Nat Med. 2009;15(10):1110–2.
67. Xu L, Hou Y, Bickhart D, Song J, Liu G. Comparative Analysis of CNV Calling Algorithms: Literature Survey and a Case Study Using Bovine High-Density SNP Data. Microarrays. 2013;2(3):171–85.
68. Haraksingh RR, Abyzov A, Urban AE. Comprehensive performance comparison of high-resolution array platforms for genome-wide Copy Number Variation (CNV) analysis in humans. BMC Genomics. 2017;18(1):321.
87. Masood N, Kayani MA. Protection against laryngeal and pharyngeal carcinoma: Heterozygous vs. homozygous deletions of GSTM1 and GSTT1. Genet Mol Biol. 2013;36(1):1–6.