

國立台灣大學醫學院分子醫學研究所

碩士論文

Graduate Institute of Molecular Medicine

College of Medicine

National Taiwan University

master thesis

台灣族群全基因組劑量變異分析：

建立供作臨床與研究應用的數據資料

Whole Genome Analysis of Copy Number Variations in  
Taiwan Populations: A Data Resource for Clinical and  
Research Applications

陳佩芬

Chen, Pei Fen

指導教授：蘇怡寧 博士

Advisor: Su, Yi-Ning , M.D./ Ph.D.

中華民國 99 年 6 月

June, 2010

國立臺灣大學 (碩) 博士學位論文

口試委員會審定書

台灣族群全基因組劑量變異分析

: 建立供作臨床與研究應用的數據資料

Whole Genome Analysis of Copy Number Variations in  
Taiwan Populations: A Data Resource for Clinical and  
Research Applications

本論文係陳佩芬 君 (P97448004) 在國立臺灣大學分子醫學所、所完成之碩士學位論文，於民國 99 年 06 月 03 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

蘇怡寧

(簽名)

(指導教授)

林煥

蘇怡寧

楊偉勳

系主任、所長

李智仁

(簽名)

## 誌謝

兩年前考上台大分醫所，是我生命中非常重要的一年，我在驕傲自己成為大人的同時，也著手完成了自己的終身大事，這兩年來好忙碌，生活像連續快拍的照相機一樣，急忙的捕捉一系列的剎那，卻沒有時間細細品味。終於，在論文即將完成的尾聲，我嘗試放慢自己的腳步，而我此刻的心情和兩年前在台大分醫所榜單上看到自己的名字，一樣的興奮，在同時面對職場工作壓力與課業壓力下完成碩士的學位，這是一場體力、意志力與熱情互相消長及對抗的過程，我終於又為自己的人生立下一個新的里程碑，彌補九年前無法繼續升學的遺憾。這兩年來，首先感謝指導教授蘇怡寧 醫師及博士班前輩林芯仔 醫師的指導，也謝謝實驗室助理久雅、瓊儀及佺瑤傳授我實驗的技巧，更感謝任職公司(生寶臍帶血銀行)長官(章修綱 董事長、陳位存 副總、李政勳 經理、羅瑋瑜 博士)及同事的支持與包容，給我足夠的空間完成我個人的目標。最後要感謝我的內子，在我想發洩的時候當我的出氣筒，在我挫折及難過時當我的垃圾桶，總是默默的陪伴我走過每一次困難，擔任我的最佳心靈捕手；同時又以其 MBA 的經歷，協助、指導我的簡報製作，家人的支持永遠是鼓勵我繼續往前走的力量，衷心的感謝！

## 摘要

基因組劑量變異(Copy Number Variation)是人類遺傳變異的主要來源，它會造成孟德爾遺傳性狀、散發性性狀，可能與某些複雜性疾病相關，但也代表著個體間良性的多型性變異。然而，亞洲族群基因組劑量變異的訊息仍然有限，我們使用 CMDX Bac Array CA2500/ CA3000 來鑑別台灣族群 1015 個個體其 CNV 的分布，總計發現了 438 個發生 CNV 訊號的位置，其中有 6 個區域其 CNV 的發生率超過 20%，而大多數 CNV(92.44%)的變異範圍小於 1 Mb，平均每個個體發生 CNV 的次數為 4.36 次。發生 CNV 的區域中，有 54.79%和片段性重複序列的位置重疊；64.49% 覆蓋基因；92.47% 已在相關的文獻報告過，只有 7.53% 是首次在此篇研究中發現。和先前的文獻相較，Bac Array 所偵測的 CNVs 比高解析度的平台要少得多，但正因為如此，可以降低臨床病理診斷的干擾，因此此篇文獻所使用的 Bac Array 較適合用於產前、臨床診斷。此篇研究建立屬於台灣族群的 CNV 圖譜，補足 CNV 資料庫的族群變異性，可供作臨床遺傳諮詢的參考。

關鍵字：基因組劑量變異、良性變異、非等位同源重組、非同源末端連結、比較式基因組雜交技術

## Abstract

Copy number variation is a source of genetic diversity in humans. CNVs can cause Mendelian or sporadic traits, be associated with complex disease, also represent benign polymorphic variants between individuals. However, CNVs information of Asian populations remains unexplored. We identified CNV distribution of 1015 individuals in Taiwan populations by using CMDX Bac Array CA2500/ CA3000 and find 438 loci with CNV signals. The CNV frequencies of six regions are over 20%. Most CNVs (92.44%) size are shorter than 1 Mb. Average CNV number in each individual is 4.36. Among CNV regions, 54.79% regions overlap with segmental duplication, 64.49% regions cover gene, 92.47% regions are reported in previous study, only 7.53% regions are novel in this study. In contrast to previous studies, CNVs identified are much less than the platform with high resolution. But it lower disturbance in clinical pathologic diagnosis for this reason. So, Bac array used in this study is more appropriate for prenatal, clinical diagnosis. This study establish CNV map of Taiwan populations for further reference of genetic consulting.

Key Word: Copy Number Variation, Heteromorphism, Non-Allelic Homomlogous

Recombination, Non Homologous End Junction, Comparative Genomic Hybridization

# 內 文 目 錄

口試委員會審定書.....	I
誌謝.....	II
中文摘要.....	III
英文摘要.....	IV
內文目錄.....	V
圖目錄.....	VII
表目錄.....	IX
第一章 研究背景與動機.....	1
第一節 顯微鏡下的結構變異.....	1
第二節 次顯微結構變異.....	3
第三節 結構變異的類型.....	4
第四節 基因組劑量變異(CNV)形成的機制.....	9
第五節 基因組劑量變異(CNV)如何影響表型變化.....	10
第六節 基因組劑量變異(CNV)對個體表型及疾病易感性的影響.....	12
第七節 基因組劑量變異(CNV)的臨床意義與重要性.....	13
第八節 偵測基因組劑量變異(CNV)的方法.....	15
第九節 近期文獻對基因組劑量變異(CNV)的研究.....	20

第十節 研究目的與動機.....	24
第二章 研究方法.....	26
第一節 研究樣本來源.....	26
第二節 以 CMDX Bac Array CGH 分析基因組中的劑量變異...	26
第三節 以 Competitive Multi-Plex PCR 驗證 CMDX Bac Array 的基因組劑量變異訊號.....	30
第四節 與現有基因組變異資料庫進行比對.....	32
第三章 結果.....	34
第四章 討論.....	43
第五章 結論.....	50
參考文獻.....	51
附錄：台灣族群 CNV 發生位置與頻率一覽表.....	56



## 圖 目 錄

圖一：CGH 與 aCGH 的操作原理.....	18
圖二：Bac Array CGH 的操作流程.....	29
圖三：人類基因組變異資料庫 <a href="http://projects.tcag.ca/variation">http://projects.tcag.ca/variation</a> 搜尋介 面.....	32
圖四：Bac Clone: CTD-2041D13 (Chr5:69364823-69448911)於資料庫的 搜尋結果.....	33
圖五：樣本 509 的全基因組基因組劑量變異掃描結果圖.....	34
圖六：樣本 313-2 在 Chr 6 的 aCGH 結果 Deletion 表現.....	35
圖七：參考檢體在 6p25.3/ DUSP22 gene 的劑量.....	35
圖八：樣本 313-2 第一次 Multi-Plex PCR 結果.....	35
圖九：樣本 313-2 第二次 Multi-Plex PCR 結果.....	35
圖十：樣本 H268 在 Chr 6 的 aCGH 結果 eletion 表現.....	36
圖十一：參考檢體在 6p25.3/ DUSP22 gene 的劑量.....	36
圖十二：樣本 H268 在 6p25.3/ DUSP22 gene 的 Deletion 表現.....	36
圖十三：樣本 H490 在的 Chr 4 的 aCGH 結果 Duplication 表現.....	36
圖十四：參考檢體在 4p16.1/ CPZ gene 的劑量表現.....	36
圖十五：樣本 H490 在 4p16.1/ CPZ gene 的 Duplication 表現.....	36
圖十六：樣本 321 在的 Chr 12 的 aCGH 結果 Duplication 表現.....	37



圖十七：參考檢體在 12p12.3/ RERG gene 的劑量表現.....	37
圖十八：樣本 321 在 12p12.3/ RERG gene 的 Duplication 表現.....	37
圖十九：參考檢體在 12p12.3/ PIK3C2G gene 的劑量表現.....	37
圖二十：樣本 321 在 12p12.3/ PIK3C2G gene 的 Duplication 表現...	37
圖二十一：台灣族群基因組基因組劑量變異分佈.....	38
圖二十二：台灣族群常見發生基因組劑量變異的位置 ( Incidence>1% ) .....	39
圖二十三：台灣族群全基因組劑量變異形式分析.....	40
圖二十四：台灣族群基因組劑量變異的基因覆蓋率.....	41
圖二十五：台灣族群基因組劑量變異與 Segmental Duplication 的 重疊率.....	41
圖二十六：台灣族群已知 CNV 與新發現的 CNV 比例.....	41
圖二十七：台灣族群基因組劑量變異的大小變化與分佈.....	42
圖二十八：在染色體 9、15、16 所觀察到的 Heteromorphism.....	45
圖二十九：多型性 CNV 的鑑別流程.....	47

## 表 目 錄

表一：分析檢體來源一覽表.....	34
表二：個體的平均基因組劑量變異.....	34
表三：在健康個體所看到的 CNV，部分座於相關遺傳疾病的致病 位置.....	44
表四：研究結果與相關文獻的比較（C 表示中國人；J 表示日本人 ；K 表示韓國人）.....	49



# 第一章 研究背景與動機

## 第一節 顯微鏡下的結構變異

正常人類的基因組的組成為 46, XX 或 46, XY, 基因組數目的改變或結構的異常會導致許多遺傳疾病, 根據統計, 早期自發性流產的胎兒有超過 50% 以上有基因組數目異常; 高齡孕婦所產下的胎兒也有 2% 有基因組數目或結構上的異常 (1), 因此臨床上只要看到非後天造成、非人為外力影響的異常表型, 像是外觀上的畸形、多重結構異常、發育遲緩、智能障礙、反覆流產或不孕等, 進行基因組的數目及結構檢查, 是臨床上進行確診非常重要的一個步驟。

在細胞遺傳學發展的早期, 是透過特殊的化學染色技術, 在顯微鏡下鑑別中期染色體(Metaphase chromosome)的數目及大於 5~8Mb 的結構變化, 並探討這些改變與人類疾病之間的關係, 一般認為大片段的結構變化, 勢必會造成某些異常表型。其中數目上的異常, 即所謂的非整倍體(Aneuploidy), 是臨床上所發現的染色體病中最為常見的一種, 例如: 唐氏症(Down Syndrome: 47, +21)、透納氏症(Turner's Syndrome: 45,X)、柯林菲特氏症(Klinefelter's Syndrome)等, 其起因為生殖細胞於減數分裂的過程中發生了染色體不分離(Chromosome Non-Disjunction)。而染色體結構上的異常則為有絲分裂或減數分裂的過程中, 染色體發生了不平衡重組(Unbalanced Rearrangement), 產生片段缺失(Deletion)、增加(Duplication)、倒位(Inversion)、易位(Translocation), 或功能性基因被打斷、破壞, 因此導致遺傳疾病。例如: 第 5 對染色體短臂末端的缺失所造成的貓哭症(Cat cry syndrome)、第 9 對染色體與第 22 對染色體發生易位(Translocation), 使得第 22 對染色體上產生一個異常的融合基因 (BCR-ABL), 因而引發慢性骨髓性白血病等。

然而, 除了與臨床疾病有關聯性的結構變異, 顯微鏡下的觀察卻也發現許多染色體的結構在不同的健康個體間有著不同的大小差異及染色特性, 這些良性

的變異稱為 Heteromorphism(3)，根據先前文獻的觀察，平均每個個體可發現 4~6 個良性的變異(3)。其可能存在於常染色質區(Euchromatic Region)或異染色質區(Heterochromatic Region)，如常見於 1、9、16 對染色體的著絲粒區(Centromere)或端粒區(Telomere)、Y 染色體長臂的末端，及近端著絲粒染色體等(Acrocentric Chromosome:13~16，21~22 對染色體)。(3, 9,10,11)。而上述這些在顯微鏡下可以觀察到的變異，則稱為“顯微鏡下的結構變異”(Microscopic Structural Variation)。列舉顯微鏡下常見的染色體良性結構變異如下(13)：

### 8p23.1

此區域由  $\beta$  Defensin 基因以 240kb 大小的單位叢聚而成，平均不同個體間約有 2~12 個基因組劑量的差異。此基因與抵抗微生物感染的能力有關。

### 9p12q12

人群中約 6~8% 的個體在第 9 號染色體著絲粒的位置有大片段的結構變異，此區域多是由高度同源性的重複序列所組成，序列的相似度>99%，使得此區域的序列擴增情形有高度的動態變化。

### 15q11.2

15q 的近端其遺傳物質的增加，主要是由包含有 GABRA5、NF1 與 IGVH 等同源性偽基因(Paralogous Pseudogene)的重複序列所組成，絕大多數的個體在此位置有 1~4 個基因組劑量的變化，甚至最多可由 20 個串聯重複序列所組成。

### 16p11.2

和 15q11.2 一樣，此區域是由包含有同源性偽基因的複製序列所組成，如 creatinine transporter SLC68A、Myosin heavy-chain、Immunoglobulin heavy chain 等，不同健康個體間可有 2~12 個基因組劑量的變化。

以上這些 Heteromorphism 皆代表了重複性 DNA 片段極端的擴增，使得正常個體間、基因組劑量變化存在著高度的多型性。而這些 Heteromorphism 有明顯

的分佈偏差，大多好發生於靠近著絲粒的位置，由於著絲粒的位置多由片段性重複序列(Segmental Duplication)整合而成，因此這樣的結構是促使基因組發生新的結構變異的主要來源。同樣的分佈偏差也存在於端粒區域，主要是因為許多染色體的末端包含嗅覺接受體基因(Olfactory Receptor Gene)、Zinc-finger Protein Gene Families、Immunoglobulin Heavy Chain Gene Families 等，這些多型性基因的分佈及劑量變化在不同的個體甚至是不同的族群間，都呈現出高度的變異性及多型性，因此造就了個體間表型多型性的樣貌，也使得端粒區在顯微鏡下經常有不同大小的染色特性。

## 第二節 次顯微結構變異

隨著分子生物學的進步，DNA 定序方法的誕生使得小於 1 Kb 的 DNA 片段可以被偵測到，這幫助我們瞭解許多更小的 Insertion、Deletion、Inversion、Duplication 與臨床疾病之間的關係，以及 Microsatellite、Minisatellite 等重複性序列(Repetitive Sequence)的存在與所扮演的角色。當然，DNA 定序方法甚至可以細微到單一核苷酸的變化，這讓我們可以確診單點突變所引起的遺傳疾病，而 DNA 定序方法的發展同時也讓我們發現單核苷酸多型性(Single Nucleotide Polymorphism, SNP)的存在，SNP 點位上不同的核苷酸，可能會改變蛋白質的表現、影響基因的調控，甚至與疾病的發生有關。這類結構變異在基因組分佈普遍且密度大，早期認為，這些小片段的結構變異與 SNP 是基因組中最主要的變異來源，是造成基因組多樣性(Genomic Diversity)、使個體間有表型差異及疾病易感性(Disease Susceptibility)的主要原因(3)。例如：依美國 NCBI dbSNP 資料庫(<http://www.ncbi.nlm.nih.gov/projects/SNP/>)之統計，人類參考基因體，大約有一千四百多萬個位置可能會有核苷酸多型性，意即平均大約每二百到三百個核苷酸，就會發現有一個單核苷酸多型性的現象。

這些在顯微鏡下無法觀察到的細微結構變化，即稱為“次顯微結構變異”

(Submicroscopic Structural Variation)。除了 DNA 定序方法外，還有許多方法可以偵測基因組裡特定位置的次顯微結構變異情形，例如：原位螢光雜交技術 (Fluorescence In Situ Hybridization)，以及許多以 PCR 為基礎的實驗方法，像是 Real-Time Quantitative PCR、Multiplex Amplifiable Probe Hybridization(MAPH)、Multiplex Ligation- Dependent Probe Amplification(MLPA)等。然而，雖然上述的方法可以有效的評估微小 DNA 片段的增加或減少，但只能針對已知的位置，或某些特定的位置(Targeted Regions)，其在基因組範圍的廣泛應用上受到很大的限制，因此對於分佈於全基因組、大小不等的結構變異，瞭解仍然有限。近幾年來，由於全基因組雜交技術有著快速的進展，其能掃描全基因組的優勢及逐漸提高的解析度，使得辨識基因組裡結構變異的情形有了革命性的突破，這讓我們發現介於幾個 Mb 的顯微鏡層次變異，到小於 1Kb 的次顯微變異之間，還存在著許多大小不等的 DNA 片段有著劑量變化的情形，並且大量的散佈於基因組中，這些介於 1Kb 到幾個 Mb 之間的 Deletion 或 Duplication 的 DNA 片段，稱為基因組劑量變異(Copy Number Variations: CNVs)。CNVs 大則可能包含數百萬個核苷酸，含蓋許多功能性基因及基因的表現調節區，因此它對於人類的遺傳多樣性、表型變異是非常重要的，影響的程度可能遠大於單核苷酸多形性所造成的遺傳變異，同時 CNVs 在物種的演化和發展中，必定也扮演著非常重要的角色。這使得人群中 CNV 存在的調查，成了近幾年基因組結構研究的熱門話題，然而這一類的次顯微變異，與人類健康、疾病之間的關係尚不清楚，仍待進一步的研究。

### 第三節 結構變異的類型

綜合以上所述，人類基因組的結構變異，小至單一個鹼基，大至幾個 Megabase、甚至是整條染色體，主要包含以下數種形式(3, 13)：

1. 單一核苷酸的變化：

即點突變或單一核苷酸多型性(SNPs)。可能起源於複製失誤、DNA 修復

機制失效，或環境因素誘發核苷酸發生置換或改變。而視核苷酸發生的位置是否位於基因功能區、表現調節區，或是否影響編碼所對應的胺基酸，決定著其對表型的影響程度。

## 2. 倒位(Inversions)：

倒位的發生為單一染色體發生斷裂，再由期間兩片段反轉而重新組成；或是某個 DNA 單股上、序列方向相反的相似性序列的不正常配對、重組，也可能導致倒位的發生。倒位一般認為並不會引起帶原者的不正常表現型，因為其為平衡重組(Balanced Rearrangement)，但會產生不正常的配子，而產生基因組不平衡的子代。然而，若倒位發生於基因表現區，則可能破壞基因結構而引起疾病，例如：近一半的 A 型血友病是由於倒位的發生破壞了基因的結構而致病。

早期，倒位的鑑別僅能藉由顯微鏡下辨識染色體帶數(Band)的順序改變來達成，因為染色體組型分析方法(Karyotyping)的解析度不夠，故對這一類結構變異的瞭解十分有限，直到定序方法及原位螢光雜交技術發展以後，才瞭解基因組裡存在更小的倒位變異。一篇研究基因組細微結構變異的文獻(14)顯示，基因組裡有許多倒位多型性(Polymorphic Inversions)，大小約介於 5 Kb 到 1.9 個 Mb，值得注意的是，這些倒位所發生的斷裂點有四分之三存在於片段性重複序列(Segmental Duplication)的位置，這說明了這些倒位的發生可能是由於染色體側端重複性結構(Flanking Repeat Structure)的重組所造成的。另外，許多倒位的發生位置經常伴隨著基因物質的增加或減少，這篇文獻首先提出倒位也可能是基因組不平衡的結構變異。

## 3. 易位(Translocations)：

為兩染色體片段之交換，通常發生於非同源染色體之間，又分為相互易位(Reciprocal Translocation)及專發生於 Acrocentric Chromosome 中心粒區的羅勃遜易位(Robertsonian Translocations)，易位亦為染色體平衡重

組，因此有染色體易位者本身並不會有異常表型，但有較高產生不正常配子及不平衡子代的危險性。和倒位一樣，易位發生時，若染色體斷裂的位置為基因的表現區，則會破壞基因的結構或於重整後產生一個融合的新基因，因此而引起疾病。例如：著名的費城染色體即為第 9 對染色體與第 22 對染色體發生易位(Translocation)，使得第 22 對染色體上產生一個異常的融合基因 (BCR-ABL)，因而引發慢性骨髓性白血病。

#### 4. 插入(Insertions)、缺失(Deletions)：

插入或缺失可能代表遺傳物質的增加或減少，插入或缺失的發生可能只是單純的某 DNA 片段或數個鹼基對插入或缺損於基因組中的某個位置；也可能是基因組中高相似性或完全相同的 DNA 序列，在有絲分裂或減數分裂的過程中發生了重組，這種不對等的互換(Unequal Crossing over)的機轉可以發生在同一條染色體、同源染色體之間(Intrachromosome)或非同源染色體之間 (Interchromatid Non-allelic Homologous Recombination)，例如： $\alpha$  地中海型貧血其  $\alpha$  血球蛋白基因的缺損，及 X 染色體上紅綠視力色素基因的數目變異，都是由於不對等互換的機轉所造成的。另外，還有一種罕見的機轉，是某些重複性序列可以藉由反轉位的過程而造成結構變異，亦即將 RNA 反轉錄成 DNA，再把自己插入基因組中的不同位置。例如：在少數的 A 型血友病患者身上，其致病的原因為 L1 家族重複序列被插入在第 VIII 凝血因子其基因的表現序列中，因此干擾了序列使基因不活化。當然，插入或缺失的發生並不一定會引起疾病，目前應用解析度日益提升的全基因組(Genome-Wide)分析方法，使得基因組裡插入與缺失發生的位置、範圍，與臨床表現的關係，有了更全面、更進一步的瞭解。

#### 5. 串連重複性序列(Tandem Repeats)：

基因組裡有許多位置是由許多相同的核苷酸小單元連續排列而成的重複性片段，有的聚集在某個或少數幾個基因座，有的則如單次出現的 DNA



序列一樣分散在整個基因組中，其構成了 10~15%的基因組(1)，依照這些串聯性重複片斷的大小差異，可將其區分為衛星(Satellite)、小衛星(Minisatellite)及微衛星(Microsatellite)DNA。衛星 DNA 的小單元由 100 個以上的核苷酸組成，主要叢聚在染色體的著絲點或端粒區，一般認為其扮演著維持染色體結構的角色，確保有絲分裂及減數分裂過程中染色體的適當分離。小衛星 DNA 由 10 至 70 個核苷酸組成，微衛星 DNA 則含有 1 至 6 個核苷酸，小衛星及微衛星兩種重複性片段遍佈於基因組中。微衛星 DNA 依據其核苷酸組成的特性，又可細分為 3 種類型。第一種是完整型，序列中的核苷酸完全由重複性小單元頭尾連接所組成。第二種是間斷型，重複性小單元的序列中有核苷酸發生置換、刪除或嵌入的情形。第三種是複合型，序列中的核苷酸由兩種以上重複性小單元組成。其遺傳過程遵循孟德爾定律，在同源等位基因上，個體間主要的差異在於小單元的重複次數，因此這類串聯性片段的多型性，廣泛的被應用於連鎖分析、親子鑑定及刑事鑑定上。

這一類的重複性序列，雖然有許多並沒有轉錄活性，但其存在著大量的劑量變異，會使得其存在的位置容易發生非對偶基因同源重組(Non-Allele Homologous Rearrangement)，而重複序列間的異常重組可能造成 DNA 序列的插入、缺失、重複或倒位，因此是一些遺傳疾病產生突變的原因。例如杭丁頓氏舞蹈症，主要與位於 DNA 轉譯區的微衛星 DNA—(CAG) n 的重複次數倍增有關，轉譯結果多出一段麩醯胺酸蛋白，而造成一種神經退化性疾病。易脆 X 染色體症候群、肌強直型進行性萎縮症，則是在不轉譯區中三核苷酸重複序列不穩定地倍增，導致智能障礙或肌肉萎縮等遺傳疾病。

6. 片段重複性序列(Segmental Duplication or Low Copy Repeat)：

重複性片段又稱為低度重複序列，其散佈於基因組中，大小介於 1 Kb~400 Kb(15)之間，通常有超過 90%的序列是相同的，依其在染色體裡

分佈的位置，可大略區分成兩種類型，某些片段在同一染色體裡的多個位置重複散佈，稱為 Intrachromosomal Duplication；其他則散佈在非同源染色體裡，稱為 Inter-或 Transchromosomal Duplication。片斷重複性序列傾向聚集於基因組中著絲粒與端粒等區域，和前述的串聯性重複序列一樣，重複性片斷的位置亦為基因組裡結構較不穩定的熱門區域(Hot Spot)，容易促使染色體的非對偶同源性重組，造成遺傳物質的缺失、插入或複製，而引起遺傳性疾病。而這些片段的長度、序列的相似性、排列順序及重複性片段之間的距離，都是影響染色體不對稱配對發生頻率的因子。當不對稱配對發生在同源染色體、排列順序相同的重複性片段間時，會導致遺傳物質交互的增加或減少；但當不對稱配對發生在排列順序相反的重複片段間時，則會導致倒位(Inversion)的發生。大部分引起已知遺傳疾病的重複性片段都很大，約介於 10~400 Kb 之間，且其序列有超過 96% 的相似性。例如最常見的微缺失疾病，Digeorge Syndrome，其重複序列的涵蓋區域約 300Kb 大小，且該區域的序列有 99.7% 的相似性。

然而，有愈來愈多的證據顯示，基因組裡重複性的片斷亦有可能為正常的變異，除了重複性片段本身在不同的個體間即有數目變化的特性外，其藉由 NAHR 的機制所構成的多型性，廣泛分佈於正常健康族群，故亦屬於基因組劑量變異(Copy Number Variation)的一部份。

#### 7. 基因組劑量變異(Copy Number Variations)：

基因組劑量變異(Copy-Number Variant, CNV)也稱基因組劑量多型性(Copy-Number Polymorphism, CNP)，是一種大小介於 1Kb 至幾個 Mb 的 DNA 片段變異，因此包含前述的重複性序列以及插入、缺失，根據文獻，至少 10-20% 的基因活性遺傳變異是由 CNV 引起的(16)，基因組劑量變化可以透過破壞基因編碼蛋白的活性區域，改變基因的表現“量”；或者破壞控制基因活性的調節區域，影響基因活性。由於其在

人類基因組中廣泛分佈，覆蓋的核苷酸總數遠超過單核苷酸多型性 (Single Nucleotide Polymorphisms, SNPs) 的總數，極大地豐富了基因組遺傳變異的多樣性。基因組劑量變異對於物種特異性、基因組組成、物種的演化和系統發育，以及基因組某些特定區域的基因表達和調控，可能具有非常重要的生物學意義。

#### 8. 其他：

除了上述的變異類型外，基因組的結構變異尚包含了於正常染色體數量外、額外多出的標記染色體 (Marker Chromosome)；因染色體兩處斷裂後又重新連結成環形構造的環狀染色體；一染色體在著絲粒的兩側有相同的長臂或短臂的等臂染色體 (Isochromosome)；兩染色體分別發生斷裂後兩者互相連結所形成的雙著絲粒染色體 (Dicentric Chromosome)；以及一雙倍體 (Diploid) 的個體其某一同源染色體皆來自單一親方染色體 (Uniparental Disomy) 等等。

### 第四節 基因組劑量變異形成的機制

承上段的敘述，CNV 通常發生在包含或兩側夾有大片段同源性序列、或片段重複序列的位置。這些同源性序列或重複性序列會藉由非等位基因同源重組 (Non-Allelic Homologous Recombination, NAHR) 的機制，使重複序列間錯誤配對，引發不同的染色體重組，造成期間序列 (Intervening Sequence) 產生基因組劑量的變化。非等位基因同源重組會造成大片段的結構多型性，或直接導致基因組不平衡或早發型、高度遺傳的疾病。

然而，非等位基因同源重組並不能解釋所有的 CNV，部分也許是由一些非同源性重組的突變機制所造成的 (5)。如，某些 CNV 被發現與非  $\beta$  的 DNA 結構有關 (與標準的右旋  $\beta$ -Helix 結構不同，包含左旋的 Z DNA 及 Cruciform)，這些結構被認為會促進染色體重組，理論上會促進某些 CNV 的產生與存在。(5)

再者，非同源性末端連結修復的機制(Non-Homologous End-Joining Repair) (14, 22, 5)，為真核生物修復雙股 DNA 斷裂的一種機制，有時雙股 DNA 斷裂並不會以同源性序列為模板進行修復，而是經由一連串酵素的作用，直接將斷裂的末端連結起來，這樣的過程會導致部分遺傳物質的丟失，因而造成 CNV 的發生；而轉位 (Transposition) 和反轉位 (Retrotransposition) 也被認為是產生 CNV 變異的重要因素之一。(26)(27)(28)

CNV 的大小與其突變的機制間是有關係的，其在基因組中的分佈並不是隨機的。最近的研究顯示 (5)，大片段的 CNV 與片段重複性序列的相關性，較小片段的 CNV 為高，雖然其間的偏差為何並不清楚。此外，缺失與重複事件之間似乎因基因組承受遺傳物質遺失或增加的差異不同而有不同的選擇壓力。無論如何，在較小的、已知 CNV 之間，非同源性片段所引發的突變是比較重要的 CNV 形成機制。

## 第五節 基因組劑量變異(CNV)如何影響表型變化

不論是核苷酸層次或大片段的 DNA 改變，皆被認為會藉由改變基因的表達量來造成表型變化，因此藉由轉錄分析或評估單一基因的表達量的研究模式，可以瞭解 CNV 對表型變化的意義 (12)。部份文獻以老鼠模型來研究 CNV 裡的內含基因在不同組織的轉錄與表達，研究發現位於 CNV 內的基因，其基因的表達量較 CNV 外、沒有劑量變化的其他基因要來的低，表達量的變異情形也較大，這顯示 CNV 的內含基因與其他非 CNV 區域內的基因相較，有較為獨特的轉錄表達模式。在分析的老鼠模型與組織中發現，約僅有 5%~18% 的 CNV 基因表達量與基因的數目呈微弱的正比關係；而有 2~15% 的 CNV 內含基因其表達量和基因的劑量呈反比關係，造成此反比效應的機制仍不清楚，但可能可以用一些僅在個體發育早期表現的基因來解釋，這一類的基因在表達的初期其表現量是與基因劑量呈正比的，但當基因表達量達到一定程度時，基因產物反而會直接或間接的

誘導 Repressor 的表達，使得原先基因的表現量下降甚至是被終止；另一個可能的解釋是，CNV 內含基因其增加的劑量，造成空間上的障礙，使得轉錄的過程中某些轉錄調控因子沒有辦法鍵結至調控區，因此當基因的劑量愈是增加，愈是使基因表現量降低。再者，有三分之二的 CNV 內含基因，不論在任何組織間，其表達量和基因的劑量的多寡沒有關聯，顯示有代償機制的存在，或缺失/重複的片斷包含了不完整的調控序列。此外，基因印痕作用也可能調控 CNV 所在位置的基因表達。

因此，CNV 的存在影響基因的表達及表型變化，大致可以區分為以下幾種機制：

1. 劑量效應 (Dosage Effect)：DNA 片段的缺失或複製導致基因劑量改變，基因表現量因基因的劑量呈比例的增加或減少。
2. 位置效應 (Position Effect)：CNV 的存在可能使得某些功能性基因與調控序列(Cis-Acting)的距離增加，使得調控序列表現的效能降低，間接影響基因的表現量。
3. 改變基因組的結構：CNV 的存在可能改變染色質的結構，干擾某些調控因子的結合；或影響細胞核內染色質在間期(Interphase)時的排列及分佈，進而干擾基因的表達。
4. CNV 其劑量的多態性影響基因表達的水準：外顯率(Penetrance)是研究遺傳性疾病時必需考慮的一個重要因素。值得注意的是，許多顯性遺傳疾病的外顯率有一定程度的變化。例如結節性硬化症以及第一型神經纖維瘤和乳腺癌就是典型的例子。針對此現象，Beckmann 等(23)提出 CNVs 可能是導致某些致病突變體基因外顯率降低的決定性原因。因此，對於一個致病的顯性突變基因而言，在未攜帶突變等位基因的染色體上多獲得一個正常等位基因的 CNV，可以通過基因劑量影響表型，使個體表現正常。

此外，基因組劑量的變化除了會影響其內含基因的表現外，也會影響其鄰近、沒有劑量變化的基因表現，影響的範圍甚至可以達到 0.5 Mb。而其改變鄰近基因的表現，亦是經由位置效應或改變基因組的結構等機制來達成。

## 第六節 基因組劑量變異(CNV)對個體表型及疾病易感性的影響

基因組廣泛分佈的劑量變異可以引起基因組和分子表型的異質性，影響複雜性疾病的發生。文獻指出(16)，大約有 40% 的 CNV 位於基因貧乏的區域(Gene Desert)，但儘管如此，仍有大量的基因位於 CNV 內。Redon 等(15)在 1447 個經 HapMap 檢測到的 CNVs 區域中，發現 2908 個 NCBI 參考序列基因和 285 個 OMIM 基因，顯示 CNV 與複雜性疾病或孟德爾遺傳性狀相關。存在 CNV 的基因通常不編碼與生長發育相關的蛋白，但是，存在 CNV 的基因常影響人體對外界環境的反應，在細胞連接、感觀理解、化學刺激、神經生理過程中發揮重要作用。不存在 CNVs 的基因往往是劑量敏感性基因，參與維持細胞的生長發育，包括細胞信號傳導、增殖、激酶化和磷酸化等過程。此外，一些 CNV 區域覆蓋非編碼的 RNAs 區域，包括 miRNAs。miRNAs 調節基因表達轉錄後加工的能力，對生長發育、正常生理過程及包括癌症在內的複雜性疾病的形成起重要作用。

目前已發現不少複雜性疾病和 CNV 有密切關聯。2008 年 9 月發表在 Nature 的一項研究證實了先前的發現(18)，在 22 號染色體的一個區域有長度為 3 百萬鹼基對缺失的人，三成患有精神疾病，如自閉症(Autism)和精神分裂症(Schizophrenia)等。2008 年 8 月發表在 Nature Genetics(19)的一項研究發現克羅恩病(Crohn's Disease)和 IRGM 基因(和對抗侵入性細菌有關)上游區域 20,000 鹼基對的缺失之間、存在著相關性，IRGM 缺失可能會使罹患克羅恩病的風險增加 40%，但一個人實際患病與否可能取決於別的遺傳或環境因素。而在 2009 年 1 月另一篇發表在 Nature Genetics 上的文章(21)發現體重指數和一個名為 NEGR1 的基因中、45,000 個鹼基對缺失具有很高的相關性，這個基因調節飢餓感、代謝

與下視丘的神經生長。

人群中一些發生頻率大於 1% 的基因組多態性結構變異也與疾病易感性有關，如與睪丸激素代謝有關的 UGT2B17 基因，其缺失會增加罹患前列腺癌的風險；CCL3L1 基因劑量增加可以減少感染 HIV 的風險；DEFB4 基因其劑量減少會增加患節結性腸炎的風險；FCGR3 基因的劑量減少則容易患血管性腎炎。另外，在 1q21.1、15q11.2、15q13.3 等位置發生結構重組，亦被認為與智能障礙或精神分裂症有關 (12)，但在這些位置有相同結構變異的不同個體，卻不一定有完全相同的表型，這說明了某些疾病的狀況並非單純的和 CNV 的劑量相關，而與基因組存在的結構改變較有關係，因此不論基因的劑量，而是干擾了某些特殊的途徑。至於多樣、不同類型的結構變異，也可能同時對某一類複雜疾病產生影響。

由於複雜性疾病染色體層次的分子機制仍未完全明朗，在基因組層次僅憑 CNVs 來解釋複雜疾病的發生是不適當的。理論上，複雜性疾病的基因組變異，對表型的影響應較為“彈性(Soft)”。例如，非編碼區序列基因組劑量的變異，儘管改變了基因的劑量但並不使基因的功能完全消失。由於 CNVs 存在多種作用機制，對於分子表型和基因表達的影響效果也不同，因此對 CNVs 的臨床意義和遺傳方式的解釋必須加倍小心，需要基於全面的基因組變異評估。

## 第七節 基因組劑量變異(CNV)的臨床意義與重要性 (5)

根據先前的「人類基因組計劃」，原本認為人類的基因組有 99.9% 的序列是相似的，隨著分析方法的進步，才揭露了基因組劑量變異的大量存在，推翻了早期認為人類基因組大致相似的結論。由於 CNV 發生的頻率遠遠高於顯微鏡下觀察到的染色體結構變異，而且在整個基因組中覆蓋的核苷酸總數大大超過 SNP 的總數，因此基因組劑量變異必定和表型變異緊密關聯，同時在物種的演化和發展中發揮著重要作用。

大片段的重複或缺失已知與某些遺傳疾病的存在有關，可能是因為對劑量敏感的生長發育基因，其劑量發生了改變。這樣的觀念引導了基因診斷方法的建立。如果一個新偶發的染色體異常在一個基因組結構異常的病人被辨識出來，而這個變異並不是已知常見的染色體多型性，那麼這個異常便會被認為可能是引起異常表型的原因。

一般而言，標準的細胞遺傳學分析方法，全基因組的染色分析通常可以辨識 5~8 Mb 及更大的染色體重組/變異。然而，更高解析度的全基因組分析方法，可以發現更多細微的基因組異常，隨著分析方法解析度的改善，伴隨而來的是增加解釋基因組不平衡的難度。為了協助全基因組、高解析度 CGH 分析結果做正確的診斷解釋，關於人類基因組不平衡與相關表型的資料庫已經被建立，並且廣泛收集許多 CGH 的分析結果及其他相關的分析訊息，這個資料庫的目標是要改善次顯微染色體不平衡的研究及醫療照護。

基因組的不平衡若是由一個表型正常的人遺傳而來，通常會被認為臨床的重要性較低。但若另一個人有著相同的基因組變異，但也許沒有遺傳其他代償性的變異，就可能就會導致不一樣的臨床表型。這些情形顯示著解釋這些基因組不平衡的臨床意義是缺乏信心的，因此瞭解基因組劑量變異在基因組中存在的位置，及其與個體健康之間的關連性，是非常重要的。

CNV 並不會導致早發型、高度遺傳的疾病，因此被認為其必然是功能正常的，但其可能會和其他的結構變異相互作用，或結合 SNP 及環境因素共同對個體的表型發生影響，因此在晚發型遺傳疾病或一般性、複雜性的疾病，仍然扮演著相當重要的角色。根據最近已知的 CNV 其功能性的貢獻，顯示特殊環境因素的刺激，影響著分子與環境之間的交互作用，例如藥物的解毒作用

(Glutathione-S-Tranferase, Cytochrome P450 Gene, Carboxylesterase Gene Families)、免疫反應和發炎反應(Leukocyte Immunoglobulin Like Receptor, Defensin, APOBEC Gene)、組織表面完整性(Late Epidermal Confide Envelop and Mucin Gene Families)、表面抗原(Galectin, Melanoma Antigen Gene, RH Blood



Typing Gene family)等。同樣的，某些包含有基因的 CNV，會使得不同的個體在藥物反應、免疫能力、疾病抗性/ 易感性等有差異性。例如不同的個體、不同的族群間 CCL3L1 基因的劑量有所不同，其是 HIV-1 的 Suppressive Cytokine 和 Ligand for HIV Coreceptor CCR5 的基因，當個體 CCL3L1 的基因劑量較少時，CCR5-CCL3L1 的複合體就較少，便會留下較多的 CCR5 供 HIV 病毒結合、進入，因此會加 HIV/AIDS 的易感性。最近，在老鼠體內發現 Fcgr3 基因的劑量變化，當其基因劑量較少時，易引發類似人類的腎絲球腎炎。Fcgr3 負責製造 Macrophage 表面的 Transmembrane Receptor，當其活化時會導致吞噬作用及毒殺作用，Fcgr3 基因的增加對 Fcgr3 的功能有抑制的效應，因此 Fcgr3 基因劑量減少卻會增加其免疫反應，在某些老鼠個體便可能導致自體免疫疾病。而人類 Fcgr3 的直系同源 (orthologous) 基因亦經常觀察到有 0~4 個基因劑量變化，相關的研究顯示，人類 Fcgr3 同源性基因的低劑量是造成個體易罹患腎絲球腎炎的危險因子。

對這些結構變異的認識愈是深入，醫學界可以瞭解基因組中基因的增加、損失或者改變與疾病的關係，以及特定藥物或者療法用在不同人身上為何會有不同反應，並有助於瞭解目前為止仍無藥可醫的疑難雜症，找出新藥或者新的治療方式。因此，基因組劑量變異圖譜可以發現特定基因差異對特定疾病的影響，幫助治療和預防疾病。所以建立基因組劑量變異的資料庫，分享基因組劑量變異知識，將對基因組劑量變異影響表型改變及疾病易感性的特質，能有更深入的瞭解及應用。

## 第八節 偵測基因組劑量變異的方法

偵測基因組劑量變異的方法有非常多，傳統的細胞生物學檢驗方法，如染色體組型(Karyotyping)、原位螢光雜交技術(Fluorescence In Situ Hybridization)，以及以 PCR 為原理的分析方法，都能偵測基因組劑量變異的存在，但上述的方法分別受限於解析度及分析範圍的限制，直到近幾年比較式基因組雜交技術及電

腦運算技術的進步，才改善了上述的問題，同時對於發現基因組裡 CNV 的廣泛存在及分佈，有著極大的貢獻。針對近幾年常用的基因組劑量變異分析方法，分別介紹如下：

### 1. 比較式全基因組雜交技術 (Genome-Wide, Comparative Genomic Hybridization)

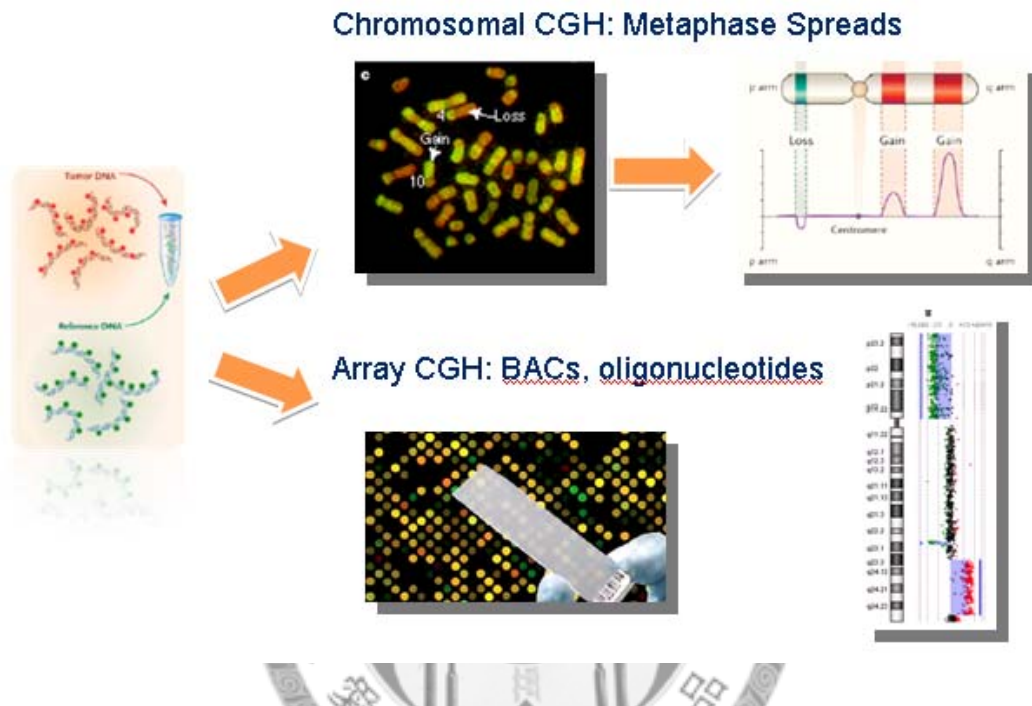
比較式基因組雜交晶片是目前檢測 CNV 最常用的方法，它能快速、有效的掃描全基因組劑量變異的情形。比較式基因組雜交技術最初用於檢驗癌症樣本與對照樣本間的染色體重組，由於是在玻片上使樣本與中期染色體(Metaphase Chromosome)進行雜交反應，因此解析度較差，通常需大於 4 Mb 的片斷才能做清楚判讀(3)。而隨著人類基因組定序工作的順利進行，BAC 資料庫、cDNA 資料庫和 Contig 資料庫在公共數據庫中大量累積，加上微陣技術的成熟，細菌人工染色體被廣泛的應用於微陣中，以細菌人工染色體製備的微陣又稱為基因晶片，由此發展出的技術稱為比較式微陣全基因組雜交晶片 (Array-CGH)，其雜交後的影像解析度可以大大的提高，並簡化整個實驗的分析步驟。Array-CGH 的原理與 CGH 的原理基本相同，區別在於將基因組 DNA 轉移到晶片(微陣)上以替代細胞分裂中期的染色體。用於建構微陣的 DNA 序列(亦指晶片上的探針)，可以有許多不同的形式，包括大小介於 75~200 Kb 左右的細菌人工染色體(Bac Artificial Clone)，較小的插入 Clone 如大小介於 30~40 Kb 的 Cosmid、大小介於 40~50 Kb 的 Fosmid，以及 cDNA Clone (0.5 ~2 kb)、PCR 產物 (100 bp ~1.5 Kb) 和寡核苷酸鏈 (25 ~85 bp) 等(29)。隨著晶片上所涵蓋的探針的數目不同、探針本身長度的大小不同，及探針所含蓋的基因組的密度不同，使得不同的 Array CGH 晶片其解析度有所差異。目前基因組雜交技術的發展趨勢致力於縮短探針的長度及增加晶片上微陣的密度，因此解析度由最早期 CGH 的幾個 Mb 提升到 Bac Array 的 1Mb 左右，再到

Oligo- array 的 30~50Kb (3)。

SNP 晶片是另一種有效檢測 CNV 的技術(29)。與比較基因組雜交晶片不同的是，SNP 晶片不需要同時使用兩個樣本的 DNA (實驗組和對照組) 和探針進行雙雜交，只需單雜交即可完成。每一個 SNP 位元點由一組探針來對應，每組探針包括 Perfect-Match(PM)和 Miss-Match(MM)探針，MM 探針除了中間鹼基不同於 PM 探針外，兩者有相同的 DNA 序列，主要作為內部對照之用。為了降低基因組的複雜性，常常使用限制性內切酶對整個基因組 DNA 消化後再進行 PCR 擴增，然後與晶片雜交。通過對 Perfect-Match、Miss-Match 探針的信號強度與其他個體相對值的比較，使用特殊的演算法，可確定基因組裡特定位置的劑量變化。此外，SNP 晶片還有另一個優點，即能提供基因型的訊息，如：能顯示雜合缺失(Loss of Heterozygosity)，或指出單一親方雙倍染色體(Uniparental Disomy)的存在。雖然 SNP 晶片探針的密度足夠大，但由於其在基因組中不是均衡分佈，在一些重複序列和複雜的 CNV 區域，SNP 密度較小，因而 CNV 圖譜的解析度較差。於是 Affymetrix 公司和 Illumina 公司提出了一種新的策略，即在新一代晶片中增加偵測 CNV 多型性的探針以提高圖譜的解析度。例如 Affymetrix 公司的 Genome-Wide Human SNP Array 5.0，除了包含約 500,000 個全基因組 SNP 探針外，還包含有 500,000 個多型性探針，可用於檢測全基因組的 CNV 多型性，建構較高解析度的 CNV 圖譜。

與傳統的細胞遺傳學方法相較，Array CGH 由於高解析度及可掃描全基因組的優勢，因此可以顯示額外的、原先非預期存在的基因組不平衡情形，同時精確的指出基因組不平衡發生的位置。此外，由於其操作不受限於需製備足量、高品質的中期染色體，因此操作較容易、需要的檢體量較少，臨床上完成檢驗報告所需的時間也較短。還有，Array CGH 可提升嵌合體(Mosaicism)診斷率、協助發現新的基因及微缺失/微增加疾

病(Microdeletion/ Microduplication Syndrome)。綜合以上優點，Array CGH 除了是分析基因組 CNV 的有效工具，更是對許多疾病的臨床診斷、研究有著非常大的貢獻，相信隨著成本的逐漸降低，Array CGH 將會逐漸成為臨床上主要的、例行性的分析工具。



圖一：CGH 與 aCGH 的操作原理

## 2. 以 PCR 原理為基礎的分析方法(Target, PCR-Base Analysis Method)

針對基因組裡特定區域的分析，以 PCR 為基礎的方法是最快速且有效的，其中定量型 PCR (Real-Time Quantitative PCR: qPCR) 是最被廣泛使用的一種，定量 PCR 也可用於 CNV 的檢測，雖然每次 PCR 都能有效的評估待測區域其基因劑量的增加或減少，但每次 PCR 只限於一個位置。近年來 PCR 檢測 CNV 的方法發展迅速，已可針對多個位置同時進行分析，如下列：

### (1) 多重可擴增探針雜交技術 (Multiplex Amplifiable Probe

Hybridization, MAPH): 根據所檢測的 DNA 序列, 製備數個具有通用引子的 PCR 產物作為探針組, 與待測基因組 DNA 雜交、擴增並回收特異性雜交探針, 透過螢光信號的強度差異分析來了解基因組中特定基因片段其基因組劑量的變化。

(2) 多重連接探針擴增技術(Multiplex Ligation-Dependent Probe Amplification, MLPA): 是一種靈敏度極高的相對定量技術, 利用簡單的雜合、連接及 PCR 擴增反應, 於單一反應管內同時檢測多個 DNA 片段的基因組劑量變化。

偵測基因組結構變異的 PCR 相關方法還包括一些常規分子生物學技術, 如短片段多重定量 PCR(Quantitative Multiplex PCR of Short Fluorescent Fragments, QMPSF)、qPCR、DNA 印痕技術(Southern blotting)、同源序列比率實驗等。

### 3. 電腦運算方法(Computational Approaches)

基因組裡的結構變異也可以經由比對不同來源的 DNA 序列而予以辨識, 由於在基因組中富含 CNV 的區域, SNP 位點的逐代傳遞可能並不符合孟德爾遺傳, 也可能不符合 Hardy - Weinberg 平衡, 因此利用 SNP 的基因型數據 (Parent - Offspring Trios), 通過特定的演算法, 可以進行 CNV 的預測。如上所述, Conrad 等(22)和 McCarroll 等(8)通過分析 HapMap Project 的 SNP 基因型資料即發現了多個 CNV。而以電腦運算方法來比對既有資料庫, 其中最簡單的方式就是直接比較兩個個體的 DNA 序列的不同處, 這種方法的好處是所有類型的結構變異都可以被偵測到, 包含平衡的變異(Balanced Variant), 且其解析度可以達到核苷酸的水平。

另一種電腦運算方法是利用已建構的基因組資料庫, 選擇特定、具

代表性的 Fosmid Clone，逐一比對 Clone 的端點序列，然後再組合 Fosmid Clone 的序列為染色體的序列，此即 Paired-End Sequence 方法，這種方法也能偵測部分倒位的存在，雖然它不像前者可達到核苷酸水平的解析度，但由於前者全基因組定序的工作相當龐大且費用昂貴，因此在成本的考量下，它仍然是一個實用的方法。

另外，末端配對作圖(Paired-End Mapping, PEM)是將基因組 DNA 打斷為 3 kb 左右的片段，經過生物素標記、連接、成環、測序，與人類基因組參考序列比。這是目前在全基因組範圍內分析小片段結構變異方法。

以上三類主要的分析方法，各有其優勢及限制，如比較式全基因組雜交技術，雖然可以快速、有效的掃描全基因組的劑量變化情形，但其只能鑑別劑量的改變，無法偵測倒位或易位等平衡變異、無法辨識在不同位置發生的 Interchromosomal Duplication，也無法有效的指出 CNV 發生範圍的起迄位置，只能進行大略的估計，且隨著不同平台所設計的探針長度、分佈位置與分度密度的不同，其解析度會有不同程度的差異。而以 PCR 為基礎的方法，雖然便宜、有效，可鑑別 1~50 kb 小片段的結構變化，卻受限於分析的範圍；至於以定序為基礎的分析方法或運算方法，可以分析倒位或易位等結構變異，且解析度最高，可以窺見 < 1Kb 等更小的結構變化，甚至可以達到核苷酸的水平，但受限於基因組廣泛定序的費用，在擴充有效的樣本數及花費時間上，是這一類方法主要面臨的問題。另外，若以電腦運算方式直接作序列的比對，則須考量基因組資料庫的完整性。鑑於不同研究方法的優缺點，使得不同方法的分析結果有不同類型及不同大小的差異，而由於分析的結果也必須排除可能存在的偽陰性或偽陽性反應，因此研究基因組的劑量變異，最好能結合兩種不同類型的方法進行驗證。

## 第九節 近期文獻對 CNV 的研究

目前基因組劑量變異(CNV)的研究內容包括偵測 CNV 在全基因組的分佈、基因組某個特定區域 CNV 的多態性與某種複雜疾病及疾病易感性的關聯分析，以及 CNV 的進化等。

2004 年有兩個研究團隊首先描述了在健康人群中，沒有明顯疾病表型的基因組劑量變異在基因組中的存在及分佈。Iafate 等(7)使用解析度為 1 Mb 的 Bac 矩陣晶片(Bac Array)，檢測 55 個正常個體的全基因組，發現了 255 個富含 CNV 的區域，這些 CNV 區域的大小介於 150~425 Kb 不等，平均每個人有 12.4 個基因組劑量變異；255 個富含 CNV 的區域(CNVR: Copy Number Variation Region) 中有 41% 在人群中重複出現的現象，而其中有 24 個 CNVR 發生率超過 10%；6 個 CNVR 的頻率超過 20%。此研究族群中，最常見的 CNVR 為含有 AMY1A 及 AMY2A 基因的 1p13.3 的位置，49.1% 的研究族群在 1p13.3 的位置有劑量變化的情形，顯示澱粉酶基因其基因組劑量的多型性變化，是造成不同個體間對澱粉的消化能力有所差異的原因之一。

Sebat 等(6) 使用解析度約為 35 Kb 的代表性寡核苷酸微矩陣晶片分析技術 (Representational Oligonucleotide Microarray Analysis, ROMA)，檢測 20 個正常個體，發現了 76 個 CNVR，其平均大小約 465 Kb，大約有一半的 CNVR 在研究樣本中有重複出現的現象，且平均每個人有 11 個基因組劑量變異。這 76 個 CNV 的位置，總共涵蓋了 70 個不同的基因，這些基因有的和調節細胞生長、代謝及神經功能有關；有的則已知會引起疾病。

接著於 2005 年，Tuzen 等(14)以電腦運算技術(Paired- End Sequencing)直接比較兩個基因組的 Fosmid DNA 序列資料 (Build 35, May, 2004，覆蓋基因組序列長度為 581 Mb，平均 Fosmid 長度為 40 Kb)，辨識出 241 個 CNVs，約 41% CNVs 覆蓋已知基因，其大小大多介於 8 Kb~40 Kb，而約有 86% 的 CNV 位置並沒有與先前 Iafate、Sebat 等所發表的文獻的分析結果有重複，這可能是因為所採用的

分析方法不同，尤其解析度的水平有很大的差異，因此此篇文獻大多數發現的 CNV 都小於前者所使用的晶片平台的解析度。而這種比較 DNA 序列的方法，由於能比較序列的排列順序，故還能偵測“倒位”這一類的結構變異。在這一篇文章中除了發現 241 個 CNVs，還辨識出 56 個倒位的斷裂點，這說明了這兩個基因組間將近有 300 個基因組結構變異。

2005 年 Sharp 等(15)於整個基因組中選擇了 130 潛在的染色體重排區域，即基因組裡已知的重複片段 (Segmental Duplication)，用以設計 Bac 基因晶片的探針(2194 個，總計長度 274 Mb，占全基因組序列的 10%)，然後使用比較基因組雜交的方法，分析 47 個健康歐洲人的基因組，鑑別出 119 個有發生 CNV 的區域，其中 73 個在先前的文獻沒有被報告過，66 在研究族群中是重複出現的。而這些區域是在基因組裡發生重組的熱門區域(Hot spot)，其 CNV 的發生頻率是基因組一般區域的 4 倍，這說明了基因組裡的重複片段因為容易引發染色體重組，所以是促進遺傳物質發生劑量變異的重要因素，特別是大片段的重複性序列(長度大於 10 Kb；序列相似性大於 95%)，愈容易透過非等位重組的機制，造成遺傳物質的增加或減少。令人驚訝的是，大部分的 CNV 並未涵蓋整個重組的熱門區域，有許多因素可能造就這樣的結果，這意味著個體之間可能還存在著許多潛在的、容易發生重組的位置。此外，由於重複性片段在不同染色體散發存在的特性，故其也具有基因組劑量變化的特質，因此藉由晶片所偵測到的大片段結構劑量變異，部分是起源於基因組本身所存在的重複性結構。

2006 年 Conrad 等(22)提出了一種透過 SNP 基因型數據 (Parent-Offspring Trios)分析來鑑別 CNV 的方法。藉由對 HapMap Project 的 SNP 基因型資料的分析，發現了 586 個區域存在較大片段的缺失，大小介於 300 bp~1.2 Mb 之間，其中有 30~50 個區域缺失片段大於 5 Kb；覆蓋的核苷酸總數為 550~750 Kb，覆蓋的基因總數為 267 個(包括已知的基因和預測的基因)。在缺失的區域，基因密度相對較小，此篇文獻推論可能是淨化選擇(Purifying Selection)的結果。



2006 年 McCarroll 等(8)使用前段所述同樣的基因型資料，開發了另一種算法，在整個基因組中共發現 541 個片段缺失，片段大小約在 1~745 Kb 之間。

而英國 Wellcome Trust (Institute, Wellcome Trust Genome Campus) Sanger 研究院，以及美國 Affymetrix 公司等多國研究人員組成的研究小組，於 2006 年公佈了人類基因組第一代 CNV 圖譜 (First-Generation CNV map) (16)。這張遺傳圖譜是透過對歐洲、非洲和亞洲 (HapMap collection) 4 個族群的 270 個樣本進行分析(30 個源於奈及利亞 Yourba 核心家系；30 個源於美國猶他州歐洲後裔核心家系；45 個來自於日本東京、無親源關係的個體；45 個來自於中國北京、無親源關係的個體)，使用單核苷酸多態性 (SNP) 基因分型和以 Bac Clone 為基礎的比較基因組雜交技術進行 CNV 掃描，共發現了 1447 個富含 CNV 的區域，CNV 的平均長度於 Bac Array 平台為 341 Kb；於 SNP Array 平台則為 206 Kb。所發現的 CNV 中，有 43% 在研究族群是重複出現的，所有的 CNV 覆蓋基因組的範圍為 360 Mb，佔整個基因組的 12%，這些 CNV 包含了數百個功能基因、多個致病基因位置及基因的調控性序列。全部的 CNV 其包含的核苷酸總數顯著高於整個基因組的 SNP 總數，說明 CNV 在遺傳的多樣性、表型的多型性以及系統演化中可能扮演著重要角色。

2007 年 Wong 等(24) 使用 Bac Clone Based 比較式基因組雜交晶片掃描整個基因組，共發現 3654 個、大於 40 Kb、位於常染色體的 CNV，77% 的 CNV 未在先前的相關文獻報導過。有趣的是，14 個 CNV 多態性豐富的區域，包含 21 個編碼有已知功能的 micro-RNAs 分子，顯示著 micro-RNAs 可能也與表型的多樣性有關。

2008 年 Kidd 等(26)使用 Fosmid Subclone，建構了 8 個個體的全基因組資料庫(4 樣本為奈及利亞約魯巴人，4 樣本來自非洲以外)，經由和已發表的人的基因組序列相比對，發現了 1695 個大片段的 DNA 變異位點(序列長度大於 6 kb)，及 525 個未在參考基因組發現的插入性序列。所發現的 CNV 中，有 50%

在研究族群中重複出現，也有 50%是在此篇文獻中新發現的。另外，261 個 CNV 的解析度可以達到單核苷酸水平。

2008 年 Wang (31) 等公佈了第一個亞洲個體的雙倍體(Diploid)基因組序列，其鑑別了超過三億個 SNP，其中有 13.6% 未在 dbSNP Database 發表過，此外，也分別偵測到 2441 個 Deletion 與 33 個 Duplication，其平均變異的大小為 492 bp，其中 Deletion 的數量較 Duplication 高出甚多，可能與所使用的分析方法 (Paired-End Sequencing) 有關，當插入序列的長度大於 Paired-End Library 的 DNA 片段長度時，Duplication 的訊息比較容易被忽略。

2009 年 Shaikh (32) 等發表了一篇研究群體極大的研究，使用 SNP-based、高密度的 Oligo-Array，分析 2026 個沒有疾病表型的個體，這篇文獻的研究對象仍就以美國白人及黑人為主，亞洲人種僅佔了 0.6%。研究發現了 3272 個富含 CNV 的位置，總計占基因組的 5.9%，有 77.8% 的 CNVs 在研究族群中重複出現，而 51.5% 所發現的 CNVs 是先前文獻未報導過的。此篇文獻由於研究的樣本數大，因此對於美國本土族群提供了一個有用的臨床參考數據。

目前在人類基因組中檢測到發生 CNV 的位置總數已超出 14478 個 (<http://projects.tcag.ca/variation/>; Last updated: Mar 25, 2010)。比較上述不同的研究，結果的重疊性並不高。由於不同的分析平台其解析度不同，所涵蓋的基因組範圍也不同，且許多研究在檢測 CNV 的過程中過度強調避免假陽性，可能導致假陰性的比例升高。因此，上述估計還很有可能是一種偏低的估計(25)。

## 第十節 研究目的與動機

隨著分子診斷技術的進步，我們對基因組變異的情形有更深入及廣泛的瞭解，尤其是微陣列比較式全基因組雜交技術(Array Comparative Genomic Hybridization)的發展，提供了一個高解析度的分析平台，使得我們能更精確、快速的瀏覽整個基因組的劑量變異情形，特別是針對部分異常的臨床表型，當其基

因組的變異位置仍不清楚時，全基因組雜交技術是一個非常有效的分析及研究工具。

然而以全基因組雜交技術進行分析的過程中，我們發現正常人的基因組也存在著大量的基因組劑量變異(Copy Number Variations)，其變異的大小為 1 Kb (Kilobases)到幾個 Mb (Megabases)不等，涵蓋基因組極大的比例，包含許多基因及功能調節區(Regulatory Element)，由於這些 CNVs 並沒有明確的疾病表型相關性(Non-Diseased Phenotype Associated)，因此這些 CNVs 的存在，會增加檢測結果判讀及遺傳諮詢的困擾。故建立 CNVs 的參考資料庫，有效的鑑別 CNVs 發生的位置，區分人群中正常存在的多型性變異及有疾病相關性的 CNVs，對於臨床的遺傳諮詢工作是非常重要的。


近年來有越來越多的研究致力於人類基因組中 CNVs 的鑑別與分析，這些研究結果幫助我們瞭解人類族群中 CNVs 在基因組裡發生的位置、頻率與形式，雖然目前(Mar, 25, 2010)人類基因組變異資料庫 <http://projects.tcag.ca/variation> 已發表了 14478 個 CNVs 的位置，但早期的研究樣本數較小，難以呈現出人類基因組變異的全面樣貌；且仍有許多不同種族的 CNVs 尚未研究，尤其在先前的相關文獻中，亞洲族群的 CNVs 研究樣本比起歐美族群要少的多，因此調查台灣族群 CNVs 的分佈情形，以豐富 CNVs 資料庫的族群變異性，祈使對人類基因組變異資料庫的完整性有所助益；同時建立屬於台灣族群的 CNVs 圖譜，期待在臨床上解釋全基因組的分析數據時，能有效的鑑別有病理意義的結構突變及造成人群多型性的多態變異，以對遺傳諮詢有所幫助。

## 第二章 研究方法

### 第一節 研究樣本來源

本研究所採用的樣本主要來源於接受產前診斷的孕婦，剔除有家族遺傳史及有病理意義診斷的樣本後(如超音波影像顯示結構或發育上的畸形)，共收集了 1015 個樣本，其中內含羊水、臍血、絨毛膜等樣本各 971、35、9 個。這些樣本採用 Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN) 來萃取 DNA，並以標準的操作步驟執行。

### 第二節 以 CMDX Bac Clone aCGH 分析基因組中的基因組劑量變異



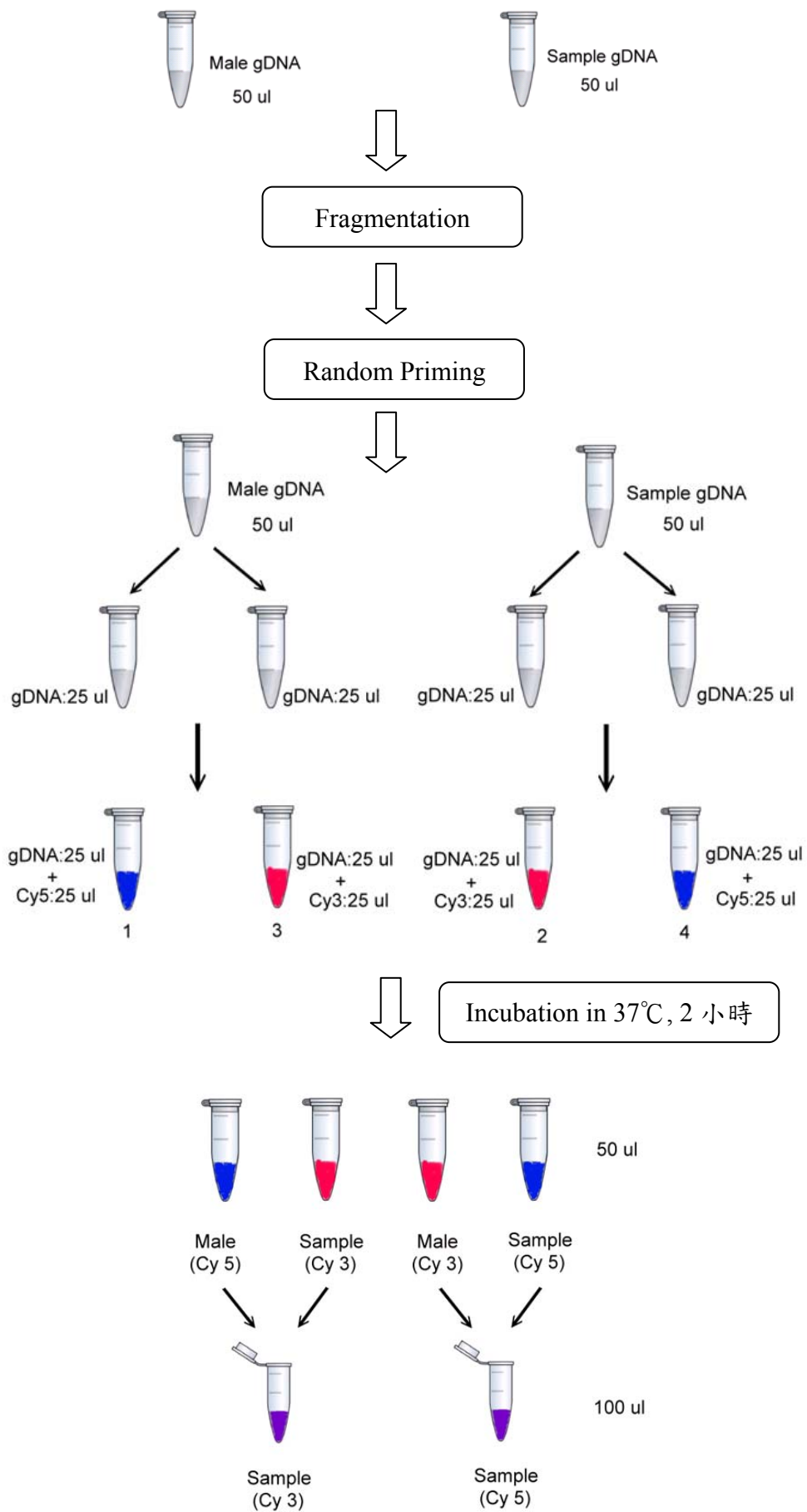
所萃取的 1015 個 DNA 樣本以微陣列基因晶片(CMDX: CA2500/ CA3000)執行基因組劑量變化的分析。CMDX Bac array CA2500/ CA3000 Platform 分別含有 2437 個與 3039 個不同的 Large-Insert Clone，分佈於整個基因組，每個 Clone 的大小約 150~250kb 不等。CMDX Bac Array Platform 採用 Dye swap，使待測檢體與參考檢體分別在 Forward 與 Reverse Array 區進行競爭型雜交反應，由於每一個 Bac Clone 在 Forward Array 區與 Reverse Array 區分別有 3 個重複的探針，因此整張晶片上每一個 Bac Clone 總計有 6 個重複的探針，由 6 個探針所測得螢光訊號的平均值來評估該相對應位置的遺傳物質劑量。當待測基因組有劑量不平衡的現象存在時，Forward Array 區與 Reverse Array 區探針的螢光值應呈相對的增加或減少(即當一區為 Duplication 的反應，另一區必為 Deletion 的反應)，而當兩區的平均螢光值呈現相對應的結果，取 Log<sub>2</sub> 後的數值若超過評估的閾值( $\pm 0.3$ )，分析軟體才會判讀為有劑量變異發生；若兩區螢光值呈現一致性的結果，或平均螢光值取 Log<sub>2</sub> 後未達閾值，則不予判定。這樣的設計提供了操作上品質管制的機制，能降低偽陽

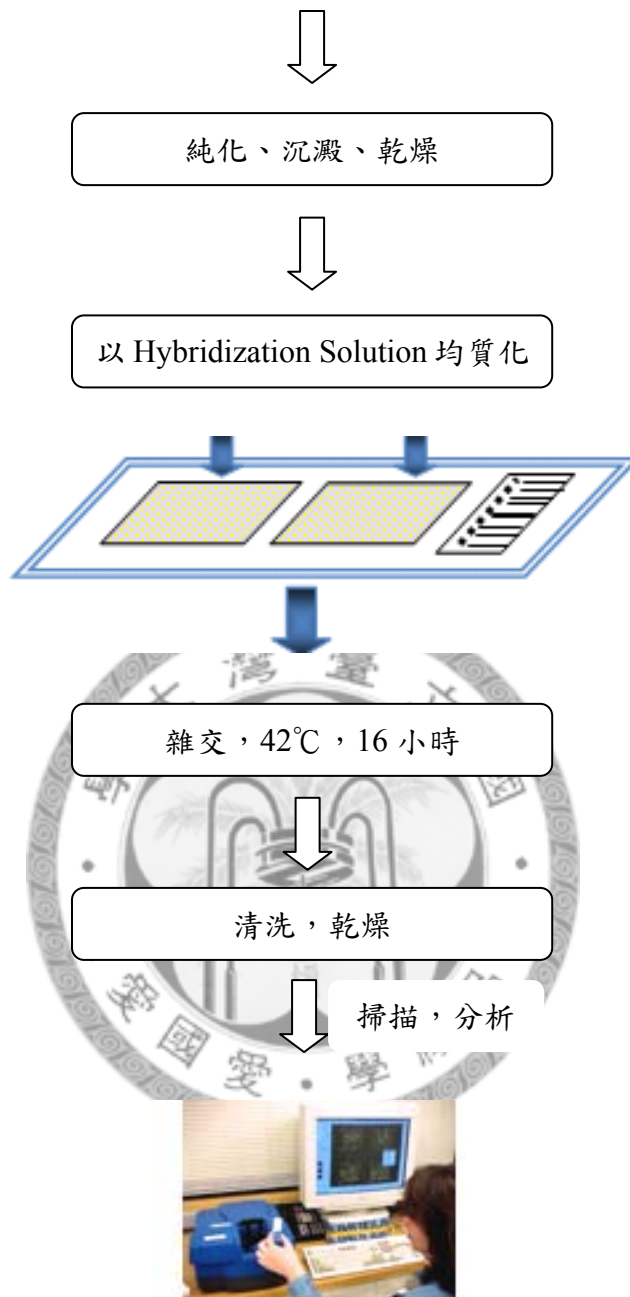
性率及雜訊的干擾，提供良好專一性及再現性的判讀結果。

以標準的操作方法進行比較式基因組雜交分析，首先配置適量( $2\mu l$ )的待測樣本及參考樣本，於超音波震盪器上將 DNA 打碎成適當的片段，然後將 DNA 樣本分裝成兩組，分別標示不同的螢光，即待測樣本+Cy3、參考樣本+Cy5；另一組則將螢光交換染色(待測樣本+Cy5、參考樣本+Cy3)。接著將已加入螢光的 DNA 樣本置於  $99^{\circ}\text{C}$  的乾浴槽加熱 5 分鐘，再放置冰上 5 分鐘，然後加入  $3\mu l$  之 Exo-Klenow，此目的為造成 DNA 變性(Denature)，促使螢光染劑嵌入 DNA 雙股結構，然後將 DNA 樣本置於  $37^{\circ}\text{C}$  水浴槽，至少作用兩個小時。

兩個小時後，將染有 Cy3 的待測樣本與染有 Cy5 的參考樣本混合，另一組染有 Cy5 的待測樣本與染有 Cy3 的參考樣本亦混合在一起，接下來便利用特殊的 Spin Column Tube (Invitrogen)，與 Binding Buffer、Wash Buffer、5M NaCl、75% Isopropanol、70%Ethanol 來純化並沉澱 DNA，過程中亦加入 Cot-1 DNA 來降低非特异性重複序列所引起的交叉干擾。

純化後的 DNA 樣本先置入乾燥箱中做完全乾燥，然後再加入  $25\mu l$  Hybridization Buffer 將 DNA 樣本打散、混合均勻，接著將兩組 DNA 樣本放置於基因晶片上 Forward 矩陣區與 Reverse 矩陣區、蓋上蓋玻片，進行競爭型雜交反應，此雜交反應須於  $42^{\circ}\text{C}$  恆溫培養箱中至少作用 16 個小時，並保持適當的溼度。雜交步驟後，以 Wash Solution 2X SSC-0.5%SDS、2X SSC-50% Formamide、2X SSC-0.1% Igepal、2X SSC 等，分別於  $50^{\circ}\text{C}$  恆溫振盪培養箱中緩慢搖晃以去除蓋玻片、清洗晶片，並使其乾燥。最後以掃描儀與 CMDX Software 偵測晶片上每一個探針的相對螢光強度，判讀是否有 Segmental DNA Gain 或 Loss 的訊號。如遇有病理意義的結構變異，則將該樣本由研究族群中予以排除。





圖二：Bac Array CGH 的操作流程

### 第三節 以 Competitive Multi-Plex PCR 驗證 CMDX Bac Array 的基因組劑量變異訊號

Competitive Multi-Plex PCR 是評估特定位置的基因其劑量增加或減少的有效方法，其原理為使目標基因與參考基因在同一個 PCR 反應中進行競爭型的基因產物擴增，然後執行電泳，依待測檢體其目標基因產物與參考基因產物的訊號比值、參考檢體其目標基因產物與參考基因產物的訊號比值，進行相除比較，以評估基因劑量的增加或減少。分析方法如下所示。(30)

$$\text{Copy Number} = \frac{\text{Peak area of candidate gene (U) / Peak area of reference gene (U)}}{\text{Peak area of candidate gene (C) / Peak area of reference gene (C)}} \times 2$$

(U): Unknown，表示未知的待測檢體

(C): Control，表示參考檢體

挑選 4 個 Bac Clones 作為驗證基因組劑量變異訊號的對象，其中兩個 Bac Clones：RP11-556H10、RP11-85G7，僅在一個進行人工流產所取得的臍帶檢體觀察到基因組劑量變異的訊號，其位於 12p12.3 的位置；另外兩個 Bac Clones 則在台灣族群中普遍存在 CNV，分別是位在 4p16.1 的 RP11-301J10 與位在 6p25.3 的 CTD-2249D1。針對上述 4 個 Bac Clones，分別挑選其所內含的基因：RERG、PIK3C2G、CPZ、DUSP22 等設計引子，並隨機挑選樣本，以 Competitive Multi-Plex PCR 驗證樣本在上述 4 個 Bac Clones 的劑量變異訊號。

將源自於上述 4 個 Bac Clone 的基因：RERG、PIK3C2G、CPZ、DUSP22，與參考基因 FGFR2、Kirt 分別在同一 PCR 反應中進行競爭型的 PCR 反應，PCR 反應的最終容積為 25 $\mu$ L，其包含的內容物與相對應的反應模組敘述如下：

(1) RERG+ FGFR2 +Kirt

PCR 反應溶液裡含 100ng DNA、200 $\mu$ M dNTP、2mM MgCl<sub>2</sub>、2.5 $\mu$ L 10X



buffer II、0.5U AmpliTaq Gold enzyme、0.025 $\mu$ M FGFR2 引子、0.03 $\mu$ M Kirt 引子、0.08 $\mu$ M RERG 引子，先於 95 $^{\circ}$ C 下使 DNA 變性 10 分鐘，接著連續在 94 $^{\circ}$ C/30 秒、57 $^{\circ}$ C/45 秒、72 $^{\circ}$ C/45 秒擴增基因產物 25 循環，最後於 72 $^{\circ}$ C 延伸 10 分鐘。

(2) PIK3C2G+ FGFR2 +Kirt

PCR 反應溶液裡含 100ng DNA、200 $\mu$ M dNTP、2mM MgCl<sub>2</sub>、2.5 $\mu$ L 10X buffer II、0.5U AmpliTaq Gold enzyme、0.025 $\mu$ M FGFR2 引子、0.03 $\mu$ M Kirt 引子、0.05 $\mu$ M PIK3C2G 引子，先於 95 $^{\circ}$ C 下使 DNA 變性 10 分鐘，接著連續在 94 $^{\circ}$ C/30 秒、57 $^{\circ}$ C/45 秒、72 $^{\circ}$ C/45 秒擴增基因產物 25 循環，最後於 72 $^{\circ}$ C 延伸 10 分鐘。

(3) CPZ+ FGFR2 +Kirt

PCR 反應溶液裡含 100ng DNA、200 $\mu$ M dNTP、2mM MgCl<sub>2</sub>、2.5 $\mu$ L 10X buffer II、0.5U AmpliTaq Gold enzyme、0.025 $\mu$ M FGFR2 引子、0.03 $\mu$ M Kirt 引子、0.05 $\mu$ M CPZ 引子，先於 95 $^{\circ}$ C 下使 DNA 變性 10 分鐘，接著連續在 94 $^{\circ}$ C/30 秒、54 $^{\circ}$ C/45 秒、72 $^{\circ}$ C/45 秒擴增基因產物 25 循環，最後於 72 $^{\circ}$ C 延伸 10 分鐘。

(4) DUSP22+ FGFR2 +Kirt

PCR 反應溶液裡含 100ng DNA、200 $\mu$ M dNTP、2mM MgCl<sub>2</sub>、2.5 $\mu$ L 10X buffer II、0.5U AmpliTaq Gold enzyme、0.035 $\mu$ M FGFR2 引子、0.04 $\mu$ M Kirt 引子、0.05 $\mu$ M DUSP22 引子，先於 95 $^{\circ}$ C 下使 DNA 變性 10 分鐘，接著連續在 94 $^{\circ}$ C/30 秒、57 $^{\circ}$ C/45 秒、72 $^{\circ}$ C/45 秒擴增基因產物 25 循環，最後於 72 $^{\circ}$ C 延伸 10 分鐘。

接著使用 GCK-500 cartridge kit (eGene, Irvine, CA, USA)來分離不同大小片段的 DNA 產物，首先取用 5  $\mu$ L 的 PCR 產物加入 15  $\mu$ L 的 dH<sub>2</sub>O 中，稀釋 4 倍，然後選用 e Gene Biocalculator Software、OM500 mtd 方法進行 DNA 產物分離，最

後 Biocaculator Software 便可自動標示出 DNA 產物的大小、分離先後順序及相對劑量。

將 Competitive Multi-Plex PCR 的結果與原 CMDX Bac Array 的基因組劑量變異訊號進行比對，分別評估 CMDX Bac Array 的偽陽性率及偽陰性率。偽陽性率=基因晶片顯示有劑量變化但 Multi-Plex PCR 未顯示出劑量變化的件數/ 基因晶片顯示有劑量變化的總件數；偽陰性率=基因晶片未顯示劑量變化但 Multi-Plex PCR 顯示有劑量變化的件數/ 基因晶片未顯示劑量變化的總件數。分析的結果可用以評估此 CMDX CA2500/ CA3000 Bac Array 平台的準確度。

#### 第四節 與現有基因組變異資料庫進行比對

人類基因組變異資料庫 <http://projects.tcag.ca/variation> 持續收錄自 2004 年以來研究基因組結構變異的文獻的研究結果，提供了人類基因組全面的結構變異摘要，透過這個平台，將 1015 個樣本所觀察到的基因組劑量變異與先前文獻的分析結果做一比對，統計台灣族群廣泛存在或具獨特性的基因組劑量變異。

The screenshot shows the homepage of the Database of Genomic Variants. The header includes the title "Database of Genomic Variants" and the subtitle "A curated catalogue of structural variation in the human genome". It is hosted by "The Centre for Applied Genomics". Navigation links include "About The Project", "Genome Browser", "Download", "Links", "Data Submissions", and "Email us". A dropdown menu for "Please select genome assembly" is set to "Build 38 (Mar. 2008)".

Search options include:

- View Data by Chromosome:** A list of links for chromosomes 1 through 22, X, and Y, plus an "All" link.
- Keyword Search:** A search box with a "Search" button and radio buttons for "Exact Match?" (Yes/No). Examples provided are "clone name, accession number, cytoband or gene".
- BLAT Search:** A text input field for "Enter sequence in FASTA format here:" and a "BLAT Search" button.

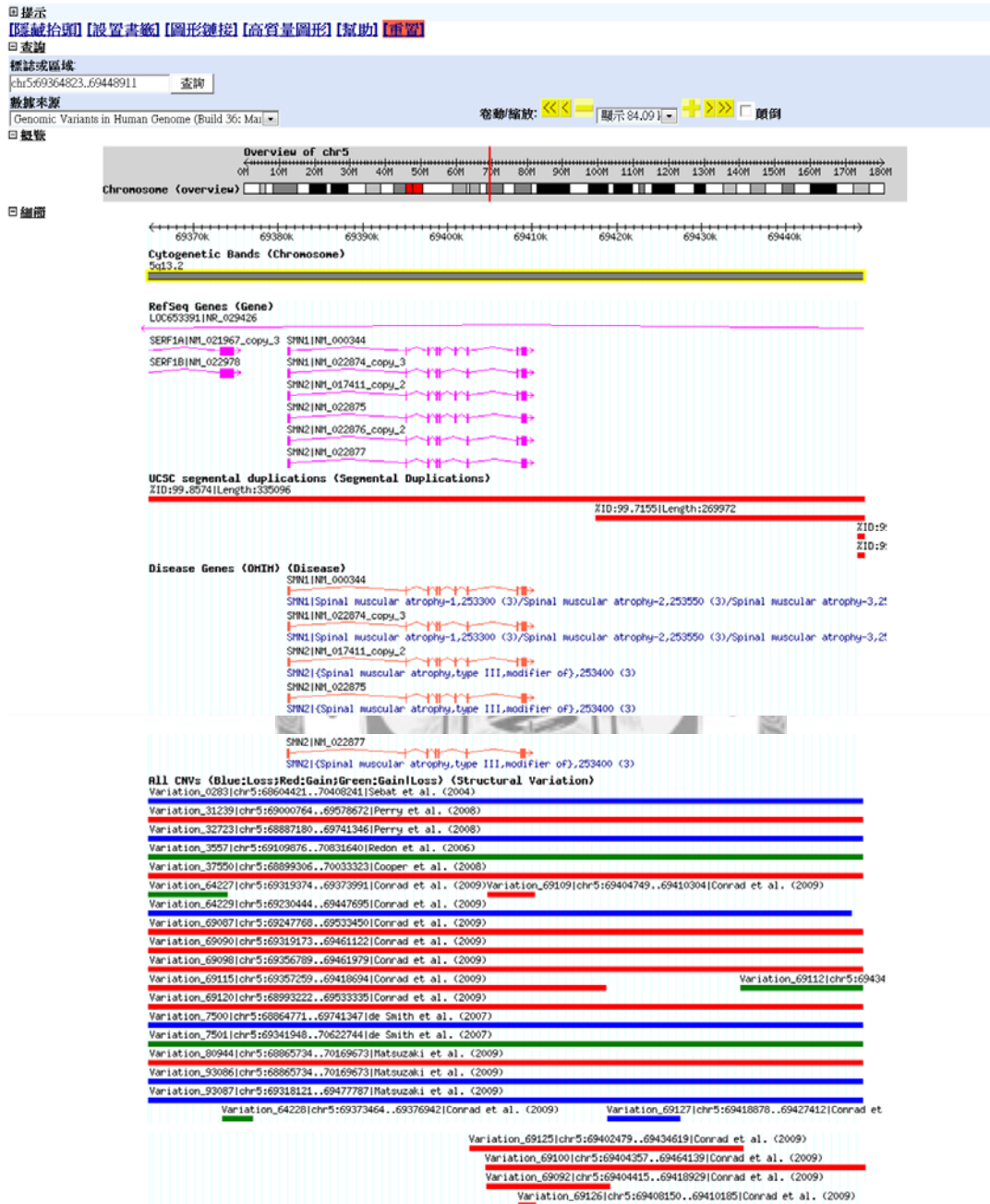
Summary Statistics on the right side:

- Total entries: 89427 (hg18)
- CHVs: 57829
- Inversions: 850
- Indels (100bp-1Kb): 30748
- Total CNV loci: 14478
- Articles cited: 38
- Last updated: Mar 25, 2010
- Join our mailing list

Footer text: "Read about the updates in our newsletter." and a citation: "In citing the Database of Genomic Variants please refer to: Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C: Detection of large-scale variation in the human genome. Nat Genet. 2004 Sep;36(9):949-51."

圖三：人類基因組變異資料庫 <http://projects.tcag.ca/variation> 搜尋介面

顯示 84.09 kbp 起始于 chr5, 位置從 69,364,823 到 69,448,911



圖四：Bac Clone: CTD-2041D13 (Chr5:69364823-69448911)於資料庫的搜尋結果

### 第三章 結果

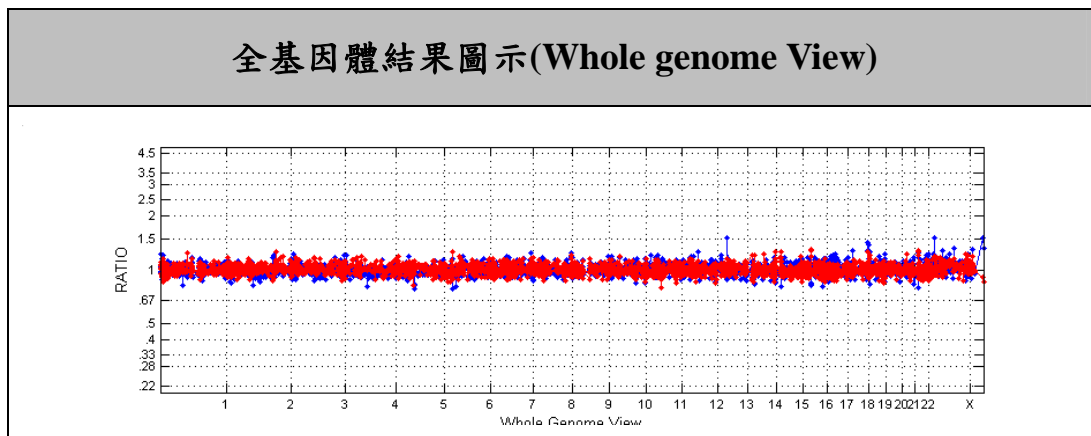
以 CMDX Bac Array CA2500/ CA3000 觀察 1015 個產前檢查樣本其基因組劑量變異在基因組的分佈，總計發現 4152 個劑量變異(gain or loss)的訊號，分佈在 438 個 Bac Clone 與 252 個 Cytobands 中。平均每個人可觀察到 4.36 個基因組劑量變異，但如果將 CA2500 與 CA3000 平台分開統計，CA2500 可觀察到的平均基因組劑量變異次數為 4.84 次，CA3000 則略低，為 3.88 次。

來源	樣本種類	數目
孕婦	絨毛	9
	羊水	971
	臍血	35
總和		1015

表一：分析檢體來源一覽表

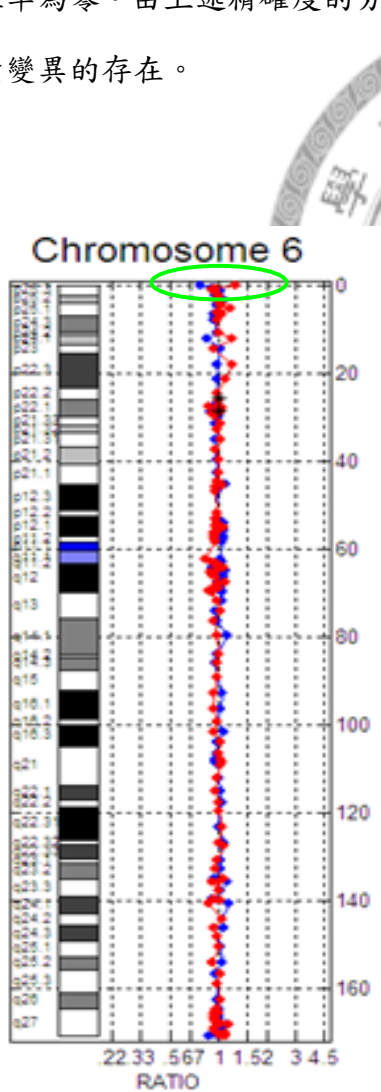
Platform	CNV平均	數量	次數
CA2500	4.84	221	1069
CA3000	3.88	794	3083
CNV平均/ Case No.	4.36	1015	4152

表二：個體的平均基因組劑量變異

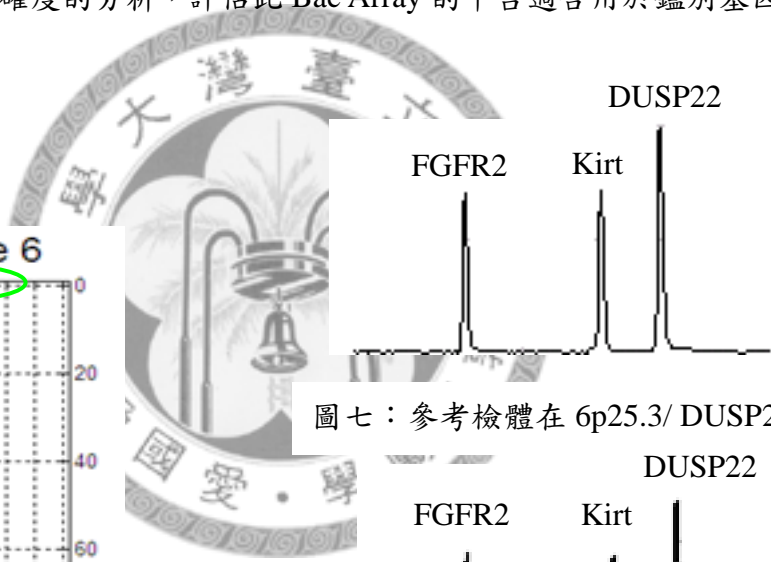


圖五：樣本 509 的全基因組基因組劑量變異掃描結果圖

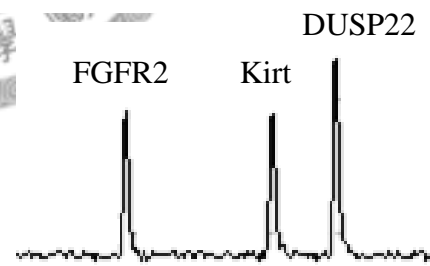
以 Multi-Plex PCR 方法驗證 CMDX Bac Array 的可信度，挑選於 6p25.3、4p16.1、12p12.3 等位置有劑量變化的樣本共 57 個，而由 Multi-Plex PCR 取得一致性的結果共 56 個，評估偽陽性率為 1.75%。其中，分析結果呈現不一致的樣本為 Sample313-2，經重覆分析後 Sample313-2 的 Multi-Plex PCR 結果仍為 Wild Type (Bac Array 結果為 Deletion Type)，此偽陽性的結果極可能為 Multi-Plex PCR 於 6p25.3 所設計的探針位置，並未和 sample313-2 產生劑量變異的位置重疊，因此使得 CMDX Bac Array 的分析結果與 Multi-Plex PCR 不一致。另於上述位置挑選沒有劑量變化的樣本 30 個，以 Multi-Plex PCR 取得一致性的結果為 30 個，評估偽陰性率為零。由上述精確度的分析，評估此 Bac Array 的平台適合用於鑑別基因組劑量變異的存在。



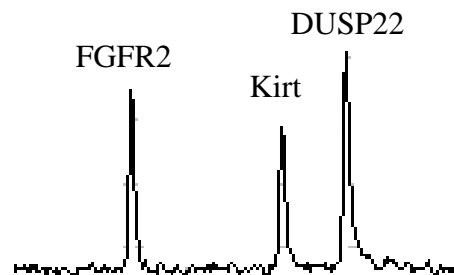
圖六：樣本 313-2 在 Chr 6 的 aCGH 結果，6p25.3 有一 clone 為 Deletion 表現



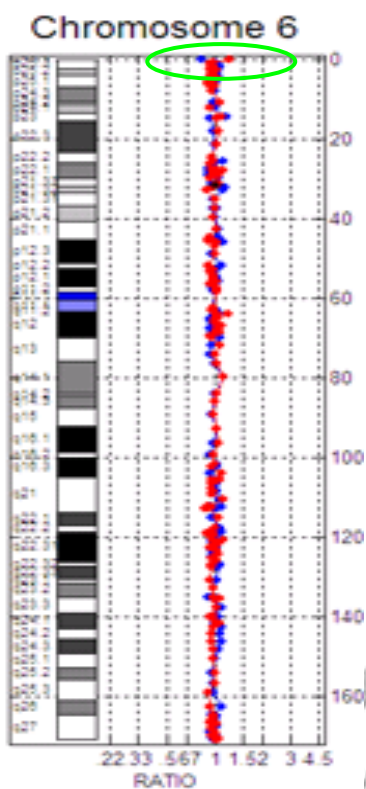
圖七：參考檢體在 6p25.3/ DUSP22 gene 的劑量



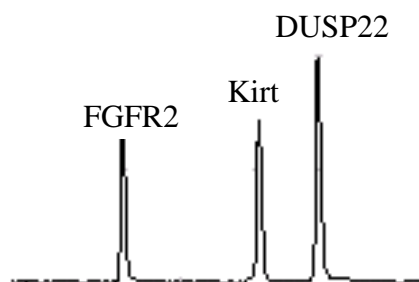
圖八：樣本 313-2 在 6p25.3/ DUSP22 第一次 Multi-Plex PCR 結果



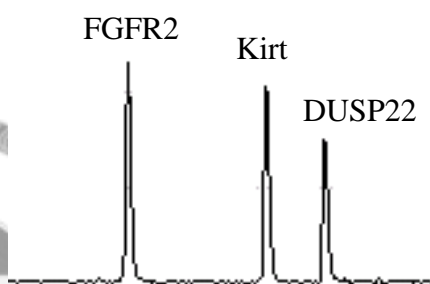
圖九：樣本 313-2 在 6p25.3/ DUSP22 第二次 Multi-Plex PCR 結果



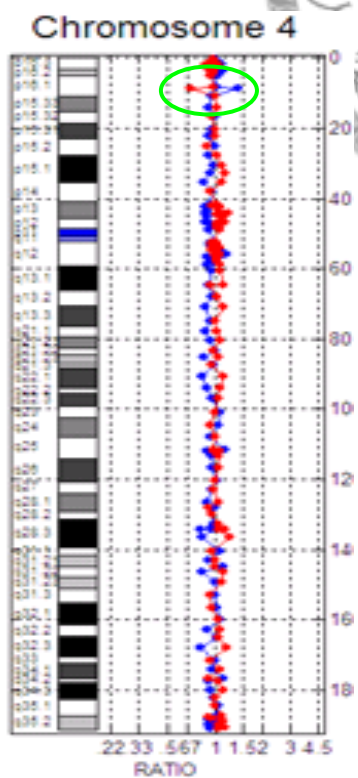
圖十：樣本 H268 在 Chr 6 的 aCGH 結果，6p25.3 有一 clone 為 Deletion 表現



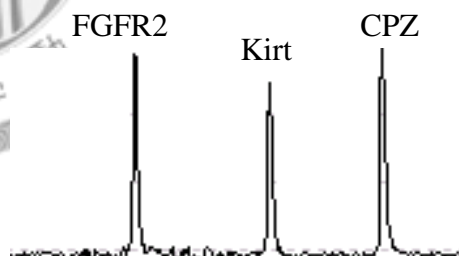
圖十一：參考檢體在 6p25.3/ DUSP22 gene 的劑量表現



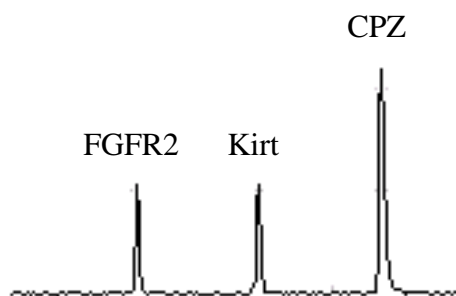
圖十二：樣本 H268 在 6p25.3/ DUSP22 gene 的 Deletion 表現



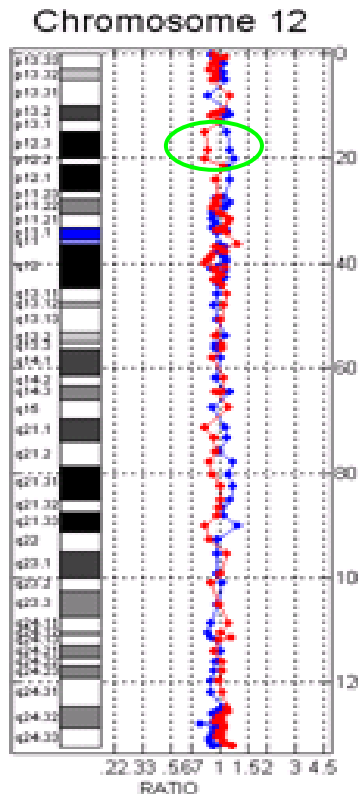
圖十三：樣本 H490 在的 Chr 4 的 aCGH 結果，4p16.1 有一 clone 為 Duplication 表現



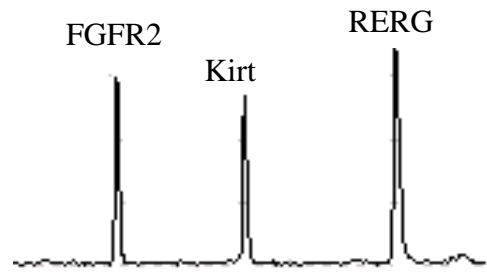
圖十四：參考檢體在 4p16.1/ CPZ gene 的劑量表現



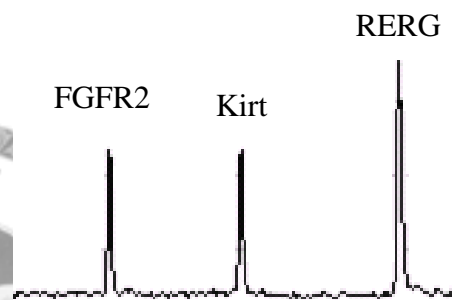
圖十五：樣本 H490 在 4p16.1/ CPZ gene 的 Duplication 表現



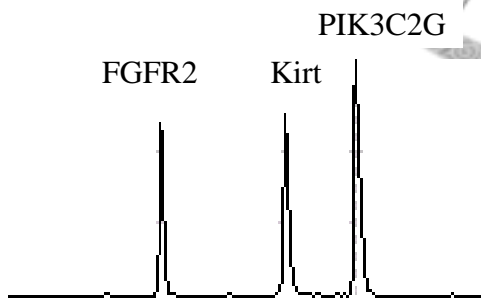
圖十六：樣本 321 在的 Chr 12 的 aCGH 結果，有 2 個 clone 為 Duplication 表現



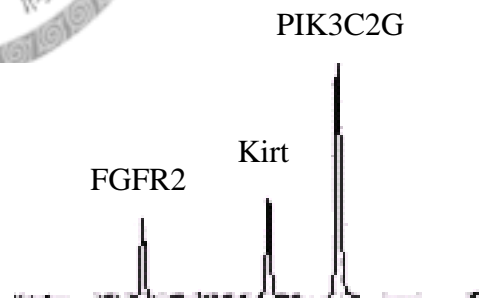
圖十七：參考檢體在 12p12.3/ RERG gene 的劑量表現



圖十八：樣本 321 在 12p12.3/ RERG gene 的 Duplication 表現



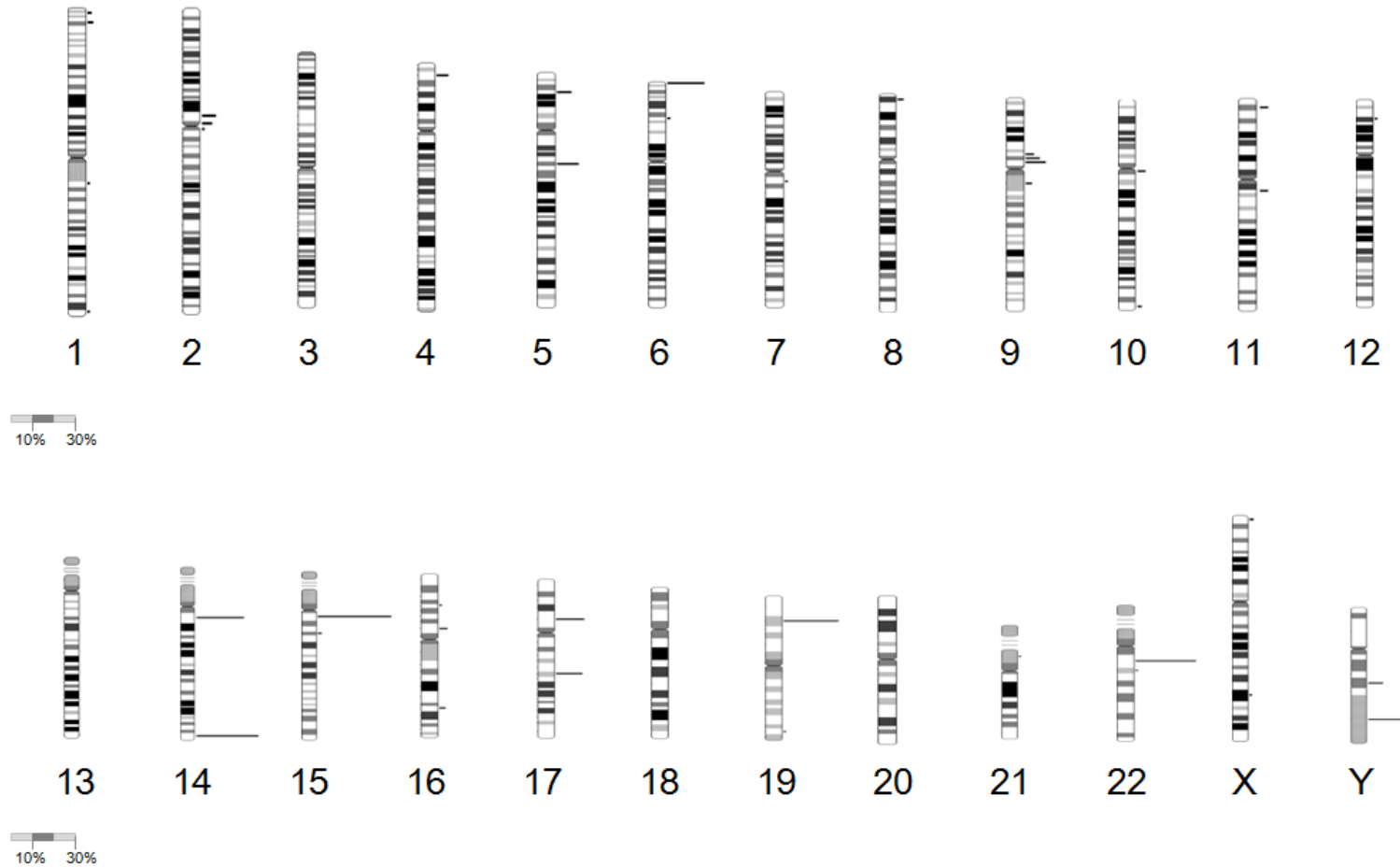
圖十九：參考檢體在 12p12.3/ PIK3C2G gene 的劑量表現



圖二十：樣本 321 在 12p12.3/ PIK3C2G gene 的 Duplication 表現

註：樣本 H268、H490 與 321 的 Multi-Plex PCR 分析結果與 Bac Array 結果一致

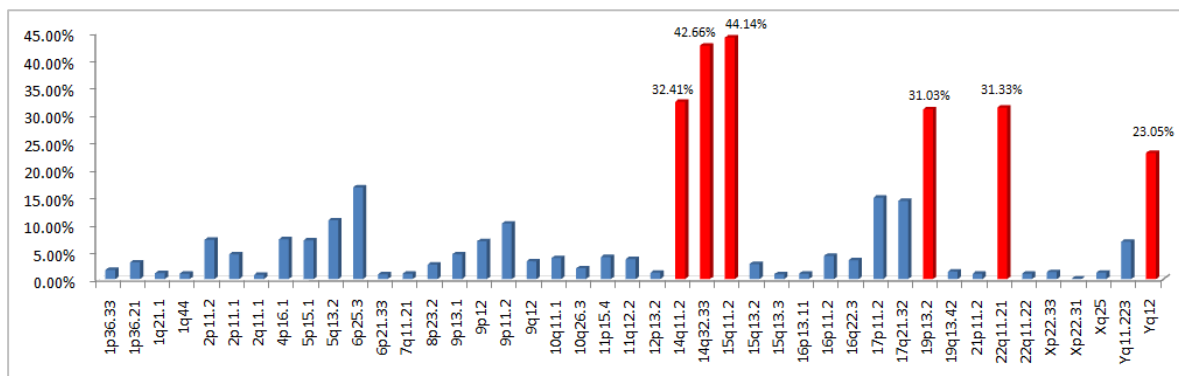
## Genomic Distribution of CNVR ( Incidence >1%)



圖二十一：台灣族群基因組劑量變異分佈



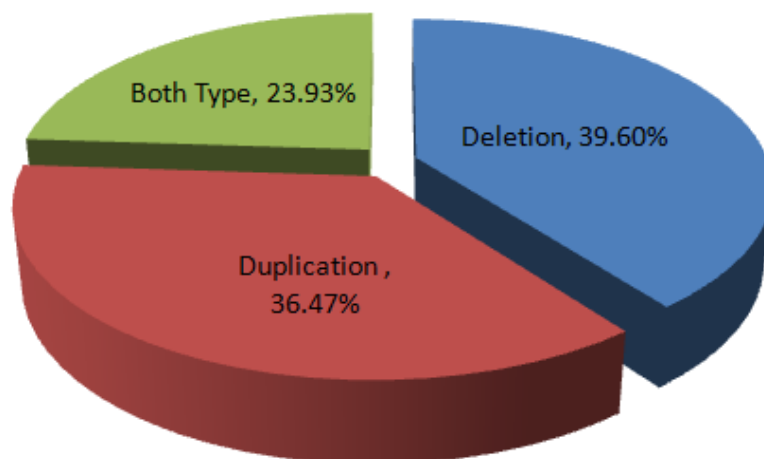
根據 CMDX Bac Array 的觀察結果，台灣族群常見的 CNV 大多都發生於 Centromere 與 Telomere 的位置，與先前文獻的觀察一致。而在台灣族群中，第 3、13、18、20、21 是比較少出現 CNV 的，最常出現 CNV 的位置是 14q11.2(32.41%)、14q32.33(42.66%)、15q11.2(44.14%)、19p13.2(31.03%)、22q11.22(31.33%)、Yq12(23.05%)等位置，這些位置在台灣族群中出現 CNV 的頻率皆超過 20%。



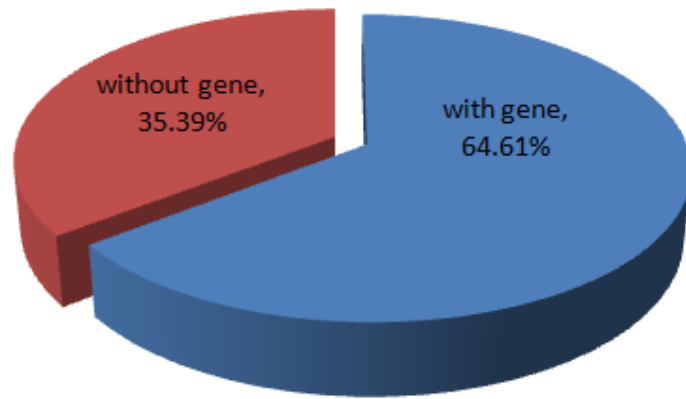
圖二十二：台灣族群常見發生基因組劑量變異的位置 (Incidence>1%)

分析於台灣族群所看到的基因組劑量變異，其中有 39.60%的區域屬於缺失的形式；36.47%的區域屬於複製的形式；而有 23.93%的區域是兩種形式兼有之。這些基因組劑量變異分佈的區域有 54.79%與基因組中已知的 Segmental Duplication 分佈的位置重複，Segmental Duplication 和基因組劑量變異的高相關性與先前文獻的觀察是一致的，顯示 Segmental Duplication 的存在及 Non-Allele Homologous Recombination 的發生是造成大片段基因組劑量變異的主要機制。這些基因組劑量變異的訊號，絕大多數(92.44%)的涵蓋範圍不超過 1~3 Bac Clones 的長度，大小不及 1Mb；只有少數(7.56%)的涵蓋範圍大於 1Mb，其主要分佈在 2p11.2~p11.1、5q13.2、9p13.1~q12、15q13.1~q13.2、15q11.2、16p13.11 等位置。甚至少數人在 9p13.1~q12 的基因組劑量變化可達 28~30Mb。而這些基因組劑量變異分佈的區域有 64.49%與已知基因的位置重複，顯示基因組劑量變異的涵蓋範圍廣泛，經由改變基因的劑量或影響調控基因表現的機制，來造成個體之間表現的多樣性。再者，於台灣族群所觀察到的 438 個 CNV 區域，其中有 405 個

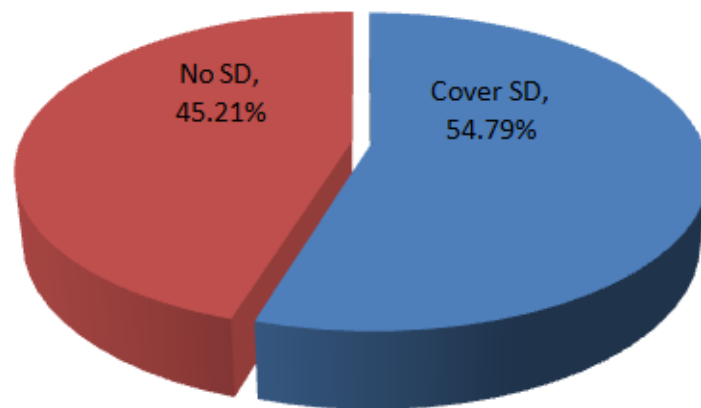
(92.47%)已在先前 Copy Number Variation 相關文獻中被報導過，或是進一步分析個案的健康雙親，確認該 CNV 屬於遺傳性的 CNV；而有 33 個(7.53%)是首次在此次研究中發現，首次發現的 CNV 中，又有 26 個（5.94%）不含有功能性基因或在研究族群中反覆出現，重複出現的位置為 1p21.3 的 RP11-89P12、2p24.3 的 RP11-794I18、4q13.1 的 RP11-91C3、5p12 的 RP11-246L1、6p22.3 的 RP1-135L22、6q22.31 的 RP11-765A3、6q24.1 的 RP1-217O16、13q31.32 的 RP11-326B4 與 14q12 的 RP11-1111O8，其 CNV 的訊號在研究族群中可在兩個人以上的個體被觀察到，推測其可能為台灣族群特有的 CNV 位置。首次發現的 CNV 中，另有 7 個(1.59%)是研究族群中罕見的 CNV，其中位於 4q21.3 的 RP11-164I21，內含 MAPK10 基因，其蛋白產物為 MAP Kinase 家族的成員之一，此蛋白在神經細胞的凋亡過程中扮演著重要的調節角色，目前 OMIM 資料庫上已知一名於 4q21.3 處發生平衡易位的個案，其 MAPK10 基因因易位而被阻斷，有癲癇的病徵；除此之外，其他 6 個罕見的 CNV 均未有相關的疾病資訊，其內含基因的主要功能為調控細胞內外的訊號傳遞，與細胞的附著(Cell Adhesion)、生長、分化、轉錄、轉譯調節或發展有關。



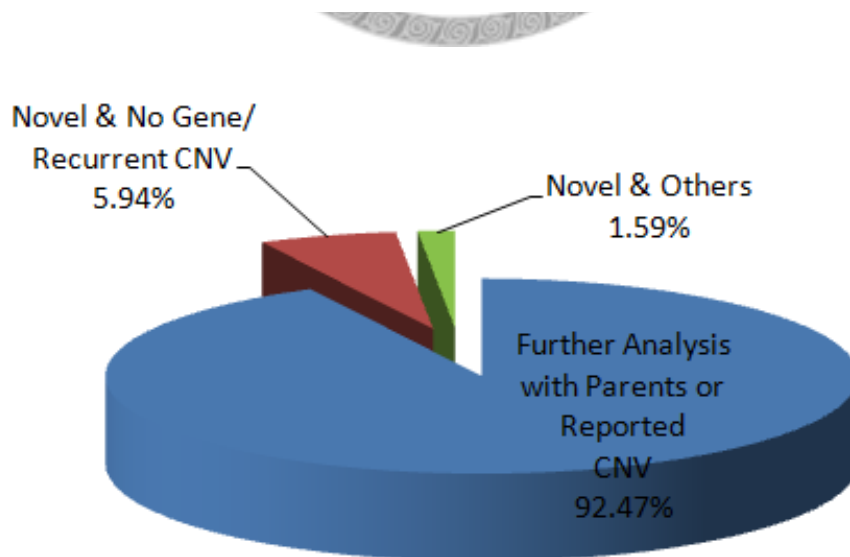
圖二十三：台灣族群全基因組劑量變異形式分析



圖二十四：台灣族群基因組劑量變異的基因覆蓋率



圖二十五：台灣族群基因組劑量變異與 Segmental Duplication 的重疊率



圖二十六：台灣族群已知 CNV 與新發現的 CNV 比例

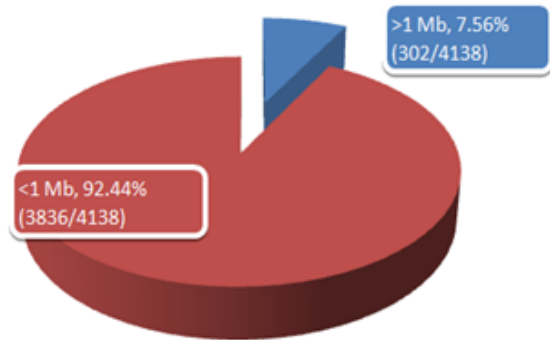
- **< 1 Mb:** Most CNVs identified are less than 1Mb (1~3 Bac Clone)

- **1~5 Mb:**

- 2p11.2p11.1 (2.13Mb)
- 5q13.2q13.2 (1.11 Mb)
- 9p13.1p11.2 (2~6 Mb)
- 15q13.1q13.2 (1.42 Mb)
- 15q11.2q11.2 (1.2~2Mb)
- 16p13.11p13.11(1.2 Mb)

- **>10 Mb**

- 9p13.1q12 (28~30 Mb)



圖二十七：台灣族群基因組劑量變異的大小變化與分佈



## 第四章 討論

研究樣本的取材來自於妊娠中的胎兒，在收集樣本的過程中已然剔除有家族遺傳病史及有相關病理診斷的來源，而所有生出的胎兒皆具有正常的表現型，如此可以排除研究中所鑑別的 CNV 與早發性、高度遺傳性(High Penetrance)疾病的相關性。然而，由於追蹤所有胎兒的生長發育直至其成年後的健康情形，有其困難度，因此研究中所鑑別的 CNV 確實無法完全排除與晚發性、複雜性疾病的相關風險。由於研究對象的測試時機為產前診斷，因此以胎兒檢體所得的 CNV Database 來說明、代表一般健康族群的 CNV 分佈，確實有其限制性。

CMDX Bac Array CA2500/ CA3000 所選用的探針大小界於 150~250 Kb，原則上要大於 85 Kb 以上的基因組劑量變化才容易被 Bac Array 平台偵測到，因此 C2500/ CA3000 平台無法鑑別較小的基因組劑量變化，而無法精確指出基因組劑量變異的起迄位置、實際大小，也是 Bac Array 平台分析基因組劑量變異的限制，因此當基因組劑量變異的大小超出了 Bac Array 平台的敏感度，或基因組劑量變異發生的位置不在 Multi-Plex PCR 所設計的探針的偵測範圍時，可能是導致 Bac Array 分析結果與 Milti-Plex PCR 結果不一致的原因。當然，人為的操作問題、螢光強度判讀上的誤差，也是造成兩種方法不一致的可能因素。

CA2500 Bac Array 平台含有探針數 2437 個，而 CA3000 平台則含有探針數 3039 個，理論上探針的密度愈高，應能提供愈豐富的基因組訊息，但 CA2500 平台所觀察到的個體平均 CNV 發生次數為 4.84 次，CA3000 平台卻略低、為 3.88 次。仔細比對 2 個平台的選用探針，發現 CA2500 平台有部份探針並未包含在 CA3000 平台中，例如 1q21.1 的 RP11-160L8、2p11.1 的 RP11-90E3、3p14.1 的 RP11-14D22 與 11p15.4 的 RP11-1031H14，而上述 4 個 Bac Clones 在 221 個以 CA2500 進行分析的樣本中，發生基因組劑量變異的頻率並不低，分別是 2.71% (6/221)、21.72% (45/221)、0.45% (1/221)、0.90% (2/221)，這應是造成 CA3000 平

台觀察到的基因組劑量變異較 CA2500 為低的主要原因。

於 Bac Array 平台所觀察到的基因組劑量變異，只出現 Deletion 訊號的 CNVR 與只出現 Duplication 訊號的 CNVR 所佔的比例沒有明顯差異，約各占 39.6% 與 36.47%；而同時兼具 Deletion、Duplication 訊號的 CNVR 則約佔了 23.93%。由於絕大多數的 CNVR 在研究族群中發生 CNV 的頻率不高，因此罕見的 CNVR 很難由少數幾個出現 CNV 的樣本來判斷該 CNVR 的變異形式究竟是僅有 Deletion 或 Duplication，但人群中高發生率的 CNVR(>1%)，則幾乎同時具有 Deletion 及 Duplication 兩種變異形式，顯示 Non-Allele Homologous Rearrangement 是促進大片段結構變異發生的主要機制，藉由同源性片段的互換、重組，因此 Deletion、Duplication 的發生、存在是相對的。

統計結果顯示 CNV 的發生與 Segmental Duplication 有極高的相關性，且經常集中於 Centromere 與 Telomere 的位置，甚至與某些致病基因的位置很接近，例如發生在 5p15 的 Cri du Chat Syndrome、5q13.2 的 Spinal Muscular Atrophy、7q11.23 的 William-Beuren Syndrome、11p15.4 微缺失導致的 Charcot-Marie-tooth disease、15q11-q13 突變引起的 Prada-Willi Syndrome/Angelman Syndrome、17p11.2 微缺失導致的 Smith-Magenies Syndrome 或 22q11.2 微缺失造成的 DiGeorge Syndrome 等，均位於 Pericentromere 的區域，這些已知的遺傳疾病，其致病位置附近的 CNV 發生率均很高，這表示 CNV 的發生並不是隨機的，而是與基因組本身的結構有關，由於重複性序列、Segmental Duplication 經常集中於 Centromere、Telomere 以及致病基因的附近，因此這些區域的基因組結構較不穩定，容易發生重組，自然也容易出現 CNV 的結構變異。當重組的 DNA 片段僅限於同源性序列或重複性序列的涵蓋範圍時，其變化的範圍不大，且不包含關鍵性的功能基因，則其重組的結果就只是造成個體之間的多型性變化；但若發生重組的片斷較大，其內含關鍵性的功能基因，則重組的結果就會造成關鍵性基因的重覆或丟失，這樣的染色體分配到配子中時，就可能使下一代出現基因表現失衡的疾病變化。

Cytogenetic Location	Taiwan Population Frequency	Disease Phenotype Associated with Region
5p15	7.09%	Cri du Chat Syndrome
5q13.2	10.74%	SMA
7q11.23	0.50%	Williams- Beuren Syndrome
11p15.4	4.04%	Charcot-Marie-Tooth Disease Type B
15q11-q13	44.14%	PWS & AS
17p11.2	14.88%	Smith-Magenis Syndrome
22q11.2	31.33%	Digeorge Syndrome

表三：在健康個體所看到的 CNV，部分座落於相關遺傳疾病的致病位置

而 CNV 對基因的高覆蓋率，顯示其對人類的遺傳變異性及表型多型性扮演重要的角色，但 CNV 並不會直接引起疾病的發生，或單一的造成某個特殊表型改變，它必須要集合多個基因的效應或接受環境因子的刺激或調控，才會導致某個複雜表型出現。例如個體間免疫力的差異、對某些疾病的易感性或風險性的差別、對某些藥物代謝或食物消化能力的不同等，都與 CNV 在基因組中存在的多型性有關。

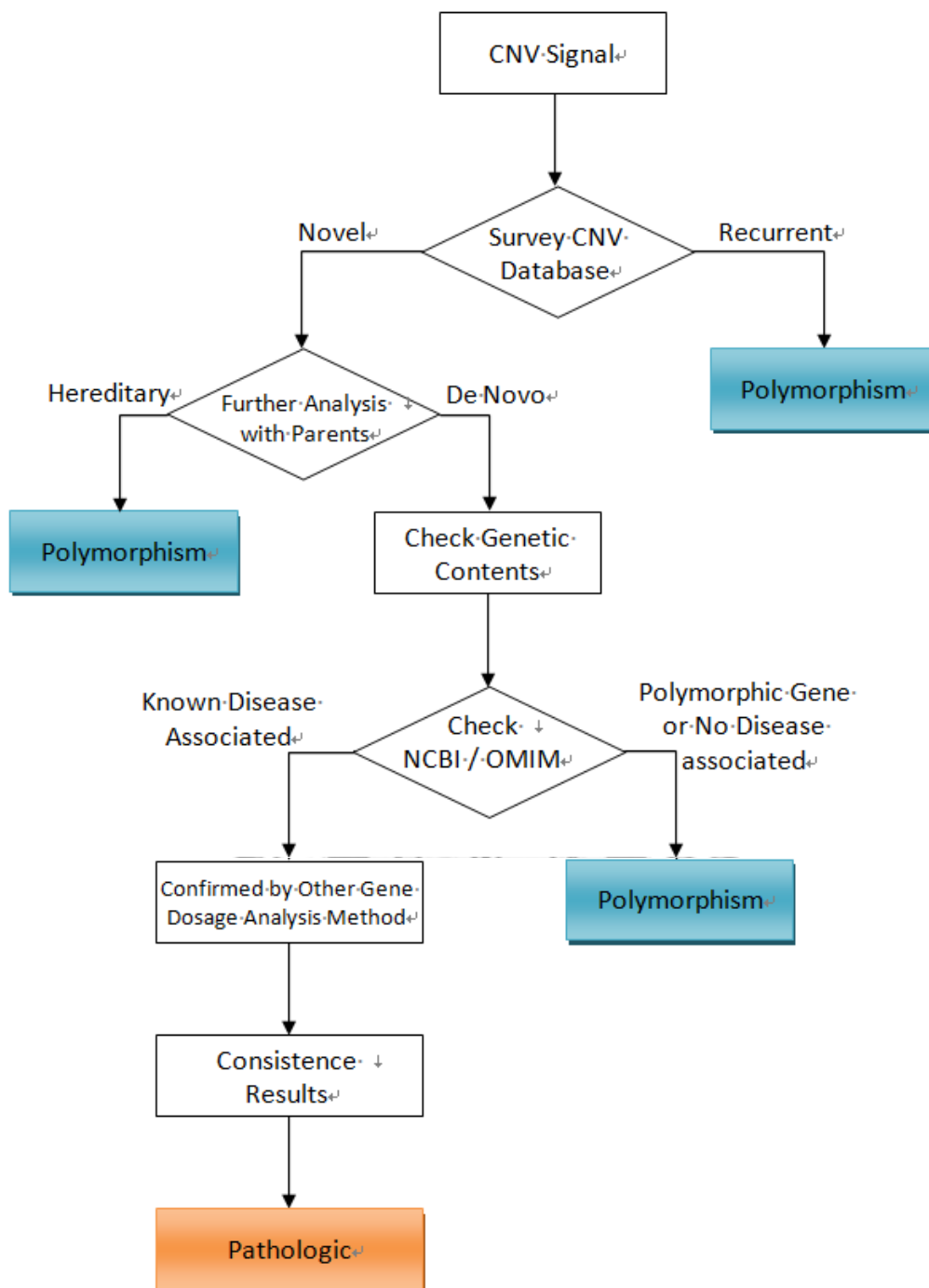
再者，大片段的基因組劑量變異(>1 Mb)，一如預期的，在基因組涵蓋的比例較低(7.30%)，涵蓋功能性基因的比例也較低，這是演化的結果，顯示大片段的結構變異承受較大的演化選擇壓力。而由 aCGH 分析結果所觀察到的大片段結構變異的好發位置，也與傳統染色體組型分析經常觀察到的 Heteromorphism 區域一致，如 9p13.1~q12、15q11~q13、16p11~p13 等位置，經常在正常、健康個體的染色體組型觀察到不同大小及染色特性。這些大片段的結構變異，通常由一些重複性的序列組成，涵蓋基因的比例較低，或是大多包含一些控制多型性表型的基因，如：嗅覺接受體基因、免疫球蛋白基因等，這也說明了這些大片段的結構變異在經過數百萬年的演化後，為何還能夠存在現行人類基因組的原因。



圖二十八：在染色體 9、15、16 所觀察到的 Heteromorphism

於研究中所發現的 438 個 CNVR，92.47% 已在先前相關的 CNV 文獻報導過，僅有 33 個 CNVR( Bac Clone)是首次在此次研究中發現，而 33 個 CNVR 中又有 26 個 CNVR 不涵蓋功能性基因，或是在研究族群中反覆出現。產前、臨床診斷上，判斷基因組中的劑量變化是否具有病理意義，已建立的 CNV 資料庫具有非常重要的參考價值，比對健康人群中是否有出現相同的劑量變化是解讀分析結果的第一步。一般而言，已被報導過的劑量變異或在研究族群中重複出現的劑量變異，原則上認同其為多型性的 CNV；若 CNV 資料庫裡查無相同的劑量變異，則建議須尋求個案的雙親做進一步的分析，以鑑別該結構變異為由健康的雙親遺傳而來，或是為一新發生(De novo)的案例。由健康的雙親遺傳而來的 CNV 可認定其為無病理意義的多型性變化；至於新發現的 CNV，則需進一步檢查該 CNV 是否涵蓋功能性基因，若不含有功能性基因，則推斷該結構變異造成病理表型的可能性極低；而涵蓋有功能性基因者，再確認基因的功能，了解其是否為多型性基因，或是 NCBI、OMIM 等疾病資料庫有無相關病例報導，若該基因屬多型性基因，或無疾病相關性，則判定其為多型性 CNV；而涵蓋的功能性基因如已知會引起相關病理表型，則宜以其他的基因劑量分析方法進一步偵測、確認功能性基因缺損或擴增的位置，並結合 NCBI、OMIM 等疾病資料庫的資訊，來提供完整的遺傳諮詢。產前、臨床診斷結構變異的流程如下圖所示。





圖二十九：多型性 CNV 的鑑別流程

註：CNV Database 來自於 DGV 資料庫及台大分子遺傳實驗室 aCGH 資料庫

將台灣族群基因組劑量變異的分析結果與先前的文獻相較，除了 2006 年 Wellcome Trust 發表的全球廣泛性的 CNV 研究，囊括了 45 個中國人的樣本、45 個日本人的樣本外，其於文獻的研究對象多集中於歐美人種，亞洲人種較少，而這些文獻基於所使用的分析方法不同、解析度不同、探針選擇不同、涵蓋基因組的比例不同、研究族群不同、樣本大小不同、參考樣本不同，使得其各自看到的 CNV、個體平均 CNV 發生次數，有很明顯的差異。原則上當方法的解析度越高時，看到的 CNV 更豐富、多樣，且 CNV 的平均長度更小。鑑於小於 50 Kb 的 CNV 佔了基因組劑量變異主要的比例，因此解析度愈高的分析方法，如探針密度愈高的 Oligo Array、SNP Array，或 DNA Sequencing 等，愈適合用來研究、鑑別族群中 CNV 的存在與分佈。但即使是高解析度的方法，也分別有其限制，如廣泛型(Whole Genome)的基因晶片雖然可以快速的掃描全基因組的劑量變異情形，但無法精確的定位核苷酸的起迄位置，所以 CNV 的大小只能做概略性的估算，同時，也無法觀察平衡性的結構變異；而基因組的定序方法其解析度可以達到核苷酸的水準，因此可以精確的指出結構變異的點位及變異 DNA 片段的實際大小，並偵測倒位等 DNA 序列倒置的結構，然而，全基因組定序曠日費時，且費用昂貴，無法在短時間內針對大規模的樣本進行分析，難以取得足以代表整個族群或不同族群的廣泛性資料，因此，研究基因組的結構變異應根據研究目的來選擇一個合適的分析平台。

表四：研究結果與相關文獻的比較 (C 表示中國人；J 表示日本人；K 表示韓國人)

Method	Sample No.	CNVR No.	CNV Size	CNV Avarage	Source
<b>Bac array</b>	<b>1011</b>	<b>445</b>	<b>150 kb~30 Mb</b>	<b>4.33</b>	<b>Current Study</b>
Bac array	55	255	150~425 Kb	12.4	lafrate(2004)
Oligo- array	20	76	465 Kb	11	Sebat (2004)
Computerization	2	241	8~40 Kb	NA	Tuzen (2005)
SD array	47	119	150~425 Kb	NA	Sharp (2005), Asian=8C+4J
Parent-offspring trios	2	586	300 bp~1.2Mb	NA	Conrad (2006)
Parent-offspring trios	2	541	1~745 Kb	NA	McCarroll (2006)
Bac array & SNP array	270	1447	341Kb/ 206 Kb	NA	Wellcome Trust (2006), Asian=45C+45J
Bac array (26363 clone)	95	3654	>40 Kb	NA	Wong (2007)
Pair end Sequencing	8	1695	15.3 Kb	NA	Kidd (2008), Asian=1C
Pair end Sequencing	1	2474	492 bp	NA	Wang(2008), Asian=1C
Pair end Sequencing	1	2920	0.1~100 Kb	NA	Ahn (2009), Asian=1 K
SNP array	2026	3272	2 bp~2.6 Mb	26.9	Shaikh (2009); Asian=12C

## 第五章 結論

依據此篇研究與先前文獻的分析結果，基本上沒有任何一種方法可以偵測基因組所有的結構變異，且依照先前文獻的發現，小於 50 Kb 的 CNV 在基因組中出現的頻率更高，所以採用解析度愈高的 Oligo-Array 或 SNP Array，愈適合用來做 CNV 的研究，然而高解析度的 Array CGH 平台雖然可以看到更多的 CNV 訊息，但在臨床報告的解釋上卻不一定是有助的，由於大多數的 CNV 與人類健康、疾病的關係仍不清楚，同時 CNV 的存在並不會直接的導致疾病發生，因此揭露大量的基因組劑量變異訊息，反而會造成遺傳諮詢的困擾，尤其絕大多數的 Mmicrodeletion/ Microduplication Syndrome 是由大於 1~2 Mb 的 DNA 片段變化所造成的，所以增加分析平台的解析度，並不會增加遺傳疾病的偵測率(33)，因此此篇研究所採用的 CMDX Bac Array，雖然解析度僅 1 Mb 左右，但其在全基因組選取適當的探針，並在相關致病基因的位置增加探針的密度，平均每個個體可觀察到的良性 CNV 次數僅為 4 次左右，比其他文獻所使用的分析平台要少得多，故這樣的分析平台較適合應用在產前或臨床診斷上，特別是當產前超音波影像分析異常，或者已知個體有異常表型，但 Karyotyping 並沒有發現異狀時，利用 Bac Array 來做進一步的檢查、診斷，是非常可靠而且有效率的，甚至可因此鑑別新的致病基因，或新的 Mmicrodeletion/ Microduplication Syndrome。

此篇研究分析了 1015 個個體，鑑別基因組劑量變異的好發位置，建立了屬於台灣族群 CNV 的圖譜，補足現行 CNV 資料庫、亞洲族群分析樣本不足、族群變異性不足的缺點，同時也發現了少數先前文獻未報導過的罕見 CNV，可以供作臨床診斷或研究的參考。

## 參考文獻

1. THOMPSON& THOMPSON GENETICS IN MEDICINE 6<sup>th</sup> edition
2. Marwan Shinawi, Sau Wai Cheung:  
**The array CGH and it's clinical application.**  
Drug Discovery Today, 2008, 9:760-770
3. Ferk L, Carson AR, Scherer SW:  
**Structural Variation in the Human Genome.**  
Nat Rev Genet, 2006, 7(2):85-97
4. Nigel P Carter:  
**Method and strategies for analyzing copy number variation using DNA microarrays.**  
Nat Genet, 2007, 7:S16-S21
5. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, *et al.*:  
**Copy number variation: new insights in genome diversity.**  
Genome Res 2006, 16(8):949-961.
6. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, *et al.*:  
**Large-scale copy number polymorphism in the human genome.**  
*Science* 2004, 305(5683):525-528
7. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C:  
**Detection of large-scale variation in the human genome.**  
*Nat Genet* 2004, 36(9):949-951.
8. McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, Gabriel SB, Lee C, Daly MJ, *et al.*:  
**Common deletion polymorphisms in the human genome.**  
*Nat Genet* 2006, 38(1):86-92.
9. B Trask, G van den Engh, B Mayall, and J W Gray:

**Chromosome heteromorphism quantified by high-resolution bivariate flow karyotyping**

Am J Hum Genet. 1989 November; 45(5): 739–752

10. Dr Ram S. Verma , Harvey Dosik , Herbert A. Lubs , Uta Francke :  
**Size and pericentric inversion heteromorphisms of secondary constriction regions (h) of chromosomes 1, 9, and 16 as detected by CBG technique in Caucasians: Classification, frequencies, and incidence**  
Am. J. Med. Genet. **2**, 331–339 (1978).
11. Hsu, L. Y., Benn, P. A., Tannenbaum, H. L., Perlis, T. E., & Carlson, A. D. :  
**Chromosomal polymorphisms of 1, 9, 16 and Y in 4 major ethnic groups : A large prenatal study**  
Am. J. Med. Genet. **26**, 95-101 (1987)
12. Charlotte N. Henrichsen, Evelyne Chaignat and Alexandre Reymond  
**Copy number variants, diseases and gene expression**  
Human Molecular Genetics **2009** 18(R1):R1-R8
13. Andrew J. Sharp, Ze Cheng, and Evan E. Eichler:  
**Structural Variation of the Human Genome**  
Human Genet. **2006**.7: 407-442.
14. Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, et al.  
**Fine-Scale Structural Variation of the Human Genome**  
Nat. Genet. **2005**:37: 727-32.
15. Sharp AJ, Loke DP, MsGrath SD, Ze Cheng, Bailey JA, et al  
**Segmental Duplication and Copy-Number Variation in the Human Genome**  
Am. J. Hum. Genet. **77**: 78-88, 2005
16. Richard Redon, Shumpei Ishikawa, Karen R. Fitch, Lars Teuk, et al  
**Global Variation in Copy Number in the Human Genome**  
Vol 444/ 23 Number 2006: 444-454
17. Melinda Wenner  
**Genetic Copy Variations and Disease**

Science American. 2009, June

18. Jennifer L. Stone, Michael C. O'Donovan, Hugh Gurling, George K. Kirov, et al  
**Rare chromosomal deletions and duplications increase risk of schizophrenia**  
Nature **455**, 237-241 (11 September 2008)
19. S.A McCarroll, Alan Huett , Petric Kuballa' Shannon D Chilewski , et al  
**Deletion polymorphism upstream of *IRGM* associated with altered *IRGM* expression and Crohn's disease**  
Nature Genetics 40, 1107 - 1112 (2008)
20. Gonzalez E., Kulkarni H., Bolivar H., Mangano A., Sanchez R., Catano G., Nibbs R.J., Freedman B.J., Quinones M.P., Bamshed M.J., et al.  
**The influence of *CCL3L1* gene—containing segmental duplications on HIV-1/AIDS susceptibility.**  
Science 2005:307:1434–1440.
21. Cristen J Willer, Elizabeth K Speliotes, Ruth JF Loos, Shengxu Li, Cecilia M Lindgren. et. al  
**Six new loci associated with body mass index highlight a neuronal influence on body weight regulation**  
Nature Genetics 2009: 41 , 25 - 34
22. Conrad,D.F., Andrews, T.D., Carter, N.P., Hurles, M.E., and Pritchard, J.K.  
**A high- resolution survey of deletion polymorphism in the human genome.**  
Nature Genetics, 2006: 38, 75-81
23. Beckmann JS, Estivill X, Antonarakis SE  
**Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability.**  
Nature Reviews Genetics(2007), **8**, 639–646
24. Wong KK, DeLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE

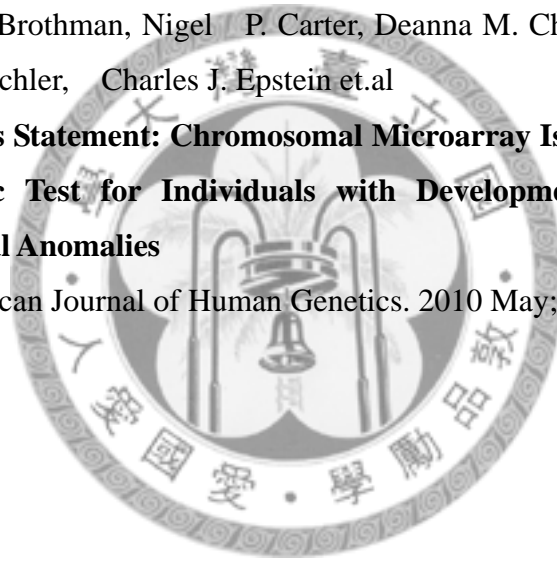
**A comprehensive analysis of common copy-number variations in the human genome.**

American Journal of Human Genetic (2007)s, **80**, 91–104.

25. Carter NP (2007)  
**Methods and strategies for analyzing copy number variation using DNA microarrays.**  
Nature, **39**, 16–21.
26. Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, Hansen N, Teague B, Alkan C, Antonacci F, Haugen E, Zerr T, et. al  
**Mapping and sequencing of structural variation from eight human genomes.**  
Nature (2008), **453**, 56–64.
27. Balakirev ES, Ayala FJ  
**Pseudogenes: are they “junk” or functional DNA?**  
Annual Review of Genetic(2003) s, **37**, 123–151.
28. Cheng Z, Ventura M, She X, Khaitovich P, Graves T, Osoegawa K, Church D, De Jong P, Wilson RK, Pääbo S, Rocchi M, Eichler EE  
**A genome-wide comparison of recent chimpanzee and human segmental duplications.**  
Nature(2005), **437**, 88–93
29. Yousheng Rao, Xiquan Zhang  
**Copy-number variant: new annotation for genetic variation across genomes**  
BIODIVERSITY SCIENCE , 2008 16(4)
30. Chia-Cheng Hung , Chien-Nan Lee , Chia-Yun Lin , Wen-Fang Cheng , Chi-An Chen , Sung-Tsang Hsieh , Chih-Chao Yang , Yuh-Jyh Jong , Yi-Ning Su, Dr. , Win-Li Lin  
**Identification of deletion and duplication genotypes of the *PMP22* gene using PCR-RFLP, competitive multiplex PCR, and multiplex ligation-dependent probe amplification: A comparison**  
Electrophoresis 2008 , 29:618-625



31. Jun Wang , Wei Wang , Ruiqiang Li , Yingrui Li , Geng Tian , Laurie Goodman , Wei Fan , Junqing Zhang , Jun Li , Juanbin Zhang et. al  
**The diploid genome sequence of an Asian individual**  
Nature. 2008 November 6; 456(7218): 60–65.
32. Shaikh TH, Gai X, Perin JC, Glessner JT, Xie H, Murphy K, O'Hara R, Casalunovo T, Conlin LK, D'Arcy M, Frackelton EC, Geiger EA, Haldeman-Englert C et. al  
**High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications**  
Genome Res. 2009 Sep; 19(9):1682-90.
33. David T. Miller, Margaret P. Adam, Swaroop Aradhya, Leslie G. Biesecker, Arthur R. Brothman, Nigel P. Carter, Deanna M. Church, John A. Crolla , Evan E. Eichler, Charles J. Epstein et.al  
**Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies**  
The American Journal of Human Genetics. 2010 May; 86: 749-764



附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr1	84	8.28%	p36.33	17	1.67%	622976	802152	RP11-206L10	5	0.08%	0.49%
						730704	893787	RP11-668E2	16	0.27%	1.58%
			p36.32	1	0.10%	4356201	4538331	RP11-111O5	1	0.02%	0.10%
			p36.22	2	0.20%	11143717	11308518	RP11-28P4	1	0.02%	0.10%
						12351275	12462906	RP4-636F13	1	0.02%	0.10%
			p36.21	31	3.05%	12753712	12934179	RP11-907J20	28	0.47%	2.76%
						13228470	13388043	RP11-204E16	23	0.38%	2.27%
			p34.2	1	0.10%	41525520	41714392	RP11-655J1	1	0.02%	0.10%
			p33	1	0.10%	47311175	47483419	RP11-624P10	1	0.02%	0.10%
						47355247	47536716	RP11-742I1	1	0.02%	0.10%
			p22.3	3	0.30%	85485485	85694513	RP11-449P23	3	0.05%	0.30%
			p21.3	3	0.30%	95708140	95861123	RP11-48A6	1	0.02%	0.10%
						99299348	99448949	RP11-89P12	2	0.03%	0.20%
			q21.1	11	1.08%	142650738	142806916	RP11-463I6	1	0.02%	0.10%
						144174221	144361891	RP11-30I17	3	0.05%	0.30%
						144706445	144862226	RP11-780O23	1	0.02%	0.10%
						147538833	147702161	RP11-160L8	6	0.10%	2.71%
q23.1	1	0.10%	155892292	156056457	RP11-68N5	1	0.02%	0.10%			
q23.3	3	0.30%	162713598	162895106	RP11-134M3	3	0.05%	0.30%			
q44	10	0.99%	244944068	245106860	RP11-102B17	1	0.02%	0.10%			
			245382614	245545837	RP11-360J3	1	0.02%	0.10%			
			245809744	245972563	RP11-77I18	1	0.02%	0.10%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
						246175941	246341428	RP11-121O22	1	0.02%	0.10%
						246768571	246944046	RP11-378I20	8	0.13%	0.79%
chr2	157	15.47%	p25.3	4	0.39%	44189	230596	RP11-664N22	2	0.03%	0.20%
						333841	497043	RP11-60K6	3	0.05%	0.30%
						883970	1026826	RP11-123C13	1	0.02%	0.10%
			p25.2	1	0.10%	4438703	4606948	RP11-59O9	1	0.02%	0.10%
			p24.3	3	0.30%	13868637	14051764	RP11-794I18	3	0.05%	0.30%
			p16.3	1	0.10%	50837941	51003975	RP11-780H1	1	0.02%	0.10%
			p16.1	1	0.10%	57007808	57160561	RP11-81L7	1	0.02%	0.10%
			p12	3	0.30%	81636068	81792654	RP11-799D9	2	0.03%	0.20%
						82146305	82293616	RP11-775D7	1	0.02%	0.10%
			p11.2	73	7.19%	87187122	87385778	RP11-1080K19	9	0.15%	0.89%
						87649620	87827692	RP11-1030A21	9	0.15%	0.89%
						89028072	89200771	RP11-483P7	52	0.87%	5.12%
						89561790	89717265	RP11-62I20	59	0.98%	5.81%
						89800710	89958566	RP11-433C18	57	0.95%	5.62%
			p11.2-p11.1	2	0.20%	90993186	91155050	RP11-641O22	2	0.03%	0.20%
			p11.1	46	4.53%	91670045	91679791	RP11-90E3	45	0.75%	20.36%
q11.1	8	0.79%	94690032	94858599	RP11-134N21	8	0.13%	0.79%			
			94953146	95132040	RP11-317D19	1	0.02%	0.10%			
q12.2	1	0.10%	105976810	106154022	RP11-798K13	1	0.02%	0.10%			
q12.3	1	0.10%	106871334	107053741	RP11-464P18	1	0.02%	0.10%			
q13	1	0.10%	111344519	111495247	RP11-1144O16	1	0.02%	0.10%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
			q14.1	1	0.10%	114990964	115147046	RP11-28I10	1	0.02%	0.10%
			q14.3	8	0.79%	125735605	125901144	RP11-11G20	8	0.13%	0.79%
			q22.2	2	0.20%	143922717	144089691	RP11-298B2	2	0.03%	0.20%
			q31.1	1	0.10%	174967371	175153504	RP11-780C22	1	0.02%	0.10%
chr3	19	1.87%	p26.3	3	0.30%	227076	227427	CTC-228K22	1	0.02%	0.10%
						943122	1120651	RP11-105L21	2	0.03%	0.20%
						1130610	1325575	RP11-696M8	1	0.02%	0.10%
			p14.2	2	0.03%	59980796	60159008	RP11-812I4	1	0.02%	0.10%
						60022762	60191344	RP11-418E13	1	0.02%	0.10%
						60556809	60707635	RP11-280N4	1	0.02%	0.10%
			p14.1	2	0.20%	64352431	64528692	RP11-14D22	1	0.02%	0.45%
						68008075	68170632	RP11-384L17	1	0.02%	0.10%
						68448766	68609004	RP11-692D4	1	0.02%	0.10%
			p12.1	1	0.10%	86540469	86711505	RP11-375C4	1	0.02%	0.10%
			q11.2	2	0.20%	95279371	95439150	RP11-259L20	1	0.02%	0.10%
						97494757	97643916	RP11-162D8	1	0.02%	0.10%
			q13.33	1	0.10%	120926865	121091380	RP11-169N13	1	0.02%	0.10%
			q26.31	3	0.30%	174574124	174732545	RP11-198D7	1	0.02%	0.10%
						176545003	176724570	RP11-292L5	2	0.03%	0.20%
			q28	1	0.10%	193557831	193718283	RP11-151M22	1	0.02%	0.10%
q29	4	0.39%	194902965	195114828	RP11-326J2	1	0.02%	0.10%			
			197126284	197313351	RP11-480A16	2	0.03%	0.20%			
p16.2-p16.1	2	0.20%	5094266	5290815	RP11-326O23	3	0.05%	0.30%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr4	97	9.56%	p16.1	74	7.29%	8520656	8687196	RP11-301J10	74	1.23%	7.29%
			p13	1	0.10%	44518741	44676568	RP11-105F21	1	0.02%	0.10%
			p11	4	0.39%	48760589	48941257	RP11-620D19	4	0.07%	0.39%
			q12	3	0.30%	55797692	55998510	RP11-504F11	3	0.05%	0.30%
			q13.1	3	0.30%	60608378	60768901	RP11-91C3	4	0.07%	0.39%
			q21.3	1	0.10%	87452063	87629591	RP11-164I21	1	0.02%	0.10%
			q28.3	1	0.10%	136343075	136488904	RP11-351L8	1	0.02%	0.10%
			q34.3	4	0.39%	179787284	179926971	RP11-630J16	3	0.05%	0.30%
			q35.1	1	0.10%	184365857	184561775	RP11-451F20	1	0.02%	0.10%
			q35.2	3	0.30%	188403361	188555573	RP11-152O24	2	0.03%	0.20%
						188879765	189042572	RP11-92A7	2	0.03%	0.20%
						189391176	189553650	RP11-1141D18	1	0.02%	0.10%
						189850975	190012097	RP11-256D4	1	0.02%	0.10%
			p15.2	1	0.10%	9664912	9839469	RP11-110O8	1	0.02%	0.10%
9816110	10020490	RP11-654J7				1	0.02%	0.10%			
15627290	15780983	RP11-135M13				1	0.02%	0.10%			
17465586	17591042	RP11-91D21				71	1.18%	7.00%			
28714511	28864672	RP11-195G17				1	0.02%	0.10%			
36941080	37112868	RP11-705K7				2	0.03%	0.20%			
41021892	41168355	RP11-722F5				1	0.02%	0.10%			
44848627	45022904	RP11-246L1				2	0.03%	0.20%			
p12	3	0.30%	45454078	45784678	RP11-97B16	1	0.02%	0.10%			
			46180429	46358622	RP11-63E15	2	0.03%	0.20%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr5	200	19.70%	q11.2	1	0.10%	57339914	57537646	RP11-941D3	1	0.02%	0.10%
			q12.1	1	0.10%	62461023	62605518	RP11-160F19	1	0.02%	0.10%
			q13.1	1	0.10%	67278115	67419794	RP11-26L9	1	0.02%	0.10%
			q13.2	109	10.74%	69364823	69448911	CTD-2041D13	86	1.43%	8.47%
						70073062	70247704	RP11-74H23	63	1.05%	6.21%
						70234685	70425356	RP11-195E2	72	1.20%	7.09%
						70296260	70475039	RP11-482B7	42	0.70%	4.14%
			q21.1	1	0.10%	100154922	100338373	RP11-143D12	1	0.02%	0.10%
			q23.3	2	0.20%	128623899	128811151	RP11-203A10	2	0.03%	0.20%
			q32	1	0.10%	145497943	145726441	RP11-876L8	1	0.02%	0.10%
q35.2	1	0.10%	172851020	173016219	RP11-346N16	1	0.02%	0.10%			
q35.3	1	0.10%	177591246	177742659	RP11-2I16	1	0.02%	0.10%			
			p25.3	170	16.75%	231044	357479	CTD-2249D1	169	2.82%	16.65%
						1381240	1568545	RP11-42D22	1	0.02%	0.10%
						1793859	1992653	RP11-1103I2	1	0.02%	0.10%
			p25.1	1	0.10%	5293472	5443772	RP11-1I14	1	0.02%	0.10%
			p22.3	3	0.30%	17955036	18121418	RP11-90M17	1	0.02%	0.10%
						21424579	21545525	RP1-135L22	2	0.03%	0.20%
			p22.1	1	0.10%	29073892	29203316	RP1-88J8	1	0.02%	0.10%
			p21.33	9	0.89%	29863159	30014486	RP11-71C20	9	0.15%	0.89%
			p21.32	1	0.10%	32843695	33023853	RP11-10A19	1	0.02%	0.10%
p12.3	2	0.20%	45964230	46145363	RP11-11H19	2	0.03%	0.20%			
p11.2	1	0.10%	57501590	57681850	RP11-637D3	1	0.02%	0.10%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr6	202	19.90%	q11.1	1	0.10%	62379845	62576993	RP11-164C22	1	0.02%	0.10%
			q12	1	0.10%	69911578	70090135	RP11-622F12	1	0.02%	0.10%
			q16.1	1	0.10%	92756130	92944280	RP11-451J5	1	0.02%	0.10%
						96397095	96562812	RP11-386J20	1	0.02%	0.10%
			q21	1	0.10%	107812833	107967471	RP1-67A8	1	0.02%	0.10%
			q22.1	1	0.10%	115088963	115242892	RP11-122B15	1	0.02%	0.10%
			q22.31	2	0.20%	119974543	120136481	RP11-765A3	2	0.03%	0.20%
			q24.1	2	0.20%	140635836	140723070	RP1-217O16	2	0.03%	0.20%
			q25.2	1	0.10%	153821550	153994956	RP11-198L11	1	0.02%	0.10%
			q26	1	0.10%	162343450	162535020	RP11-456M6	1	0.02%	0.10%
			q27	3	0.30%	168170160	168359681	RP11-114E13	1	0.02%	0.10%
170655599	170792198	CTD-2011D14				2	0.03%	0.20%			
			p22.2	1	0.10%	3026370	3185513	RP11-234N20	1	0.02%	0.10%
			p21.3	1	0.10%	12020181	12160876	RP11-79O21	1	0.02%	0.10%
			p21.1	1	0.10%	17213145	17376129	RP11-1007J4	1	0.02%	0.10%
			p14.1	1	0.10%	42264550	42467165	RP11-429N13	1	0.02%	0.10%
			p12.3	1	0.10%	46811922	46974907	RP11-51N15	1	0.02%	0.10%
			p12.1	1	0.10%	52601653	52771400	RP11-871A24	1	0.02%	0.10%
			p11.2	1	0.10%	54607136	54784743	RP11-449G3	1	0.02%	0.10%
			q11.1-q11.21	7	0.69%	61096752	61270745	RP11-1025O20	7	0.12%	0.69%
			q11.21	10	0.99%	61608951	61789025	RP11-626C15	1	0.02%	0.10%
						62966021	63119691	RP11-248G15	1	0.02%	0.10%
64668252	64771093	RP11-644D10				6	0.10%	0.59%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr7	52	5.12%				64708585	64864970	RP11-489G16	3	0.05%	0.30%
						65966133	66119326	RP11-214K14	1	0.02%	0.10%
			q11.22-q11.23	4	0.39%	71658313	71830847	RP11-35P20	4	0.07%	0.39%
			q11.23	5	0.49%	73343780	73537332	RP11-351B3	1	0.02%	0.10%
						73785125	73944776	RP11-89A20	1	0.02%	0.10%
						74852506	75029107	RP11-605K23	2	0.03%	0.20%
						75967888	76118659	RP11-1113B10	2	0.03%	0.20%
						76490995	76687312	RP11-467H10	1	0.02%	0.10%
			q21.13	2	0.20%	89163364	89305339	RP11-17A16	2	0.03%	0.20%
			q22.2	1	0.10%	105004303	105185493	RP11-135H16	1	0.02%	0.10%
			q22.3	1	0.10%	106485513	106660211	RP11-120N6	1	0.02%	0.10%
			q31.1	7	0.69%	108768839	108908853	RP11-590B16	1	0.02%	0.10%
						110773231	110940607	RP11-635E6	6	0.10%	0.59%
			q31.31	2	0.20%	118727269	118925701	RP11-590F14	2	0.03%	0.20%
			q36.2	5	0.49%	153024535	153229490	RP11-105P6	3	0.05%	0.30%
						153229742	153389164	RP11-30J12	4	0.07%	0.39%
153507185	153713711	RP11-1073N5				1	0.02%	0.10%			
153585437	153735765	RP11-317C13				1	0.02%	0.10%			
q36.3	1	0.10%	158536683	158689697	RP11-133J16	1	0.02%	0.10%			
				27	2.66%	2303777	2411033	RP11-104H14	22	0.37%	2.17%
						3849244	4007372	RP11-243D4	4	0.07%	0.39%
						4218952	4386750	RP11-113B7	1	0.02%	0.10%
						4789354	4938784	RP11-96P9	1	0.02%	0.10%



附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)	
chr8	44	4.33%				5271311	5431541	RP11-36I5	1	0.02%	0.10%	
						5822620	5966061	RP11-95C8	2	0.03%	0.20%	
			p21.3	1	0.02%	19740415	19927437	RP11-402K21	1	0.02%	0.10%	
			p21.2	2	0.20%		24912114	25057768	RP11-161N9	1	0.02%	0.10%
							25119935	25301066	RP11-652E5	1	0.02%	0.10%
			p11.1	6	0.59%		25245698	25406408	RP11-76B12	1	0.02%	0.10%
							43602365	43764092	RP11-467N1	5	0.08%	0.49%
							43735846	43895573	RP11-32P10	5	0.08%	0.49%
			q11.1	2	0.20%	47019365	47126364	RP11-970F14	2	0.03%	0.20%	
			q11.21	2	0.20%		49230594	49231494	RP11-268N2	1	0.02%	0.10%
							49626944	49794697	RP11-208I6	1	0.02%	0.10%
							50257354	50395540	RP11-10H3	2	0.03%	0.20%
			q12.1	1	0.10%	58007449	58157371	RP11-60J19	1	0.02%	0.10%	
			q21.13	1	0.10%	81884881	82095247	RP11-819O6	1	0.02%	0.10%	
q21.3	1	0.10%	90197490	90371703	RP11-99K4	1	0.02%	0.10%				
q24.21	1	0.10%	130828111	130999900	RP11-79E8	1	0.02%	0.10%				
p24.3	2	0.20%		264605	419765	RP11-1112G24	2	0.03%	0.20%			
				442177	642967	RP11-675G5	2	0.03%	0.20%			
				615306	812246	RP11-130C19	2	0.03%	0.20%			
p24.1	3	0.30%		4979026	5111688	RP11-42A14	2	0.03%	0.20%			
				8228656	8410865	RP11-290M6	1	0.02%	0.10%			
p13.3	1	0.10%	33988396	34157504	RP11-148N9	1	0.02%	0.10%				
p13.1	46	4.53%		39766263	39963954	RP11-927H19	46	0.77%	4.53%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr9	268	26.40%	p12	70	6.90%	41071154	41307321	RP11-797M14	59	0.98%	5.81%
						41581167	41740982	RP11-105J4	49	0.82%	4.83%
						42198320	42370032	RP11-86I14	47	0.78%	4.63%
			p11.2	103	10.15%	43662586	43856412	RP11-723O15	86	1.43%	8.47%
						44179849	44373479	RP11-984C16	65	1.08%	6.40%
						44666649	44848283	RP11-160N1	33	0.55%	3.25%
						45506659	45680377	RP11-435A4	26	0.43%	2.56%
			q12	33	3.25%	66426891	66602753	RP11-318K12	15	0.25%	1.48%
						67221726	67374086	RP11-76I5	15	0.25%	1.48%
						68635600	68779111	RP11-1114I8	10	0.17%	0.99%
						69796362	69975287	RP11-154P18	23	0.38%	2.27%
			q21.2	1	0.10%	79838950	79911276	RP11-795D8	1	0.02%	0.10%
			q21.31	1	0.10%	81585778	81746774	RP11-31D18	1	0.02%	0.10%
			q21.32	1	0.10%	84501653	84667659	RP11-122C19	1	0.02%	0.10%
			q22.1	3	0.30%	89865365	90040361	RP11-106L15	3	0.05%	0.30%
			9q34.11-q34.12	2	0.20%	132470385	132643715	RP11-57C19	2	0.03%	0.20%
q34.2	1	0.10%	135572168	135778033	RP11-451N3	1	0.02%	0.10%			
q34.3	1	0.10%	137706134	137897355	RP11-498D23	1	0.02%	0.10%			
p15.3	4	0.39%	214538	366264	RP11-10D13	4	0.07%	0.39%			
			461744	638615	RP11-1077M19	4	0.07%	0.39%			
p11.21	1	0.10%	34651099	34842954	RP11-452H8	1	0.02%	0.10%			
			35203904	35380538	RP11-383B9	1	0.02%	0.10%			
p11.1	6	0.59%	38946155	39125182	RP11-109B16	6	0.10%	0.59%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr10	74	7.29%	q11.1	39	3.84%	41704729	41716128	RP11-81J3	39	0.65%	3.84%
			q21.1	2	0.20%	59260247	59432213	RP11-237L11	2	0.03%	0.20%
			q21.2	1	0.10%	62550850	62706448	RP11-79H15	1	0.02%	0.10%
			q24.32	1	0.10%	103291815	103442009	RP11-68M5	1	0.02%	0.10%
						103323201	103508676	RP11-765O11	1	0.02%	0.10%
			q26.3	20	1.97%	133629686	133778293	RP11-45A17	1	0.02%	0.10%
						134340939	134548854	RP11-288G11	1	0.02%	0.10%
						135067049	135251707	RP11-431G18	15	0.25%	1.48%
135079030	135240347	RP11-108K14				18	0.30%	1.77%			
chr11	112	11.03%	p15.5-p15.4	7	0.69%	2759787	2871057	CTD-2226C18	7	0.12%	0.69%
			p15.4	41	4.04%	4168465	4353116	RP11-1118C1	39	0.65%	3.84%
						3190460	3319716	RP11-1031H14	2	0.03%	0.90%
			p12	1	0.10%	40322149	40503595	RP11-1068E11	1	0.02%	0.10%
						40455382	40654414	RP11-768P16	1	0.02%	0.10%
			p11.2	1	0.10%	44196265	44375216	RP11-280B8	1	0.02%	0.10%
			p11.2-p11.12	1	0.10%	48641632	48820012	RP11-463K8	1	0.02%	0.10%
			p11.12	2	0.20%	51245305	51433879	RP11-828P18	2	0.03%	0.20%
			q11	3	0.30%	54983642	55174873	RP11-475G4	3	0.05%	0.30%
			q12.1	7	0.69%	58565900	58765261	RP11-1012N1	7	0.12%	0.69%
			q12.2	37	3.65%	60573957	60758403	RP11-465F2	37	0.62%	3.65%
			q12.3	1	0.10%	62458891	62622065	RP11-488C4	1	0.02%	0.10%
			q14.1	2	0.20%	81471789	81648504	RP11-1E8	2	0.03%	0.20%
q14.3	1	0.10%	91047295	91156775	RP11-74L24	1	0.02%	0.10%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
			q22.1-q22.2	1	0.10%	101500552	101687165	RP11-90M3	1	0.02%	0.10%
			q22.2	3	0.30%	101724663	101939112	RP11-315O6	3	0.05%	0.30%
			q23.3	3	0.30%	118829889	118992228	RP11-62A14	3	0.05%	0.30%
			q24.11	1	0.10%	120907159	121086046	RP11-349K14	1	0.02%	0.10%
chr12	31	3.05%	p13.33	1	0.10%	152513	329664	RP11-283I3	1	0.02%	0.10%
			p13.31	6	0.59%	7918347	8062130	RP11-277E18	6	0.10%	0.59%
			p13.2	12	1.18%	10845115	11018538	RP11-144O23	1	0.02%	0.10%
						11321141	11480339	RP11-90N7	12	0.20%	1.18%
			p12.1	1	0.10%	23903839	24058953	RP11-162F10	1	0.02%	0.10%
			p11.22	1	0.10%	28345839	28516515	RP11-924M22	1	0.02%	0.10%
			p11.21	1	0.10%	33170627	33333327	RP11-460N10	1	0.02%	0.10%
			p11.1	1	0.10%	34127734	34291672	RP11-313F23	1	0.02%	0.10%
			q12	1	0.10%	38344092	38526440	RP11-553H9	1	0.02%	0.10%
			q14.2	1	0.10%	63110965	63280538	RP11-60O8	1	0.02%	0.10%
			q21.32	1	0.10%	85151469	85314955	RP11-58K21	1	0.02%	0.10%
			q21.33	1	0.10%	89933545	90114298	RP11-917O5	1	0.02%	0.10%
			q24.11	1	0.10%	108715934	108907571	RP11-319B22	1	0.02%	0.10%
			q24.32	1	0.10%	127319957	127481773	RP11-1103G10	1	0.02%	0.10%
			q24.33	2	0.20%	130310708	130471421	RP11-146E8	1	0.02%	0.10%
					0.00%	131366783	131574912	RP11-357I22	1	0.02%	0.10%
			q12.11	1	0.10%	18729484	18809111	CTD-3053A5	1	0.02%	0.10%
			q12.12	2	0.20%	23756529	23919964	RP11-521E7	2	0.03%	0.20%
			q21.1	1	0.10%	52892079	53074594	RP11-454H21	1	0.02%	0.10%

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr13	15	1.48%	q21.2	1	0.10%	60785575	60972894	RP11-509O24	1	0.02%	0.10%
			q21.31	1	0.10%	63500884	63656227	RP11-10J15	1	0.02%	0.10%
			q21.32	5	0.49%	65078855	65222198	RP11-148N3	2	0.03%	0.20%
						66064171	66234235	RP11-1007P20	1	0.02%	0.10%
			q34	4	0.39%	66614909	66786429	RP11-326B4	2	0.03%	0.20%
						112739483	112940263	RP11-326K11	2	0.03%	0.20%
						112777833	112939731	RP11-98F14	3	0.05%	0.30%
			113805195	113967059	RP11-262A6	1	0.02%	0.10%			
chr14	777	76.55%	q11.2	329	32.41%	19273689	19485419	RP11-831B15	279	4.65%	27.49%
						20341345	20523996	RP11-89F2	2	0.03%	0.20%
						21474792	21636629	RP11-170P20	9	0.15%	0.89%
						21636748	21813687	RP11-246A2	35	0.58%	3.45%
						21930249	22109381	RP11-990K12	44	0.73%	4.33%
			q12	4	0.39%	22138643	22343786	RP11-1083M21	1	0.02%	0.10%
						24722688	24860156	RP11-180O19	1	0.02%	0.10%
						25297721	25474958	RP11-264E8	1	0.02%	0.10%
						25791020	25986265	RP11-997B23	1	0.02%	0.10%
						26798461	26947540	RP11-194P13	1	0.02%	0.10%
						29293930	29458342	RP11-1111O8	2	0.03%	0.20%
			q21.1-q21.2	1	0.10%	40873782	41029779	RP11-1134E12	1	0.02%	0.10%
			q21.2	8	0.79%	41905205	42060712	RP11-117P10	8	0.13%	0.79%
			q24.3	1	0.10%	75570349	75733729	RP11-71M17	1	0.02%	0.10%
q31.2	1	0.10%	83227463	83392181	RP11-799D22	1	0.02%	0.10%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
			q32.33	433	42.66%	105265054	105433202	RP11-448N5	3	0.05%	0.30%
						105476398	105645662	RP11-259B19	55	0.92%	5.42%
						105858039	105929313	CTD-3039D5	427	7.11%	42.07%
						105912193	106065910	RP11-683L4	22	0.37%	2.17%
						105991433	106084983	RP11-379F22	14	0.23%	1.38%
						106099276	106230230	RP11-76N15	10	0.17%	0.99%
						106117962	106301951	CTD-2011A5	12	0.20%	1.18%
						106224623	106318106	CTD-2313E3	8	0.13%	0.79%
chr15	499	49.16%	q11.2	448	44.14%	18717261	18874512	RP11-90P3	268	4.46%	26.40%
						19106050	19287262	RP11-624A1	265	4.41%	26.11%
						19203738	19367754	RP11-138C5	245	4.08%	24.14%
						19763871	19914846	RP11-415B7	374	6.23%	36.85%
						20361407	20536752	RP11-80H14	5	0.08%	0.49%
						20523632	20703023	RP11-452H4	3	0.05%	0.30%
						20732386	20912999	RP11-228M15	1	0.02%	0.10%
						21439045	21638078	RP11-373J1	1	0.02%	0.10%
						22235690	22401831	RP11-104J9	19	0.32%	1.87%
			q13.1	5	0.49%	26185054	26346267	RP11-1107G17	3	0.05%	0.30%
						27029453	27209685	RP11-650N15	1	0.02%	0.10%
						27121297	27288357	RP11-150A20	1	0.02%	0.10%
						27297142	27481488	RP11-718G7	1	0.02%	0.10%
						27573730	27748282	RP11-300A12	2	0.03%	0.20%
27740710	27905905	RP11-92F7	1	0.02%	0.10%						

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)		
			q13.1-q13.2	1	0.10%	27830813	28024200	RP11-360J18	1	0.02%	0.10%		
			q13.2	28	2.76%	28290234	28449080	RP11-1122O13	28	0.47%	2.76%		
			q13.3	11	1.08%	30179961	30339317	RP11-115G22	11	0.18%	1.08%		
								30654159	30830644	RP11-420B6	1	0.02%	0.10%
			q21.2	1	0.10%	48810464	48983569	RP11-846K6	1	0.02%	0.10%		
			q22.1-q22.2	2	0.20%	57065159	57246081	RP11-306H6	2	0.03%	0.20%		
			q26.2	2	0.20%	94552710	94747821	RP11-134D15	1	0.02%	0.10%		
								92491963	92661067	RP11-109P8	1	0.02%	0.10%
			q26.3	1	0.10%	99707174	99872903	RP11-497M17	1	0.02%	0.10%		
								99747135	99914584	RP11-530H6	1	0.02%	0.10%
			p13.3	8	0.79%	266993	431489	RP11-243K18	1	0.02%	0.10%		
								990274	1170981	RP11-656J16	1	0.02%	0.10%
								2084432	2304090	RP11-846C9	1	0.02%	0.10%
								2416830	2417096	CTD-2117I2	3	0.05%	0.30%
						2552643	2730046	CTD-2270M10	2	0.03%	0.20%		
			p13.13	1	0.10%	11420097	11635674	RP11-788A3	1	0.02%	0.10%		
			p13.11	9	0.89%	15110713	15273942	RP11-97N13	1	0.02%	0.10%		
								15519951	15681435	RP11-81L19	1	0.02%	0.10%
								15681614	15882688	RP11-629E8	2	0.03%	0.20%
								15790648	15973627	RP11-844E20	1	0.02%	0.10%
						16112598	16307524	RP11-963L5	8	0.13%	0.79%		
			p12.3	1	0.10%	18035829	18197022	RP11-758H24	1	0.02%	0.10%		
			p12.2	8	0.70%	21354957	21532552	RP11-98D10	8	0.13%	0.79%		

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr16	120	11.82%	p12.2	1	0.17%	21509195	21675241	RP11-120F20	1	0.02%	0.10%
			p12.2-p12.1	3	0.30%	21655258	21833696	RP11-449K24	3	0.05%	0.30%
			p12.1	3	0.30%	22338081	22508571	RP11-368J21	2	0.03%	0.20%
						27178754	27319548	RP11-103D22	1	0.02%	0.10%
			p11.2	43	4.24%	28058998	28203578	RP11-915B15	4	0.07%	0.39%
						29551001	29714202	RP11-180O13	3	0.05%	0.30%
						29683738	29869247	RP11-301D18	2	0.03%	0.20%
						29893466	30059319	RP11-114A14	2	0.03%	0.20%
						29981473	30174456	RP11-455F5	2	0.03%	0.20%
						31957577	32156413	RP11-1072P17	14	0.23%	1.38%
						32562118	32773200	RP11-499D5	33	0.55%	3.25%
			33556226	33754381	RP11-347N4	17	0.28%	1.67%			
			p11.1	5	0.49%	34476299	34626960	RP11-80F22	5	0.08%	0.49%
q22.1	3	0.30%	65253655	65412816	RP11-114M2	1	0.02%	0.10%			
			65946409	66146378	RP11-433A16	2	0.03%	0.20%			
q22.3	35	3.45%	72892550	73055468	RP11-252A24	35	0.58%	3.45%			
q23.3	1	0.10%	81230729	81407271	RP11-606H18	1	0.02%	0.10%			
chr17	313	30.81%	p13.3	4	0.39%	816310	976400	CTD-2237N10	3	0.05%	0.30%
						2533625	2726771	RP11-300D15	1	0.02%	0.10%
			p11.2	151	14.88%	18948013	19165438	RP11-816G17	150	2.50%	14.78%
						20492969	20692357	RP11-142H5	2	0.03%	0.20%
			p11.2-p11.1	1	0.10%	21967843	22110995	RP11-125D2	1	0.02%	0.10%
			q11.2	1	0.10%	26246069	26422085	RP11-525H19	1	0.02%	0.10%



附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr17	315	50.64%	q12	3	0.30%	34003230	34174152	RP11-607B2	3	0.05%	0.30%
			q12-q21.1	1	0.10%	35207013	35434692	RP11-387H17	1	0.02%	0.10%
			q21.31	1	0.10%	41349283	41528260	RP11-111I23	1	0.02%	0.10%
			q21.32	145	14.29%	42087498	42244563	RP11-79O18	145	2.42%	14.29%
			q22	5	0.49%	48434265	48595724	RP11-248L3	5	0.08%	0.49%
			q24.11	1	0.10%	60311024	60488913	RP11-829P10	1	0.02%	0.10%
chr18	4	0.39%	p11.31	1	0.10%	6269078	6418978	RP11-19E14	1	0.02%	0.10%
			q11.1-q11.2	1	0.10%	17267518	17448560	RP11-311F3	1	0.02%	0.10%
			q21.31	1	0.10%	53192567	53367231	RP11-79L5	1	0.02%	0.10%
				1	0.10%	73414443	73593637	RP11-90L3	1	0.02%	0.10%
chr19	338	33.30%	p13.3	3	0.30%	902925	1095436	RP11-75H6	1	0.02%	0.10%
						1331140	1495632	RP11-317H11	1	0.02%	0.10%
						1541695	1585189	G248P8411D6	1	0.02%	0.10%
						1560821	1602142	G248P81027G3	2	0.03%	0.20%
			p13.2	315	31.03%	7789817	7950264	CTD-3193O13	3	0.05%	0.30%
						8721973	8863829	RP11-79F15	314	5.23%	30.94%
			p12	2	0.20%	20455617	20644162	RP11-106E12	1	0.02%	0.10%
						24107907	24108553	RP11-89N18	1	0.02%	0.10%
			q12	2	0.20%	34827680	35023872	RP11-345J21	1	0.02%	0.10%
						36556817	36742905	RP11-18J6	1	0.02%	0.10%
q13.11	2	0.20%	38449070	38638734	RP11-939O22	1	0.02%	0.10%			
			39390486	39546101	RP11-31D10	1	0.02%	0.10%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
			q13.42	14	1.38%	59909223	60083800	RP11-155P5	14	0.23%	1.38%
chr20	6	0.59%	p13	1	0.10%	2989984	3165774	RP11-520F6	1	0.02%	0.10%
			p12.3	1	0.10%	8035960	8174995	RP11-26K18	1	0.02%	0.10%
			p11.23	1	0.10%	20252074	20440763	RP11-470C13	1	0.02%	0.10%
			p11.21	1	0.10%	25459775	25610059	RP11-109I22	1	0.02%	0.10%
			q13.33	2	0.20%	60192432	60385082	RP11-157P1	2	0.03%	0.20%
chr21	22	2.17%	p11.2	10	0.99%	9769099	9959643	RP11-430M17	10	0.17%	0.99%
			p11.1	6	0.59%	10052189	10187262	RP11-24D17	6	0.10%	0.59%
			q11.2	1	0.10%	14590704	14724640	RP11-1144F24	1	0.02%	0.10%
						14647682	14865524	RP11-846C20	1	0.02%	0.10%
						14765093	14925811	RP11-392H8	1	0.02%	0.10%
			q21.1	2	0.20%	21984344	22147565	RP11-990C8	1	0.02%	0.10%
						22438185	22617050	RP11-203K3	1	0.02%	0.10%
			q22.3	3	0.30%	45594025	45808493	RP11-345F15	3	0.05%	0.30%
			q11.21	318	31.33%	16756605	16926481	RP11-168B14	1	0.02%	0.10%
						17195001	17392290	RP11-690P21	112	1.87%	11.03%
						17448597	17639741	RP11-71J20	3	0.05%	0.30%
						18002551	18171177	RP11-652F11	2	0.03%	0.20%
						18073347	18286507	RP11-316L10	1	0.02%	0.10%
						18259448	18440797	RP11-260L7	1	0.02%	0.10%
						18859172	18981817	RP11-586I18	320	5.33%	31.53%
						19310295	19458332	RP11-155F20	7	0.12%	0.69%
			19565460	19711786	RP11-1116M14	1	0.02%	0.10%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
chr22	339	33.40%	q11.22	10	0.99%	20579317	20752897	RP11-481L22	10	0.17%	0.99%
			q11.23	2	0.20%	22415519	22596122	RP11-124F9	1	0.02%	0.10%
						22800851	22995231	RP11-76E8	2	0.03%	0.20%
						22871034	23040012	CTD-2338K23	1	0.02%	0.10%
			q13.1	3	0.30%	37259048	37435352	RP11-59O4	2	0.03%	0.20%
						37642077	37856194	RP11-1033I2	1	0.02%	0.10%
			q13.31	2	0.20%	43960327	44128818	RP11-78A15	1	0.02%	0.10%
						46658255	46820515	RP11-66M5	1	0.02%	0.10%
			q13.32-q13.33	1	0.02%	48083933	48278974	RP11-98J21	1	0.02%	0.10%
			q13.33	3	0.05%	48216045	48259549	G248P83973A7	2	0.03%	0.20%
						48255788	48295808	G248P86144C3	2	0.03%	0.20%
						48393298	48566243	RP11-697C11	1	0.02%	0.10%
						49406228	49444089	G248p85566F9	1	0.02%	0.10%
			p22.33	13	1.28%	272927	407714	RP11-91D5	1	0.02%	0.10%
402169	547936	CTD-3128A11				2	0.03%	0.20%			
433104	614144	RP11-800K15				1	0.02%	0.10%			
482453	619358	CTD-2584M8				1	0.02%	0.10%			
1109108	1328943	RP11-892B14				1	0.02%	0.10%			
1782036	1929740	RP11-74L17				5	0.08%	0.49%			
2201001	2375939	RP11-640A16				5	0.08%	0.49%			
2689420	2873864	RP11-802I24				1	0.02%	0.10%			
3147893	3332868	RP11-809F20				1	0.02%	0.10%			
			6513904	6692138	RP11-294K6	8	0.13%	0.79%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)			
chrX	49	4.83%	p22.31	10	0.99%	6958009	7136799	RP11-483M24	8	0.13%	0.79%			
						7455636	7633868	RP11-323F16	1	0.02%	0.10%			
			p21.3	1	0.10%	28360036	28535208	RP11-487M22	1	0.02%	0.10%			
			p11.21-p11.1	2	0.20%	56571345	56751495	RP11-431N15	2	0.03%	0.20%			
			q11.1	1	0.10%	63953972	64102032	RP11-922E2	1	0.02%	0.10%			
			q12	1	0.10%	66222560	66389922	RP11-109N24	1	0.02%	0.10%			
			q21.33-q22.1	1	0.10%	98110888	98259699	RP11-91K14	1	0.02%	0.10%			
			q22.2	2	0.20%	102918729	103115179	RP11-771H14	2	0.03%	0.20%			
			q22.3	1	0.10%	109102727	109309460	RP11-471A8	1	0.02%	0.10%			
			q24	1	0.10%	120338591	120511529	RP11-893C7	1	0.02%	0.10%			
			q25	12	1.18%	125404777	125555340	RP11-664P12	12	0.20%	1.18%			
			q28	3	0.30%				148315715	148484642	RP11-195I5	1	0.02%	0.10%
									148935632	149102282	RP11-792M20	1	0.02%	0.10%
									149154798	149306899	RP11-221G13	1	0.02%	0.10%
150987321	151147751	RP11-110C24							1	0.02%	0.10%			
151701063	151881428	RP11-665O24							1	0.02%	0.10%			
154597372	154763617	RP11-722P3							2	0.03%	0.20%			
chrY	311	30.64%	p11.31	5	0.49%	2500466	2690572	RP11-990G10	1	0.02%	0.10%			
						2530482	2721249	RP11-639O7	1	0.02%	0.10%			
						2550572	2721250	RP11-975F3	1	0.02%	0.10%			
			p11.2	3	0.30%				10193916	10345898	RP11-1097O22	1	0.02%	0.10%
									10283378	10456941	RP11-155J5	2	0.03%	0.20%
q11.223	69	6.80%				23429774	23574981	RP11-247J14	68	1.13%	6.70%			

附錄：台灣族群CNV發生位置與頻率一覽表

Chromo-some #	Case No.	Popul. (%)	Cyto	Case No.	Popul. (%)	Clone Start	Clone End	Clone Name	CNVR Event	CNVR (%)	Popul. (%)
			q12	234	23.05%	57236087	57310616	RP11-80F8	222	3.70%	21.87%
						57239667	57326818	RP11-88F4	213	3.55%	20.99%

註：藍色字體顯示者，表示為僅在CA2500 platform選用的探針；黃色欄位者為在此篇研究中首次發現CNV的位置；粉紅色欄位標示出CNV發生頻率大於1%者。

