計畫編號：NHRI-EX91-8933SL

國家衛生研究院九十一年度整合性醫藥衛生科技研究計畫

人類染色體之結構與功能

計畫名稱

年度成果報告

執 行 機 構：中山醫學大學

計 畫 主 持 人：林齊強

執 行 期 間：91 年 1 月 1 日 至 91 年 12 月 31 日

＊＊本研究報告僅供參考用，不代表本院意見＊＊

計畫名稱：人類染色體之結構與功能

計畫編號：NHRI-EX91-8933SL

執行機構：中山醫學大學

計畫主持人：林齊強

研究人員：李月君、江佩宜、張偉修、陳怡貞、林鉦展、黃仁裕、黃天龍

關鍵字：著絲點 DNA、著絲點蛋白質、Glioma cell line、TRAIL、apoptosis、SKY、CGH、三連核苷酸重覆擴增、myotonic dystrophy type I 、prelingual deafness

壹、九十一年度計畫研究成果摘要

本年度研究計劃之進行可分爲三個主要部份：（一）著絲點衛星 DNA 的結構，演化和功能之研究。（二）利用新近發展之分子細胞遺傳方法如 SKY 和 CGH 等系統對神經膠瘤(glioma)細胞株 sensitive or resistant to TRAIL induced ocpotosis 之研究。（三）參與對遺傳疾病機制之研究，如三連核苷酸重覆擴增(trinucleotide repeat expansion)疾病中 myotonic dystrophy type I （DM1）發病率及遺傳特性之探討和對 Cx26 gene mutation 與台灣地區 prelingual deafness 相關性之研究。並將過一年對上述三方向之研究成果摘要報告如下:

1. 完成對新近發現的一個哺乳類著絲點衛星 DNA 族系的研究：

高等生物如哺乳類動物的著絲點衛星DNA核苷酸序列在不同種類之間差異很大（Choo 1997），而且其基因體組織也非常複雜（Lee et al.1997）。本實驗室最近發現及鑑定出一個新的哺乳類著絲點衛星 DNA 家族，命名爲 cervid satellite IV DNA。Cervid satellite IV DNA 原本只用一對來自 white tailed deer satellite II DNA 的 primers，從印度山羌和中國山羌的 genome 中由 PCR amplify 得來。爲了進一步證明用一種

satellite DNA 片段作爲 primers，可以同時 amplify 得到兩種不同家族的 satellite DNA，我們再用兩種不同於山羌之鹿類，The black tailed deer 和 Canadians woodland caribou 之 genomic DNA 以同樣 satellite II DNA primer 做 PCR amplification，也同樣獲得除了 satellite II DNA 之外還有專屬於此兩種鹿類的 satellite IV DNA。進而從此四種不同鹿類的 satellite IV的株群（clones）中發現彼此之間存有極高（超過 90%）的 DNA 核苷酸序列的相似性。

利用螢光原位雜交法（FISH）和免疫螢光法（Immunofluorescence）偵測到 satellite IV DNA 緊附著絲點蛋白質，同時出現在著絲點的 kinetochore 之位置。由於 satellite IV DNA 在不同鹿類中保持有高度 DNA 核苷酸序列的相同性與其存在的位置非常靠近著絲點 kinetochore，我們推測這種新發現的 satellite DNA 對著絲點在細胞分裂的過程中可能扮演重要的角色。此項研究成果已發表於著名的染色體研究國際期刊 Chromosoma （2002）111:176-180。

2. 神經膠瘤分子細胞遺傳學之研究

神經膠瘤（glial tumors）爲中樞神經中最常見之腫瘤，惡性的神經膠瘤（malignant glioma）病人雖經手術切除及放射治療，其存活率仍不超過一年，由於先前發現一種與腫瘤壞死因子（tumor necrosis factor）有關且能引導細胞自殺（apoptosis）的 ligand （TRAIL）能導致腫瘤細胞自殺而不危害正常細胞。因此認爲 TRAIL 可能有助於神經膠瘤之治療。近來加拿大阿爾巴他大學病理系教授（Dr. Chunhai Hao）從十三個神精細胞株中發現有些細胞株對 TRAIL 有敏感性（Sensitive），有一部份細胞株對 TRAIL 有部份抵抗力（Partial resistant），其餘一部份則完全不受 TRAIL 的影響（Resistant）（Hao et al.2001）。因此本實驗室與

Dr. Hao 合作，研究是否由於不同 gioma 細胞株其基因體物質如 chromosome 之異常性不同而導致對 TRAIL 敏感度有所差異，進而期能找出對 TRAIL induced apoptosis 最相關之基因與機制。我們採用 9 個 glioma cell line ，其中三個爲 TRAIL sensitive:LN71，U343GM 和 T98G，三個爲 partial resistant : U138GM，U118 和 LN443，三個爲 TRIAL resistant : U373GM，LN215 和 LN464 做細胞遺傳分析。由於這些神精膠瘤細胞株的染色體異常非常複雜，並非以一般傳統細胞遺傳分析法就能明確鑑定，因此以分子細胞遺傳學技術，如:比較基因體雜交法（Comparative genomic hybridization ，CGH）和多顏色螢光雜染頻譜分析系統（SKY）配合傳統的 G-banding 對這 9 種不同 glioma 細胞之染色體異常做詳細分析。

CGH 分析結果發現這些 glioma cell lines 在下列染色體或染色體部分如 1p，3q，7，8q，19 和 20，有普遍擴增的現象，而在其它部分如：4q，9p，10，12q 和 18q 則有缺失的情形，其中最爲重要的發現是所有對 TRAIL induced apoptosis 有抵抗力的細胞株都有 chromosome regions 10p11.2-p15 和 12q 的缺失，因爲一個誘導細胞自殺有關的基因 CRADD（caspase and R1P adaptor with death domain）是位於染色體 12q21.33-q23.1 正好落在 12q 上，因此可能由於 CRADD 或與 CRADD 有關基因的缺失使這些細胞株不受 TRAIL 的影響而導致自殺，此一發現可能有助於對未來類似 TRAIL 之抗癌藥之研發。G-banding 配合 SKY 之細胞遺傳分析發現這些細胞株有基因體不平衡（genomic inbalance）和複雜的染色體互換之現象，染色體數目從超三倍體至低五倍體皆在這些細胞株中發現，許多用傳統 G-banding 無法判別其來源之 marker chromosomes 也可用 SKY system 鑑定出來，同時也偵測到超過 100 個

已以上的染色體斷點，這些斷點在染色體上分佈情形現在正進一步分析中。

3. 遺傳疾病機制的研究

（A） 繼續參與對三連核苷酸重覆擴增疾病之研究：將從 1990 至 2001 年期間登錄到的 MD1 病人之 DNA 重新鑑定，同時對 MD1 locus 的 CTG repeat size 做明確的判斷，確認 96 位 MD1 病人分別屬於 26 家族，確認是屬於肌強直肌肉萎縮症 myotonic dystrophy type1（MD1）。其發病率在台灣族群中爲 0.46/100,000，同時對帶有 MD1 家族後代而言，發病性之提早或嚴重性之增加一般是可以預測到的，甚至連一些有 MD1 的家族中，雖然其子代的 CTG 三連核苷酸之重覆擴增情形有減少之趨勢，但其提早發病率及嚴重性仍可預測出來。

最重要的發現是發病年齡與三連核苷酸重覆擴增長度之反相關（inverse correlation）性，是僅出現在那些病人有著輕度的 CTG 重覆擴增，此外又發現一位 MD1 的攜帶者（carrier）有一個在幼兒時期就發病（childhood -onset）的兒子，這位 MD1 carrier 本人具有 CTG repeat 長度之非一致性（heterogeneity）從 40 到 50 個 repeats 不等，因此指出其 premutation alleles 可能在精子形成過程中或在體細胞組織上有不穩定性，總而言之，這些研究結果發現 DM1 是一種在台灣的罕見疾病，而且病人後代之三連核苷酸重覆擴增減低之可能性只發生在那些具有高度重覆之對偶基因之病人。這些研究成果將發表於國際期刊 Neuroepideminology。

（B） Cx26（GjB2）基因突變與台灣地區先天性學語前發生耳聾症（Prelingual deafness）：最常見的遺傳性感覺異常症候是嚴重的聽覺失靈，在 1000 個兒童中就有一個失去這種聽覺。到目前爲止大約有 22

個 cloned genes 被證明與造成先天性無症狀性耳聾有關。在高加索族系和日本族系中有一半的先天性無症狀性耳聾，可能是由於 Cx26 基因突變而來，由 Cx26 基因所產生的 connexin 26 蛋白質與內耳區域性鉀離子循環有密切關係，因此我們對台灣地區 169 位學語前發生耳聾症病人及 100 位聽覺正常的人，針對其 GjB2 基因之密碼區做序列鑑定；在耳聾病人中發現兩個新的核苷酸變異，551G→A 和 109G-300del AT，同時也發現一個過去已被偵察到的 235del C 的變異；除此之外，也在病人和正常人中偵測到 4 個先前已被發現的多型性（polymorphism），79G→A，109G→A，341A→G 和 608T→C，和一個可能是新的多型性 -558G→A，發現在病人群中。很有趣的是，我們沒有發現一種在高加索族系中最常見的變異-35del G，我們的研究結果發現 235del C 是台灣地區最常見的 Cx26 變異（占所有變異的 57%）因此根據此一發現，我們已研發出一種針對 235 del C 突變之簡單分子偵測法，這種偵測法將有助於那些有先天性學語前耳聾的家庭試圖了解其耳聾的原因。此研究成果刊登於 European Jornal of Human Genetics。

參考文獻：

Choo KH (1997) The centromere . Oxford University Press, Oxford, New York, Tokyo .

Hao C, Beguinot F, ,Condorelli G, Trencia A, Van Meir E G, Young V W, Pareny I F, Roa W H, Petruk K C(2001) .Induction and intracellular regulation of tumor necrosis factor –related apoptosis –inducing legend (TRAIL)mediated apoptosis in human maglignant glioma cells.Cancer Reseach 61:1161-1170

Lee C, Wevrick R, Fusher RB, Ferguson-Smith MA. Lin CC(1997) Human centromeric DNA . Hum Genet 100:291-304

貳、九十一年度計畫著作一覽表

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整合之資料

若爲群體計畫，請勾選本表屬於: □子計畫；或 □總計畫(請自行整合)

1. 列出貴計畫於本年度中之<u>所有計畫產出</u>於下表，包含已發表或已被接受發表之文獻、已取得或被接受之專利、擬投稿之手稿（manuscript）以及專著等。

2. 「計畫產出名稱」欄位：請依「臺灣醫誌」參考文獻方式撰寫；

3. 「產出型式」欄位：填寫該產出爲國內期刊、、國外期刊、專利、手稿或專著等。

4. 「SCI」欄位：Science Citation Index，若發表之期刊爲 SCI 所包含者，請在欄位上填寫該期刊當年度之 impact factor。

5. 「致謝與否」欄位：請註明該成果產出之致謝單位。若該成果產出有註明衛生署資助字樣者，請以 DOH 註明；若該成果產出有註明國家衛生研究院委託資助字樣者，請以 NHRI 註明；若該成果產出有註明衛生署及國家衛生研究院資助字樣者，請合併以 DOH & NHRI 註明；若該成果產出有註明非上述機構資助字樣者，請以機構全銜註明。舉例如下：

序號	計畫產出名稱	產出型式	SCI*	致謝與否
例	Hsiu SL, Huang TY, Hou YC., Chin DH, Chao PDL. Comparison of Metabolic Pharmacokinetics of Naringin and Naringenin in Rabbits. Life Sciences 2002 Feb 15, 70:1481-1489.	國外期刊	1.808	NHRI

1.	Li YC, Lee C, Chang WS, Li SY, Lin CC. Isolation and Identification of a Novel Satellite DNA Family Highly Conserved in Several Cervidae Species. Chomosoma (2002) 111:176-183.	國外期刊	3.286	NHRI
2.	Wang YC, Kung CY, Su MC, Su CC, Hsu HM, Tsai CC,Lin CC, Li SY. Mutations of Cx26 Gene (GJB2) for Prelingual Deafness in Taiwan. Euro J Hum Genet (2002) 10:495-498.	國外期刊	3.173	NHRI
3.	Hsiao KM, Chen SS, Li SY, Chiang SY, Lin HM, Huang CC, Kuo HC, Jou SB, Su SS, Ro Ls, Liu CS, Lo MC, Chen CM, Lin CC. Epidemiological and Genetic Studies of Myotonic Dystrophy Type 1 in Taiwan. Neuroepidemiology (In Press)	國外期刊	1.45	NHRI
4.	Li, Y. C., Lee, C., Cheng, W. S., Li, S.Y. and **Lin, C. C.** Isolation and identification of a novel centromeric satellite DNA family highly conserved in several mammalian species. 52nd Annual American Society of Human Genetics Meeting. Baltimore, U.S.A. October 15 – 19, 2002 (*Am J Hum Genet* 71: 302)	國外期刊 Abstrate		NHRI
5.	Lin, C. C., Li, Y. C. and Lee, C. How telocentric mouse chromosomes are. 52nd Annual American Society of Human Genetics	外國期刊 Abstrate		NHRI

	Meeting. Baltimore, U.S.A. October 15 -19, 2002 (*Am J Hum Genet* 71:302)			

*本表如不敷使用，請自行影印。

參、九十一年度計畫重要研究成果產出統計表

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整合之資料

若爲群體計畫，請勾選本表屬於: □子計畫；或 □總計畫(請自行整合)

(係指執行九十一年度計畫之所有研究產出成果)

<table>
<tr><td colspan="3">科 技 論 文 篇 數</td><td colspan="3">技 術 移 轉</td><td>技術報告</td></tr>
<tr><td></td><td>國 內</td><td>國 外</td><td>類 型</td><td>經 費</td><td>項 數</td><td>篇</td></tr>
<tr><td rowspan="2">期 刊
論 文</td><td rowspan="2">篇</td><td rowspan="2">3 篇</td><td rowspan="2">技 術
輸 入</td><td rowspan="2">千元</td><td rowspan="2">項</td><td>技術創新
項</td></tr>
<tr><td rowspan="2">著作權
(核准)
項</td></tr>
<tr><td rowspan="2">研討會
論 文</td><td rowspan="2">篇</td><td rowspan="2">2 篇</td><td rowspan="2">技 術
輸 出</td><td rowspan="2">千元</td><td rowspan="2">項</td></tr>
<tr><td>專利權
(核准)
項</td></tr>
<tr><td>專 著</td><td>篇</td><td>篇</td><td>技 術
擴 散</td><td>千元</td><td>項</td><td></td></tr>
</table>

〔註〕:

期刊論文：指在學術性期刊上刊登之文章，其本文部份一般包含引言、方法、結果、及討論，並且一定有參考文獻部分，未在學術性期刊上刊登之文章（研究報告等）與博士或碩士論文，則不包括在內。

研討會論文：指參加學術性會議所發表之論文，且尚未在學術性期刊上發表者。

專　　著：爲對某項學術進行專門性探討之純學術性作品。

技術報告：指從事某項技術之創新、設計及製程等研究發展活動所獲致的技術性報告且未公開發表者。

技術移轉：指技術由某個單位被另一個單位所擁有的過程。我國目前之技術轉移包括下列三項：一、技術輸入。二、技術輸出。三、技術擴散。

技術輸入：藉僑外投資、與外國技術合作、投資國外高科技事業等方式取得先進之技術引進國內者。

技術輸出：指直接供應國外買主具生產能力之應用技術、設計、顧問服務及專利等。我國技術輸出方包括整廠輸出、對外投資、對外技術合作及顧問服務等四種。

技術擴散：指政府引導式的技術移轉方式，即由財團法人、國營事業或政府研究機構將其開發之技術擴散至民間企業之一種單向移轉（政府移轉民間）。

技術創新：指研究執行中產生的技術，且有詳實技術資料文件者。

肆、九十一年度計畫重要研究成果

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整

合之資料

若爲群體計畫，請勾選本表屬於: □子計畫；或 □總計畫(請自行整合)

1. 計畫之新發現或新發明

（A） 利用 PCR-cloning 方法成功的發現一個新的哺乳類動物的著絲點衛星 DNA 族系命名爲 Cervid Satellite Ⅳ DNA ，這個新的 satellite DNA family 有一特點就是在觀察到的數種不同品種之鹿類其 Cervid Satellite Ⅳ DNA 序列保持有極高的相似性，指示著此新的 satellite DNA family 可能和著絲點之功能有密切關係。

（B）用分子細胞遺傳分析法如 Comparation Genomic Hybridization （CGH）等發現對 TRAIL 所引導之細胞自殺（apoptosis）有抵抗力的神精膠瘤細胞株其第 12 條染色體的長臂（12q）有缺失現象，在 TRAIL 引導 apoptosis pathway 中扮演重要角色的基因 CRADD 位於 12q 上之 12q21.33-q23.1 位置上。CRADD 之缺失可能使神精膠瘤細胞對 TRAIL induced apoptosis 有抵抗力性，如能進一步證實此關連性，對未來治療神精膠瘤藥物研發上應有很重要的貢獻。

（C）由於對台灣地區三連核苷酸重覆擴疾病如 myotonic dystrophy type 1（MD1）繼續研究，發現台灣族群 MD1 發病率只有 0.46/100,000，因此 MD1 雖然在歐美是一種常見的肌肉萎縮疾病，在台灣卻是一種罕見疾病。但對 MD1 家族後代其提早發病率及增加嚴重性是可預知的，同時還發現發病年齡與重覆擴張越長的反相關性只存在於那些 MD1 病人本身有著低度的

trinucleotide 重覆擴增，而且只有那些病人本身帶有高度的重覆的三連核苷酸對偶基因，其子代才會得到較低度的重覆擴增。

（D） 在台灣地區從 169 位有先天性學語前耳聾病人中發現兩種新 C_X26 基因突變（551G→A，299-300 del AT）和一個新的多型性（558G→A），同時還發現在高加索 prelimgual dealiness 的病人中最常見 C_X26 mutation 35 del G 對偶基因卻沒有在台灣耳聾病人群中出現，反而 235 del C 是最普遍的變異，因此也研發出一種對 235 del C 基因突變的分子偵測法，有助於對先天性耳聾家族了解其兒童耳聾的原因。

2. 計畫對學術界或產業界具衝擊性（impact）之研究成果

（A）新的著絲點 DNA family 的發現和研究不但能增加對著絲點本身結構和功能的了解，而且有助於對未來人類人工染色體（aritificial human chromosome）的建造，人類人工染色體爲理想的基因載體，其成功的建造將有助於對遺傳疾病或癌症做基因治療之用。

（B）發現可能由於 TRAIL induced apoptosis pathway 中 CRADD 基因缺失而使神經膠瘤細胞有抗 TRAIL 引導細胞自殺功能，此項發現如能更進一步證實，將有助於對抗神經膠瘤新藥的研發。

3. 計畫對民眾具教育宣導之研究成果

現時一般民眾對高齡產婦可能產生唐氏症等先天性染色體疾病有相當的興趣和了解，其實一些後天性的染色體異常也常發生於癌症

病人的腫瘤細胞中，著絲點結構和功能的研究將有助於對細胞染色體變異的了解，同時和是否能成功的建造人工染色體作爲治療遺傳疾病和癌症的工具有密切的關係。另一方面由於對腫瘤細胞細胞染色體變異的研究，發現某須些部份染色體的缺少或擴張會造成某種致癌或抑癌基因的缺少或擴張，將有助於讓民眾了解癌症發病的原因，和治癌新藥的研發。

其它如先天性肌肉萎縮和耳聾等遺傳疾病的研究發現台灣本土對某些遺傳疾病的發病機率與歐美國家不相同，某種疾病基因突變也有族群性之差異，至於研發出來的簡便遺傳疾病基因突變偵測法則有助於病人和其家屬對某種遺傳疾病原因之了解。

伍、九十一年度計畫所培訓之研究人員

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整合之資料

若爲群體計畫，請勾選本表屬於：□子計畫；或　□總計畫(請自行整合)

種　　類			人數	備　　註
專任人員	1. 博士後研究人員	訓練中		
		已結訓		
	2. 碩士級研究人員	訓練中		
		已結訓		
	3. 學士級研究人員	訓練中		
		已結訓		
	4. 其他	訓練中	2	

		已結訓		
兼任人員	1. 博士班研究生	訓練中		
		已結訓		
	2. 碩士班研究生	訓練中		
		已結訓		
醫師		訓練中		
		已結訓	1	
特殊訓練課程			25	人類細胞遺傳學特論
			20	第一屆生物技術科技教育改進計劃分子細胞遺傳學研習會

註： 1.特殊訓練課程請於備註欄說明所訓練課程名稱

2.本表如不敷使用，請自行影印

陸、參與九十一年度計畫所有人力之職級分析

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整合之資料

若爲群體計畫，請勾選本表屬於： □子計畫；或 □總計畫(請自行整合)

職級	所含職級類別	參與人次
第一級	研究員、教授、主治醫師	人
第二級	副研究員、副教授、總醫師	1人
第三級	助理研究員、講師、住院醫師	3人

第四級	研究助理、助教、實習醫師	1人
第五級	技術人員	2人
第六級	支援人員	人
合計		人

〔註〕

第一級：研究員、教授、主治醫師、簡任技正，若非以上職稱則相當於博士滿三年、碩士滿六年、或學士滿九年之研究經驗者。

第二級：副研究員、副教授、助研究員、助教授、總醫師、薦任技正，若非以上職稱則相當於博士、碩士滿三年、學士滿六年以上之研究經驗者。

第三級：助理研究員、講師、住院醫師、技士，若非以上職稱則相當於碩士、或學士滿三年以上之研究經驗者。

第四級：研究助理、助教、實習醫師，若非以上職稱則相當於學士、或專科滿三年以上之研究經驗者。

第五級：指目前在研究人員之監督下從事與研究發展有關之技術性工作，且具備下列資格之一者屬之：具初（國）中、高中（職）、大專以上畢業者，或專科畢業目前從事研究發展，經驗未滿三年者。

第六級：指在研究發展執行部門參與研究發展有關之事務性及雜項工作者，如人事，會計、秘書、事務人員及維修、機電人員等。

柒、參與九十一年度計畫所有人力之學歷分析

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整合之資料

若爲群體計畫，請勾選本表屬於：□子計畫；或□總計畫(請自行整合)

類別	學歷別	參與人次
1	博士	2人
2	碩士	2人
3	學士	1人
4	專科	人
5	博士班研究生	人
6	碩士班研究生	人
7	其他(大學部工讀生)	2人
合計		人

捌、參與九十一年度計畫之所有協同合作之研究室

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整合之資料

若爲群體計畫，請勾選本表屬於：□子計畫□總計畫(請自行整合)

機構	研究室名稱	研究室負責人
中山醫學大學	遺傳疾病研究室	蕭光明副教授
中山醫學大學	分子遺傳學研究室	李宣佑教授

玖、九十一年度之著作抽印本或手稿

依「貳、九十一年度計畫著作一覽表」所列順序附上文獻抽印本或手稿。

拾、九十一年度計畫執行情形

註：群體計畫(PPG)者，不論是否提出各子計畫資料，都必須提出總計畫整合之資料。若爲群體計畫，請勾選本表屬於：□子計畫；或□總計畫(請自行整合)

1. 請簡述原計畫書中，九十一年預計達成之研究內容

原計劃中本年度研究工作預計達成部份爲(1)Study on human centromeric DNA 和 mammalian cytogenetic.(2)molecular cytogenetic study on glima cells sensitive or resistant to TRAIL induced apoptosis.

2. 請詳述九十一年度計畫執行情形，並評估是否已達到原預期目標（請註明達成率）

原計劃中 mammalian cytogenetic 如各種鹿類染色體(包括台灣山羌)之研究已完成而對人類著絲點 DNA 之研究而發現新的 centromeric satellite DNA、cervid IV DNA 等。至於對 TRAIL 引導細胞自殺性神經膠瘤之分子細胞遺傳學研究工作，包括 CGH、G-banding/SKY 等，已完成操作工作，只有 G-banding/SKY 部份結果之分析仍在進行中，因此認爲大體上達到原預期目標(達成率 85%)。

Epidemiological and Genetic Studies of Myotonic Dystrophy Type 1 in Taiwan

K.M. Hsiao[a] S.S. Chen[b] S.Y. Li[a] S.Y. Chiang[c] H.M. Lin[c] H. Pan[a] C.C. Huang[d] H.C. Kuo[d] S.B. Jou[e] C.C. Su[c,f] L.S. Ro[d] C.S. Liu[g] M.C. Lo[g] C.M. Chen[h] C.C. Lin[i]

[a]Department of Life Sciences, [c]Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan; [b]Department of Neurology, Chang Gung Memorial Hospital, Kaohsiung, Taiwan; [d]Department of Neurology, Chang Gung Memorial Hospital, Taipei, Taiwan; [e]Department of Neurology, China Medical College Hospital, Taichung, Taiwan; [f]Tian-Sheng Memorial Hospital, Tong Kang, Pin-Tong, Taiwan; [g]Department of Neurology, Changhua Christian Hospital, Changhua, Taiwan; [h]Department of Neurology, Chung Shan Medical University Hospital, Taichung, Taiwan; [i]Department of Medical Research, China Medical College Hospital, Taichung, Taiwan

Correspondence: KM Hsiao, Department of Life Sciences, Chung Shan Medical University, No.110, Sec.1, Chien-Kuo North Road, Taichung, Taiwan 402, Republic of China. Tel: 886-4-24730022 Ext. 1802; Fax: 886-4-24723229; E-mail: kmh@csmu.edu.tw

Running title: DM1 in Taiwan

Key words: prevalence; CTG trinucleotide repeat; myotonic dystrophy type 1; genetic study; Taiwan

Abstract

To investigate the prevalence and genetic characteristics of myotonic dystrophy type 1 (DM1) in Taiwan, DM-suspected patients and their families identified during the period of 1990-2001 had their clinical records re-evaluated and the CTG repeat sizes at DM1 locus examined. A total of 96 subjects belonging to 26 families were identified as DM1 patients, which gave a minimal disease prevalence of 0.46/100,000 inhabitants. Clinical anticipation was frequently observed in affected families even in some parent-child pairs with transmission contraction of the CTG repeat size. The inverse correlation between age of onset and CTG repeat length was significant only in patients with small expansions. In addition, a DM1 carrier with a childhood-onset son was found to have CTG length heterogeneity in the range of 40 to 50, indicating that premutation alleles could be unstable during gametogenesis as well as in somatic tissues. Our data demonstrated that DM1 is a rare disease in Taiwan and showed that transmission contraction of repeat size are more likely to occur in alleles with large repeats.

Introduction

Myotonic dystrophy (DM) is a dominantly inherited disease characterized by multiple tissue involvement, including myotonia, progressively myopathy, cataracts, cardiomyopathy, testicular tabular atrophy, ptosis, and other developmental and degenerative manifestations [1]. The onset age and clinical severity show marked variability both between and within affected families. Genetic mutations at two different loci could underlie the pathogenesis of the multisystemic disorders, which were sequentially named as DM1 and DM2 [2]. DM1 is caused by an expansion mutation of the CTG trinucleotide repeat in the 3' untranslated region (UTR) of the DM protein kinase gene (*DMPK*) [3-5]. DM2 was recently confirmed to be caused by a tetranulcoetide CCTG repeat expansion located in intron 1 of the zinc finger protein 9 (*ZNF9*) gene [6].

At DM1 locus, the number of CTG repeat varies from 5 to 37 in normal alleles and is greater than 50 in disease alleles. *DMPK* transcripts from the disease allele remain in the nucleus [7,8] and cause pathogenic effects [9]. The expanded CTG repeats are unstable at both mitotic and meiotic levels, with a bias towards length increase in the successive generations, which accounts for the phenomenon of anticipation in affected families [10]. Premutation alleles containing 40 to 50 repeats are not associated with any detectable phenotype and are rarely detected. However,

individuals carrying premutation alleles are at high risk of having affected offspring within a limited number of generations [11], indicating that premutation alleles could be unstable during meiosis. Although premutation alleles could further expand and become disease alleles intergenerationally, somatic instability of CTG repeat length within this range has rarely been reported.

DM is considered to be one of the most common forms of adult muscular dystrophies, with a prevalence of 1 in 8,000 overall for Western European and North American populations [1] and one in 20,000 in Japan [12]. DM is less prevalent in Southeastern Asian populations and is rare or absent among Africans [13]. The prevalence of DM is found to be closely associated with the frequency of large normal CTG alleles in several different ethnic populations [14,15]. For the Taiwanese population, the allelic frequency of large normal CTG repeats (>18) is much lower than that in the European and Japanese population and is very close to that in the Southern African Negroids [16]. This indicates that the reservoir pool for further expansion towards DM1 could be extremely small in this area. To verify this hypothesis and to investigate the genetic characteristics of DM1 in Taiwanese population, we have reevaluated the clinical records and have examined the CTG repeat size of DM patients since 1990. In the present study, we confirmed that the DM1 prevalence rate is low in Taiwan. Meanwhile, the inverse correlation between

age of onset and CTG repeat size was found to be significant only for patients with small expansions. Although the phenomenon of anticipation is common in DM1 families, it may not correlate with CTG repeat size increase, especially in families with large CTG expansions in parents. Finally, we documented a new observation of CTG length heterogeneity in leukocytes of a premutation carrier.

Material and methods

Subjects

To carry out a study of DM1 prevalence in Taiwan, patient records during 1990-2001 at the Department of Neurology in six Medical Centers were re-evaluated. The records included the following information: name, date of birth, sex, level of education, occupation, onset age, date of diagnosis of the disease, clinical phenotypes and form of the disease. Once the cases were identified, we contacted the patients in order to confirm whether they had DM1 mutation using DNA analysis. This would exclude other myotonic syndromes or DM2. Family history for reconstructing the genealogical pedigree has also obtained. When symptomatic or possible at-risk relatives of a family were identified, molecular analysis was then performed, which was eventually extended to their asymptomatic, first-degree relatives. Individuals showing DM phenotype in a family confirmed to have DM1 mutation in at least one

family member were regarded as DM1 patients. Informed consent was obtained from patients and their family members to participate in the study.

Population

Taiwan has an area of 36,188 km^2, a population of 22,393,488 as of November 2001 and a density of 618 inhabitants/km^2. The six Medical Centers from which DM patients were referred cover an area of 25,592 km^2, with a population of 21,172,626 as of November 2001 and a density of 827 inhabitants/km^2.

DNA analysis

Genomic DNA was isolated from peripheral blood leukocytes or lymphoblasts transformed with Epstein-Barr virus using the Puregene DNA isolation kit (Gentra Systems, Minneapolis, MN, USA). To determine the CTG repeat length at lower range, DNA was amplified by PCR using primers H and ER as described previously [17]. Briefly, each 20 μl reaction consisted of 1x reaction buffer (containing 1.5 mM $MgCl_2$, 50mM KCl, 10 mM Tris-HCl (pH9.0), and 0.1% Triton X-100 (Promega, Madison, Wisconsin, USA), 0.2 mM dNTP (Promega), 0.5 μM each primer, 4 μCi α-[^{32}P]dCTP (10 mCi/ml; NEN Life Science Products, Boston, Massachusetts, USA), and 1 unit of *Taq* DNA polymerase (Promega, Madison, Wisconsin, USA). After an

initial template denaturation of 5 min at 96^{0}C, amplification was carried out for 35 cycles of 45 seconds at 96^{0}C, 45 seconds at 60^{0}C, and 1 minute 30 second at 72^{0}C, then followed by a final extension of 10 minutes at 72^{0}C in a Perkin Elmer thermal cycler (model 480; PE Applied Biosystems, Foster City, California, USA). The PCR products were resolved by electrophoresis on 6% urea-polyacrylamide gels in parallel with a pGEM sequencing ladder and were visualized by autoradiograph. Cloning followed by sequencing analysis of the PCR products was performed to confirm the CTG repeat number in the cases of Fig. 3.

The CTG repeat sizes of disease alleles were determined by PCR-based Southern blot analysis as follows: 100 picogram DNA in a total volume of 20 µl was PCR amplified using Advantage-GC cDNA polymerase mix (Clontech, Palo Alto, CA). Southern blot of the PCR product was then performed using a non-isotopic Quick-Light detection system (Life Codes, Stamford, Connecticut, USA) as described [18].

Results

DM1 prevalence

A total of 96 individuals, belonging to 26 families, were identified as DM1 patients. Among them, 59 probands were confirmed to have an abnormal CTG expansion at

DM1 locus. The remaining 37 probands, all of them are relatives of proven CTG expansion cases, did not agree to undergo DNA testing. The age of patients (58 males and 38 females) vary from 1 to 80 years (mean age ± standard deviation: 39.3 ± 15.1). None of the probands are Taiwanese aboriginals. All DM1 families identified are descendants of immigrants from China. The population covered in this study was about 21 million people as of November 2001, and a minimal prevalence of the disease was estimated to be 0.46/100,000 inhabitants.

Based on clinical manifestations, those 59 DM1 patients with DNA analysis were subdivided as follows: 6 cases (10.2%) presented no or minimally affected symptoms, 51 cases (86.4%) showed classical phenotypes, and 2 cases (3.4%) were congenital patients. The CTG lengths in classical cases vary from 51 to 1200 repeats. All minimally affected subjects have 150 or less CTG repeats and two maternally transmitted congenital patients have more than 1000 repeats.

Age at onset and CTG repeat length

To investigate whether age of onset correlates with the CTG repeat size, we analyzed the relationship in 47 patients with confirmed onset age who exhibited a range of symptoms including frontal baldness, cataract, muscle weakness, and myotonia. The result was shown in Figure 1. 89% (42/47) of patients started showing clinical

phenotypes within the age range of 10 to 50 years old. The shortest CTG repeat size among the three patients with onset age less than 10 was 250. The repeat length of five patients with age of onset larger than 50 was 51, 64, 78, 740, and 1100, respectively. This analysis indicates that, although huge variation in CTG repeat length could be observed in patients with similar age of onset, the original data set as a whole can be divided into two groups: one with CTG sizes greater than 250 repeats (34 data points) and one with 250 repeats or less (13 data points). The equation for regression analysis for those below 250 repeats is $y=-0.147x + 55.8$ and for those above 250 repeats is $y=0.006 + 19.2$. These data suggest that the inverse correlation between age of onset in DM1 patients and CTG repeat size in leukocytes is significant only for patients with small expansions (less than 250 repeats).

Subsequently, the change of onset age and CTG repeat number during transmission from parent to offspring were evaluated in eight families with probands found in at least two generations (Figure 2). Nine out of fourteen parent-child pairs (64%) showed size increase and earlier age of onset on transmission. The repeat sizes of the parents in those pairs vary from 42 to 800. However, there are 5 pairs in three families (family Q, S, and W) where the CTG repeat size of offspring had decreased from parental alleles of 250 or more repeats. Although the repeat size decreased, the offspring started showing symptoms at an earlier age (4-18 years old) than their

parents (28 to 33 years old). Meanwhile, the clinical phenotypes in the contracted patients were found to be more diverse. In addition to the major clinical presentations, including action myotonia and muscle weakness, the offspring acquired additional symptoms not present in their parents, such as cardiac conduction defect and memory impairment in family S, frontal baldness in family Q, and pneumonia in family W. In general, the parents have a more severe classical DM phenotype and the offspring show a wider spectrum of secondary manifestations in these pairs.

Somatic instability of CTG repeat length of the premutation allele

During our survey on the CTG repeat length at DM1 locus, an asymptomatic male in the DM1 family (family C in Fig. 2) was found to have CTG repeat length heterogeneity in the range of 40 to 50 (lane 8 in Fig. 3). Subsequent cloning and sequencing analysis demonstrated that there are at least four different alleles with 5, 42, 43, and 50 CTG repeats, respectively, in the genome of this carrier. Using the same PCR conditions, a DM1 carrier in another family (family F in Fig. 2) was shown to have 5 and 75 CTG repeat sizes only (lane 11 in Fig. 3), suggesting that the observed CTG length instability in the leukocytes of the male carrier in family C is not due to the PCR artifact. Although the somatic instability of CTG repeats was frequently seen in patients, it is the first case observed in a carrier with multiple repeat

sizes under 50.

Discussion

In the present study, we have confirmed 96 DM1 patients belonging to 26 families in Taiwan, which gives a minimal prevalence rate of 0.46/100,000, as of November 2001. The low DM1 prevalence obtained here further confirms the prediction based on the allelic distributions of CTG repeat sizes in normal Taiwanese [16] and supports the notion that there is close association between disease prevalence and the frequency of normal alleles with higher CTG repeat in populations [14]. In China, the frequencies of large normal alleles ($(CTG)_{\geq 19}$) are close to that in Taiwanese (1.0 *vs* 1.4%) [16,19]. In addition, the major proportion of Taiwanese populations and all the DM1 families identified in this study are originally from the southeastern part of China, suggesting that the DM1 prevalence rate in these two areas could be very similar.

As far as the genotype-phenotype relationship was concerned, as with previous reports [20], in our sample a significant correlation was found between CTG expansion size and clinical category. Of the 96 DM1 patients in our DM1 population, only six benign and two congenital cases were identified. The remaining were all classical DM patients. The reasons for this could be that the index cases most likely referred to the neurologist when disturbed by serious clinical symptoms. Considering

that up to 85% of apparently normal relatives of our DM1 patients refused to undergo DNA analysis, our results may be an underestimation of the real prevalence of DM1 if we consider the relative stability and less pathogenic effects of minimally expanded disease loci in affected families [21]

Several reports on the correlation of the expansion size in lymphocytes and the age of onset of DM1 had shown a negative linear correlation without plateau effect [10,22-24]. However, our data favors the hypothesis that there exists an expansion size threshold at ~250 repeats, beyond which the expansion size does not influence the onset age [25,26]. The contradictory observations could be attributed to differences in the recruitment of patient samples and the way the data were presented. If DM1 patients of different clinical categories were included [10], the correlation could be biased towards a subgroup of patients. In addition, when the age of onset of DM1 patients was plotted against the CTG repeat number on a logarithmic scale, a plateau effect would be easily neglected especially if the threshold size is small. Therefore, it would be more likely to show the threshold size by original scale. The implication of threshold hypothesis originally came from the observation of the existence of a critical size between 150 and 400 repeats in myoblasts and fibroblasts, above which mutant DMPK RNA was not exported to the cytoplasm [8;27]. Because the reduced DMPK expression may account for particular aspects of DM [28,29],

further investigation of the effects of various CTG repeat lengths on the expression of other DM-related genes, such as chloride channel type 1 (ClC-1) [30,31], and on the neuromuscular functions *in vivo* [32] may provide more evidence for this hypothesis.

In our collections, there were five parent-child pairs in three families (family Q, S, and W) with decreased repeat size in the next generation if the repeat size in the parents was 250 or more (Fig. 2). In agreement with previous reports [33-35], most of the observed intergenerational contractions of CTG repeats were in association with paternal transmission. Male DM patients tend to have decreased sperm function [36] and there was a selection against extreme expansion in sperm [37]. Thus, a selective advantage for sperm with the shorter repeats could account for the higher frequency of repeat size contraction during paternal transmission. Here, we observed that the minimal parental size of the intergenerational contraction was around 250 repeats. Although the repeat size determined in blood may not reflect the actual status in sperm, these results lead us to speculate that there exists a threshold size above which the sperm function is inversely correlated with the CTG repeat size.

In the present study, the offspring with repeat size contraction in leukocytes presented a wider spectrum of DM phenotype compared to their parents. One possible explanation could be the somatic mosaicism, *i.e.*, there were different amplification rates in various tissues. Indeed, it was demonstrated that in most of the tissues, except

cerebellum, the CTG repeat size of disease alleles is larger than that in blood [38-41]. In consistent with this notion, some symptoms in DM1 patients, such as respiratory insufficiency, cataract, cardiac abnormalities, *etc*, were not correlated with the CTG repeat size in blood [23,24]. Therefore, the CTG contraction in leukocytes might not compare to the size expansion or contraction in other tissues. On the other hand, it was reported that the CTG repeat size in patients' blood is directly related to the clinical phenotype and is accurate for prognostic assessment [20,22]. In such cases, the size change in leukocytes may compare to those in other tissues and the acquired symptoms in the contracted patients could be due to tissue-specific pathogenic mechanisms.

The meiotic and mitotic instability of the repeat length is frequently observed in DM1 patients, probably due to the secondary structure formed within the repeat tract [42]. In premutation alleles, however, somatic mosaicism of the repeat lengths has rarely been documented. As shown in figure 3, the somatic instability of CTG repeat would occur even in an asymptomatic acrrier with repeat sizes ranging from 40 to 50. The son of this asymptomatic carrier is a childhood-onset patient with 1100 CTG repeats (family C in Fig. 2), indicating that premutation alleles could be intergenerationally and somatically unstable. Recently, mismatch repair proteins were shown to affect somatic expansion behavior of the CTG repeat [43]. Further mutation

analysis of such genes may provide a deeper insight into the underlying mechanisms influencing the repeat stability in DM1 patients.

In summary, the present studies confirmed that the DM1 prevalence rate is low in Taiwan and demonstrated the importance of DNA genotyping in genetic counseling. Genetic analysis of this DM1 pool indicated that clinical anticipation could occur in families with transmission contraction of CTG repeats. In consistent with previous reports [25,26], we also found that the correlation between age of onset and CTG repeat size is significant only for patients with small expansions. Finally, we presented the evidence for the somatic mosaicism of CTG repeat sizes in a premutation carrier.

Acknowledgements

We are grateful to the DM patients and their families for their cooperation. This work was supported by grants from the National Science Council of the Republic of China (NSC89-2316-B-040-001, NSC89-2320-B-040-068, and NSC90-2320-B-040-025), and National Health Research Institute, Taiwan (NHRI-EX92-9207 SI).

References

1. Harper PS: Myotonic dystrophy. 2nd edn. WB Saunders: London, 1989.

2. The international Myotonic dystrophy Consortium (IDMC). New nomenclature and DNA testing guidelines for myotonic dystrophy type 1 (DM1). Neurology 2000;54:1218-1221.

3. Fu YH, Pizzuti A, Fenwick RG Jr, King J, Rajnarayan S, Dunne PW, Dubel J, Nasser GA, Ashizawa T, de Jong P, Wieringa B, Korneluk R, Perryman MB, Epstein HF, Caskey CT: An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 1992;255:1256-1258.

4. Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barcelo J, O'Hoy K, Leblond S, Earle-Macdonald J, de Jong PJ, Wieringa B, Korneluk RG: Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science 1992;255:1253-1255.

5. Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, Sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Johns C, Juvonen V, Johnson K, Harper PS, Shaw DJ, Housman DE: Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. Cell 1992;68:799-808.

6. Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP: Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science 2001;293:864-867.

7. Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH: Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. J Cell Biol 1995;128:995-1002.

8. Davis BM, McCurrach ME, Taneja KL, Singer RH, Housman DE: Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. Proc Natl Acad Sci U S A 1997;94:7388-7393.

9. Seznec H, Agbulut O, Sergeant N, Savouret C, Ghestem A, Tabti N, Willer JC, Ourth L, Duros C, Brisson E, Fouquet C, Butler-Browne G, Delacourte A, Junien C, Gourdon G: Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. Hum Mol Genet 2001;10:2717-2726.

10. Harley HG, Rundle SA, MacMillan JC, Myring J, Brook JD, Crow S, Reardon W, Fenton I, Shaw DJ, Harper PS: Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. Am J Hum Genet 1993;52:1164-1174.

11. Martorell L, Monckton DG, Sanchez A, Lopez De Munain A, Baiget M: Frequency and stability of the myotonic dystrophy type 1 premutation. Neurology 2001;56:328-335.

12. Osame M, Furusho T: Genetic epidemiology of myotonic dystrophy in Kagoshima and Okinawa districts in Japan. Rinsho Shinkeigaku 1983;23:1067-1071.

13. Ashizawa T, Epstein HF: Ethnic distribution of the myotonic dystrophy gene. Lancet 1991;338:642-643.

14. Goldman A, Ramsay M, Jenkins T: Absence of myotonic dystrophy in southern African Negroids is associated with a significantly lower number of CTG trinucleotide repeats. J Med Genet 1994;31:37-40.

15. Mor-Cohen R, Magal N, Gadoth N, Shohat T, Shohat M: Correlation between the incidence of myotonic dystrophy in different groups in Israel and the number of CTG trinucleotide repeats in the myotonin gene. Am J Med Genet 1997;71:156-159.

16. Pan H, Lin HM, Ku WY, Li TC, Li SY, Lin CC, Hsiao KM: Haplotype analysis of the myotonic dystrophy type 1 (DM1) locus in Taiwan: implications for low prevalence and founder mutations of Taiwanese myotonic dystrophy type 1. Eu J Hum Genet 2001;9:638-641.

17. Monckton DG, Coolbaugh MI, Aashizawa KT, Siciliano MJ, Caskey CT:

Hypermutable myotonic dystrophy CTG repeats in transgenic mice. Nature Genet 1997;15:193-196.

18. Hsiao KM, Lin HM, Pan H, Li TC, Chen SS, Jou SB, Chiu YL, Wu MF, Lin CC, Li SY: Application of FTA® sample collection and DNA purification system on the determination of CTG trinucleotide repeat size by PCR-based Southern blotting. J Clin Lab Anal 1999;13:188-193.

19. Zhang S, Wu H, Pan A, Xiao C, Zhang G, Hou Y, Chu J: Low incidence of myotonic dystrophy in Chinese Hans is associated with a lower number of CTG trinucleotide repeats. Am J Med Genet 2000;96:425-428.

20. Gennarelli M, Novelli G, Andreasi Bassi F, Martorell L, Cornet M, Menegazzo E, Mostacciuolo ML, Martinez JM, Angelini C, Pizzuti A, Baiget M, Dallapiccola B: Prediction of myotonic dystrophy clinical severity based on the number of intragenic [CTG]n trinucleotide repeats. Am J Med Genet 1996;65:342-347.

21. Simmons Z, Thornton CA, Seltzer WK, Richards CS: Relative stability of a minimal CTG repeat expansion in a large kindred with myotonic dystrophy. Neurology 1998;50:1501-1504.

22. Novelli G, Gennarelli M, Menegazzo E, Mostacciuolo ML, Pizzuti A, Fattorini C, Tessarolo D, Tomelleri G, Giacanelli M, Danieli GA, Rizzuto N, Caskey CT, Angelini C, Dallapiccola B: (CTG)n triplet mutation and phenotype

manifestations in myotonic dystrophy patients. Biochem Med Metab Biol 1993;50:85-92.

23. Jaspert A, Fahsold R, Grehl H, Claus D: Myotonic dystrophy: correlation of clinical symptoms with the size of the CTG trinucleotide repeat. J Neurol 1995;242:99-104.

24. Marchini C, Lonigro R, Verriello L, Pellizzari L, Bergonzi P, Damante G: Correlations between individual clinical manifestations and CTG repeat amplification in myotonic dystrophy. Clin Genet 2000;57:74-82.

25. Hamshere MG, Harley H, Harper P, Brook JD, Brookfield JF: Myotonic dystrophy: the correlation of (CTG) repeat length in leucocytes with age at onset is significant only for patients with small expansions. J Med Genet 1999;36:59-61.

26. Savic D, Rakocvic-Stojanovic V, Keckarevic D, Culjkovic B, Stojkovic O, Mladenovic J, Todorovic S, Apostolski S, Romac S: 250 CTG repeats in DMPK is a threshold for correlation of expansion size and age at onset of juvenile-adult DM1. Hum Mutat 2002;19:131-139.

27. Hamshere MG, Newman EE, Alwazzan M, Athwal BS, Brook JD: Transcriptional abnormality in myotonic dystrophy affects DMPK but not neighboring genes Proc Natl Acad Sci, U S A 1997;94:7394-7399.

28. Reddy S, Smith DB, Rich MM, Leferovich JM, Reilly P, Davis BM, Tran K,

Rayburn H, Bronson R, Cros D, Balice-Gordon RJ, Housman D: Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. Nat Genet 1996;13:325-335.

29. Berul CI, Maguire CT, Aronovitz MJ, Greenwood J, Miller C, Gehrmann J, Housman D, Mendelsohn ME, Reddy S: DMPK dosage alterations result in atrioventricular conduction abnormalities in a mouse myotonic dystrophy model. J Clin Invest 1999;103:R1-R7.

30. Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, Cannon SC, Thornton CA: Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. Mol Cell 2002;10:35-44.

31. Charlet-B N, Savkur RS, Singh G, Philips AV, Grice EA, Cooper TA: Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. Mol Cell 2002;10:45-53.

32. Mankodi A, Logigian E, Callahan L, McClain C, White R, Henderson D, Krym M, Thornton CA: Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. Science 2000;289:1769-1773.

33. Abeliovich D, Lerer I, Pashut-Lavon I, Shmueli E, Raas-Rothschild A, Frydman M: Negative expansion of the myotonic dystrophy unstable sequence. Am J Hum

Genet 1993;52:1175-1181.

34. Ashizawa T, Anvret M, Baiget M, Barcelo JM, Brunner H, Cobo AM, Dallapiccola B, Fenwick RG Jr, Grandell U, Harley H, Junien C, Koch MC, Korneluk RG, Lavedan C, Miki T, Mulley JC, Lopez de Munain A, Novelli G, Roses AD, Seltzer WK, Shaw DJ, Smeets H, Sutherland GR, Yamagata H, Harper PS: Characteristics of intergenerational contractions of the CTG repeat in myotonic dystrophy. Am J Hum Genet 1994;54:414-423.

35. Lopez de Munain A, Cobo AM, Saenz A, Blanco A, Poza JJ, Martorell L, Marti-Masso JF, Baiget M: Frequency of intergenerational contractions of the CTG repeats in myotonic dystrophy. Genet Epidemiol 1996;13:483-487.

36. Hortas ML, Castilla JA, Gil MT, Molina J, Garrido ML, Morell M, Redondo M: Decreased sperm function of patients with myotonic muscular dystrophy. Hum Reprod 2000;15:445-448.

37. Jansen G, Willems P, Coerwinkel M, Nillesen W, Smeets H, Vits L, Howeler C, Brunner H, Wieringa B: Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic events in (CTG)n repeat variation and selection against extreme expansion in sperm. Am J Hum Genet 1994;54:575-585.

38. Ishii S, Nishio T, Sunohara N, Yoshihara T, Takemura K, Hikiji K, Tsujino S, Sakuragawa N: Small increase in triplet repeat length of cerebellum from patients

with myotonic dystrophy. Hum Genet 1996;98:138-140.

39. Anvret M, Ahlberg G, Grandell U, Hedberg B, Johnson K, Edstrom L: Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. Hum Mol Genet 1993;2:1397-1400.

40. Thornton CA, Johnson K, Moxley III RT: Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. Ann Neurol 1994;35:104-107.

41 Martorell L, Johnson K, Boucher CA, Baiget M: Somatic instability of the myotonic dystrophy (CTG)n repeat during human fetal development. Hum Mol Genet 1997;6:877-880.

42. McMurray CT: DNA secondary structure: a common and causative factor for expansion in human disease. Proc Natl Acad Sci, USA 1999;96:1823-1825.

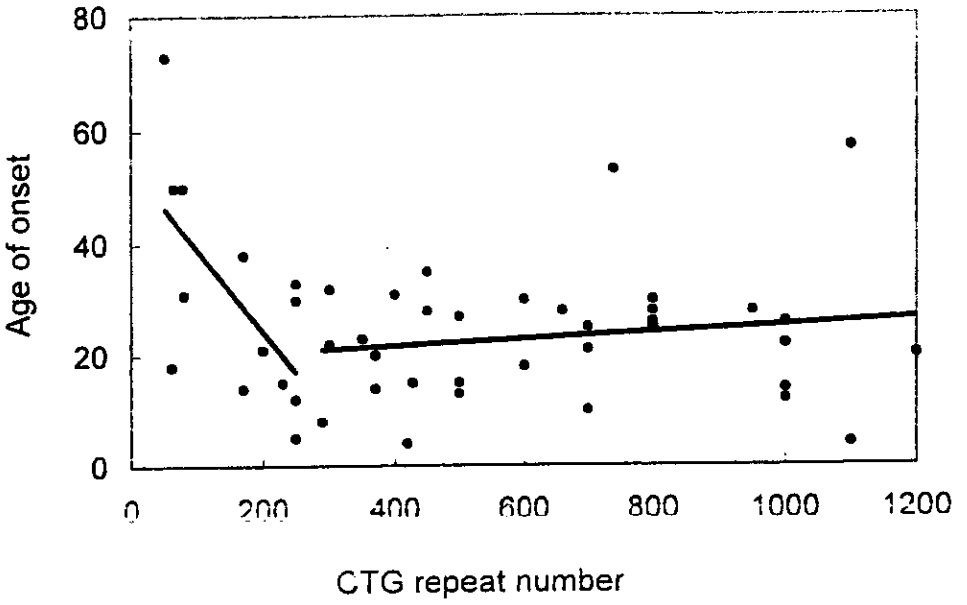
43. van den Broek WJ, Nelen MR, Wansink DG, Coerwinkel MM, te Riele H, Groenen PJ, Wieringa B: Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. Hum Mol Genet 2002;11:191-198.

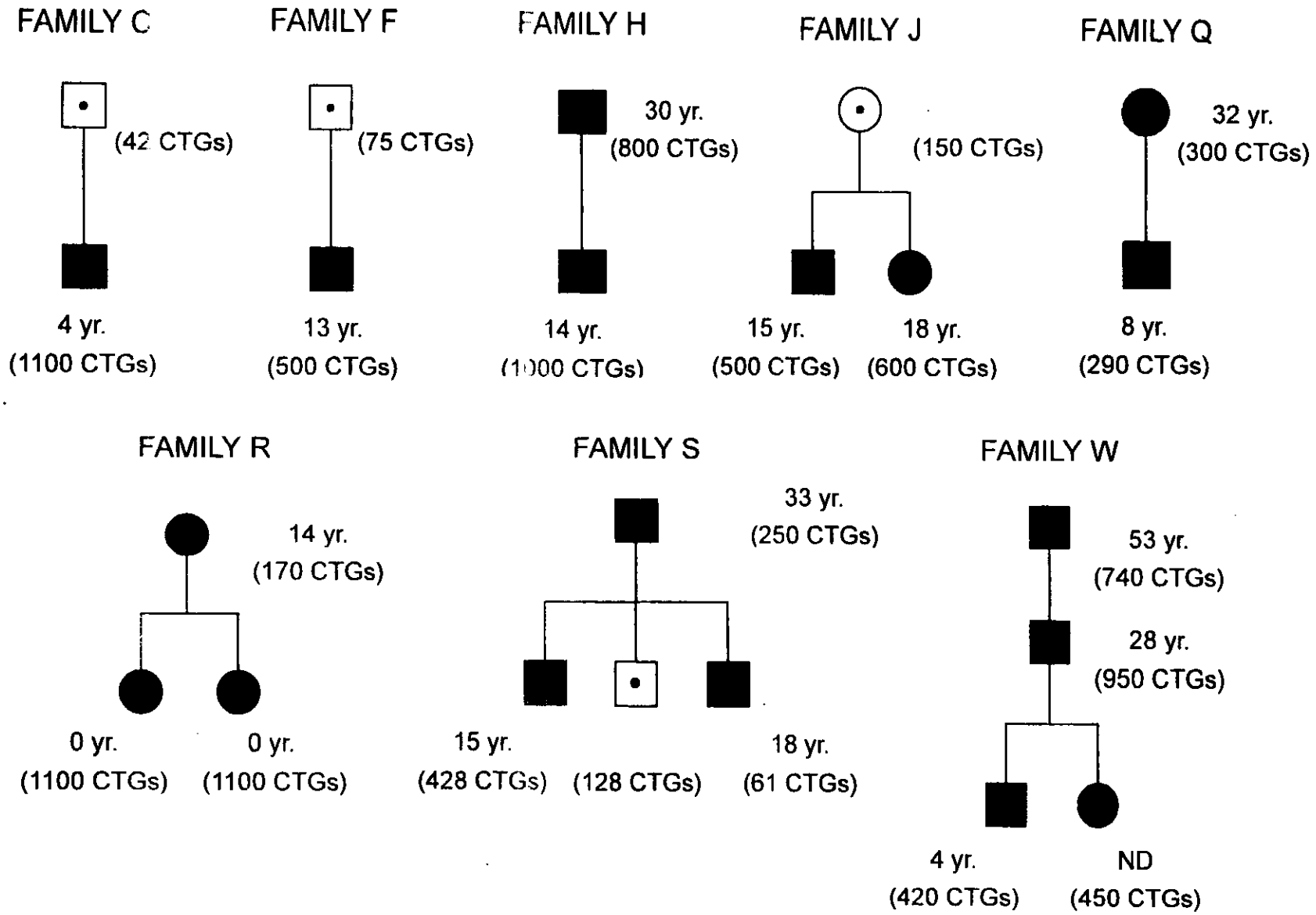
Figure legends

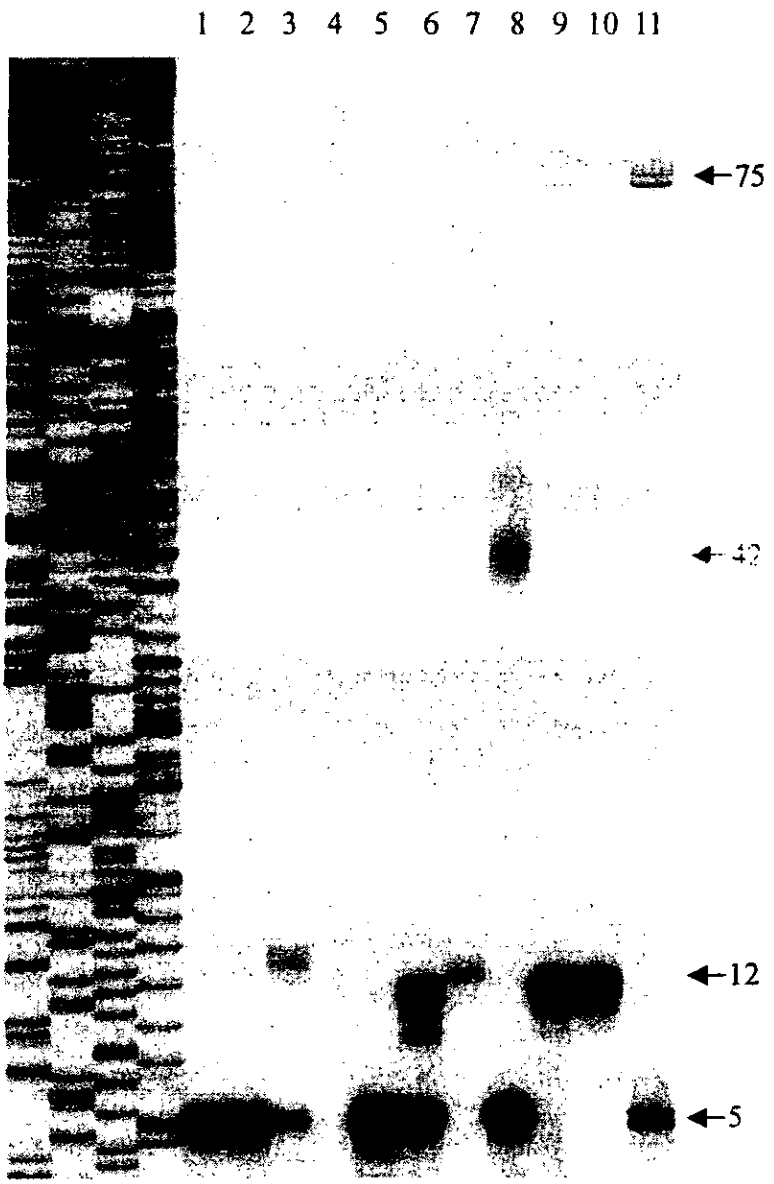
Figure 1 Age of onset for 47 DM1 patients, plotted against CTG repeat length.

Figure 2 Intergenerational transmission in DM1 families. The pedigrees of eight DM1 families and the age of onset for each patient were shown. Number in parenthesis represents the approximate expanded allele size in leukocytes. ND: not determined.

Figure 3 PCR analysis of the CTG repeat length in the DM1 locus. The CTG repeat size in blood DNA sample from four patients (lanes 1, 2, 7, 9), two carriers (lanes 8, 11), four normal DNA (lanes 3, 5, 6, 10), and one negative control of no DNA (lane 4) was analyzed as described in Material and Methods. The four tracks to the left contain pGEM sequence, which is used to determine the size of alleles. The corresponding CTG repeat number of signals is indicated on the right.


80
60
40
20
0
Age of onset
0
200
400
600
800
1000
1200
CTG repeat number


FAMILY C
(42 CTGs)
4 yr.
(1100 CTGs)
FAMILY F
(75 CTGs)
13 yr.
(500 CTGs)
FAMILY H
30 yr.
(800 CTGs)
14 yr.
(1000 CTGs)
FAMILY J
(150 CTGs)
15 yr.
(500 CTGs)
18 yr.
(600 CTGs)
FAMILY Q
32 yr.
(300 CTGs)
8 yr.
(290 CTGs)
FAMILY R
14 yr.
(170 CTGs)
0 yr.
(1100 CTGs)
0 yr.
(1100 CTGs)
FAMILY S
33 yr.
(250 CTGs)
15 yr.
(428 CTGs)
(128 CTGs)
18 yr.
(61 CTGs)
FAMILY W
53 yr.
(740 CTGs)
28 yr.
(950 CTGs)
4 yr.
(420 CTGs)
ND
(450 CTGs)


1 2 3 4 5 6 7 8 9 10 11
75
42
12
5

Chromosoma (2002) 111:176–183
DOI 10.1007/s00412-002-0200-x

ORIGINAL ARTICLE

Y.C. Li · C. Lee · W.S. Chang · S.-Y. Li · C.C. Lin

Isolation and identification of a novel satellite DNA family highly conserved in several Cervidae species

Received: 15 January 2002 / Accepted: 11 June 2002 / Published online: 23 July 2002

Abstract In an attempt to amplify cervid satellite II DNA from the genomes of Indian muntjac and Chinese muntjac, a pair of primers derived from the white tailed deer satellite II DNA clone (OvDII) yielded a prominent ~1 kb polymerase chain reaction (PCR) product (in addition to the expected 0.7 kb satellite II DNA fragments) in both species. The ~1 kb products were cloned, sequenced, and analyzed by Southern blotting and fluorescence in situ hybridization (FISH). This revealed that the ~1 kb cloned sequences indeed represent a previously unknown cervid satellite DNA family, which is now designated as cervid satellite IV DNA. Approximately 1 kb PCR clones were also obtained from the genomes of the black tailed deer and Canadian woodland caribou with similar primer pairs. Extremely high sequence conservation (over 90% homology) was observed among the clones generated from all four deer species and PCR-Southern hybridization experiments further verified the co-amplification of two kinds of satellite DNA sequences with the same pair of primers. This satellite DNA was found to co-localize with centromeric proteins at the kinetochore by a simultaneous FISH and immunofluorescence study. Due to its high sequence conservation and close association with kinetochores, the newly identified satellite DNA may have a functional centromeric role.

Edited by: J. Rattner

Accession numbers: AY 064466, AY 064467, AY 064468, AY 064469

Y.C. Li · W.S. Chang · S.-Y. Li · C.C. Lin (✉)
Department of Life Sciences, Chung Shan Medical University, No.110, Sec.1, Chien Kuo N. Rd., Taichung, Taiwan 40203
e-mail: cclin@csmu.edu.tw

C. Lee
Department of Pathology, Brigham and Woman's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA

C.C. Lin
Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada T6G 2B7

Introduction

One unique feature of higher eukaryotic genomes is the universal presence of large quantities of tandemly repeated DNA (satellite DNA), which can account for up to 10%–20% of some mammalian genomes (Beridze 1986). Satellite DNA is located mainly in the centromere, and up to tens of megabases can reside in a single human centromere (Choo 1997). Furthermore, the genomic organization of satellite DNAs in the centromeric region appears quite complex (Lee et al. 1997b) and no sequence conservation can be found among a wide range of species. Whether centromeric satellite DNA has a definitive functional role is still debatable (Choo 2000; Henikoff et al. 2001) and due to its abundance and highly repetitive nature, few attempts have been made to sequence this region of the chromosome completely. The existence of a functional "core" centromeric DNA sequence, similar to that found in the budding yeast (*Saccharomyces cerevisiae*), has not been reported in higher eukaryotic species. However, it cannot be ruled out that within the complex centromere, less prominent yet functional sequence motifs have gone undetected (Henikoff et al. 2001). Most recently, efforts have been made to map an ~450 kb region of the centromeric DNA on the human X chromosome and identify a candidate functional centromeric sequence that comprises ~3 Mb higher-order repeats of DXZ1 α-satellite DNA (Schueler et al. 2001).

Three centromeric satellite DNA families have been identified in the Cervidae (Table 1). Satellite I is the predominant centromeric satellite DNA family found in all deer species examined (Bogenberger et al. 1987; Scherthan 1991; Lee et al. 1997a). This cervid satellite DNA has a repeat unit of 0.8 kb in plesiometacarpalia deer and 1 kb in telemetacarpalia deer (Lee et al. 1997a). These monomers are thought to represent the higher-order structures of internal 31 bp subrepeats (Bogenberger et al. 1985; Lee and Lin 1996). The cervid satellite II family was initially isolated from the white tailed deer, and is characterized by monomeric repeats of 0.7 kb (Qureshi

Table 1 Cervid satellite DNA families

Family	Species		Satellite DNA clone	Reference
Satellite I	Indian muntjac	*Muntiacus muntjak vaginalis*	MMVsatIA	Bogenberger et al. (1982)
	Chinese muntjac	*Muntiacus reevesi*	C5	Lin et al. (1991)
	Roe deer	*Capreolus capreolus*	CCsatI	Scherthan (1991)
	Caribou	*Rangifer tarandus caribou*	Rt-Pst3	Lee et al. (1994)
	Red deer	*Cervus elaphus*	Ce-PstI	Lee and Lin (1996)
	Moose	*Alces alces*	Aa-Msp	Lee et al. (1997a)
	White tailed deer	*Odocoileus virginianus*	Ov-Msp	Lee et al. (1997a)
	Mule deer	*Odocoileus hemionus*	Oh-Msp	Lee et al. (1997a)
	Fallow deer	*Dama dama*	Dd-PstI	Lee et al. (1997a)
Satellite II	White tailed deer	*Odocoileus virginianus*	OvDII	Qureshi and Blake (1995)
	Caribou	*Rangifer tarandus caribou*	Rt-0.7	Li et al. (2000a)
	Indian muntjac	*Muntiacus muntjak vaginalis*	Mmv-0.7	Li et al. (2000b)
Satellite III	Roe deer	*Capreolus capreolus*	CCsatIII	Buntjer et al. (1998)
	Chinese water deer	*Hydropotes inermis*	HI-III	Li et al. (unpublished)

and Blake 1995). It was later also found in the genomes of the Indian muntjac (Vafa et al. 1999; Li et al. 2000b) and other deer species (Li et al. 2000a). Cervid satellite II DNA appears to be complexed with centromeric protein A (CENP-A), as demonstrated by immunoprecipitation of Indian muntjac DNA with human anticentromere autoantibodies (Vafa et al. 1999). The cervid satellite III DNA family has a repeat unit of 2.2 kb and has so far only been reported in the genome of the roe deer (*Capreolus capreolus*) (Buntjer et al. 1998).

In this study, we have obtained ~1 kb repetitive DNA fragments using a pair of primer sequences designed to amplify satellite II DNA from the Indian muntjac or Chinese muntjac genomes. Similar ~1 kb repeated DNA elements were also generated from the genomes of the black tailed deer and the Canadian woodland caribou using a modified pair of primers. The polymerase chain reaction (PCR) products were cloned, analyzed by Southern blotting, sequenced and studied by fluorescence in situ hybridization (FISH). This suggested the identification of a novel cervid centromeric satellite DNA with a monomeric unit of 1 kb. The ~1 kb satellite DNA clones from these four cervid species share extremely high sequence homology and are located close to the kinetochore regions. These findings imply that this new cervid satellite DNA may be functionally relevant.

Materials and methods

Cell culture and DNA isolation

Four cervid cell lines were used in this study. Cell lines from a male Indian muntjac *(Muntiacus muntjak vaginalis)*, a male Chinese muntjac *(Muntiacus reevesi)*, a male Canadian woodland caribou *(Rangifer tarandus caribou)*, were established in our laboratories previously. The female black tailed deer *(Odocoileus hemionus hemionus)* cell line was acquired from the American Type Culture Collection. Cells were grown in Dulbecco's modified Eagle's medium (Gibco/BRL), supplemented with 10% fetal calf serum, 1% glutamine and 1% penicillin-streptomycin. Procedures for cell harvesting and genomic DNA isolation were described elsewhere (Li et al. 2000a).

Polymerase chain reaction amplification and cloning

Amplification of the ~1 kb fragment from genomic DNA samples of the two muntjac species was achieved by PCR using a set of primers (forward: GAGCTGCCTGACAGACTCG; reverse: CAGAGCCGACCTAGGATCAC; Li et al. 2000a) derived from the published white tailed deer *(Odocoileus virginianus)* satellite II sequence (OvDII) (Qureshi and Blake 1995). Polymerase chain reaction was performed in a 25 µl reaction volume with 10 mM TRIS-HCl, pH 9.0, 50 mM KCl, 0.1% Triton X-100, 1.5 mM of $MgCl_2$, 200 µM each of dNTP, 200 nM of each primer, 100 ng of the genomic template DNA, and 2.5 U of Taq polymerase (Promega). Predenaturation at 94°C for 5 min was followed by 30 cycles of 30 s at 94°C for denaturation, 30 s at 50°C for annealing and 1 min at 72°C for extension. The final extension was carried out at 72°C for 10 min. The PCR products were electrophoretically fractionated on a 1.2% agarose gel, and prominent ~1 kb DNA fragments were excised, purified, and ligated into the pGEMT-easy vector (Promega). The recombinant plasmids were used to transform XL1-blue competent cells. Subsequently, the transformants were screened and randomly chosen for further characterization following the standard procedure (Maniatis et al. 1982).

Amplification and cloning of the ~1 kb DNA fragment from the black tailed deer and the caribou genomes were similar to that mentioned above except that the primer sequences were modified based on the sequence of the 1 kb clone obtained from the Indian muntjac (forward: GACTGATTTCCTGGGTTAAGAG; reverse: CACACAGAATGCTAGGAAATCC) and the annealing temperature was set at 48°C.

DNA sequencing and analysis

The ~1 kb PCR clones from the genome of Indian muntjac, Chinese muntjac, caribou and black tailed deer were designated as MMV-1.0, MR-1.0, RTC-1.0 and OHH-1.0, respectively, and sequenced from both ends using a dideoxy chain termination kit (United States Biochemical) and read on a Perkin-Elmer ABI DNA sequencer (model 377). The complete DNA sequences of these four clones were deposited in the GenBank database (Accession numbers AY 064466, AY 064467, AY 064468 and AY 064469). Single base shift self-comparisons (Plucienniczak et al. 1982) were also conducted to evaluate the presence of internal subrepeats.

Southern and slot-blot analysis

For Southern blot experiments, 10 µg aliquots of genomic DNA of muntjac species were each digested with one of six different re-

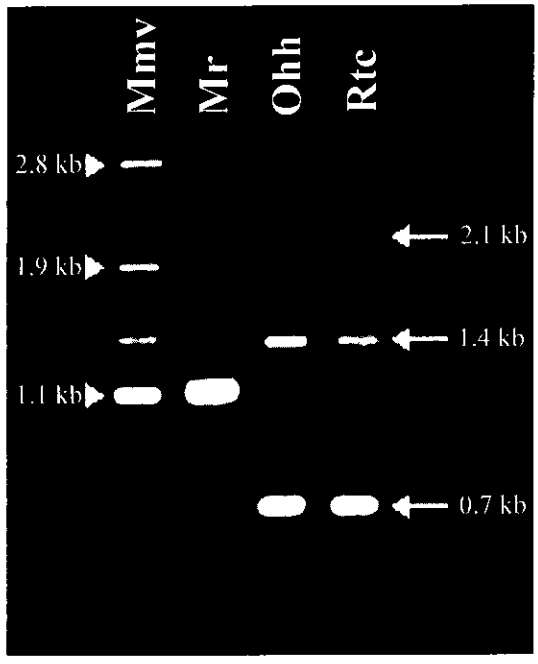


Fig. 1 Electrophoretic analysis of polymerase chain reaction (PCR) products. 100 ng of PCR products from each of the four deer species (*Mmv, Muntiacus muntjak vaginalis*; *Mr, M. reevesi*; *Ohh, Odocoileus hemionus hemionus*; *Rtc, Rangifer tarandus caribou*) amplified with a primer pair derived from a cervid satellite II DNA clone (OvDII) of white tailed deer were electrophoretically separated in a 1.2% agarose gel. A prominent 1.1 kb band and a 1.9 kb band are observed in the *Mmv* and *Mr* lanes. In addition three bands of 0.7, 1.4 and 2.8 kb are seen in the *Mmv* lane and a 1.4 kb light band is found in the *Mr* lane. Three bands of 0.7 kb register (0.7, 1.4 and 2.1 kb) are seen in both the *Ohh* and *Rtc* lanes

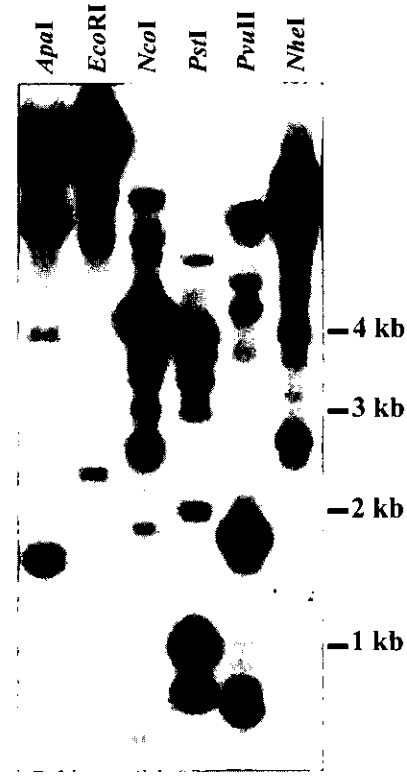


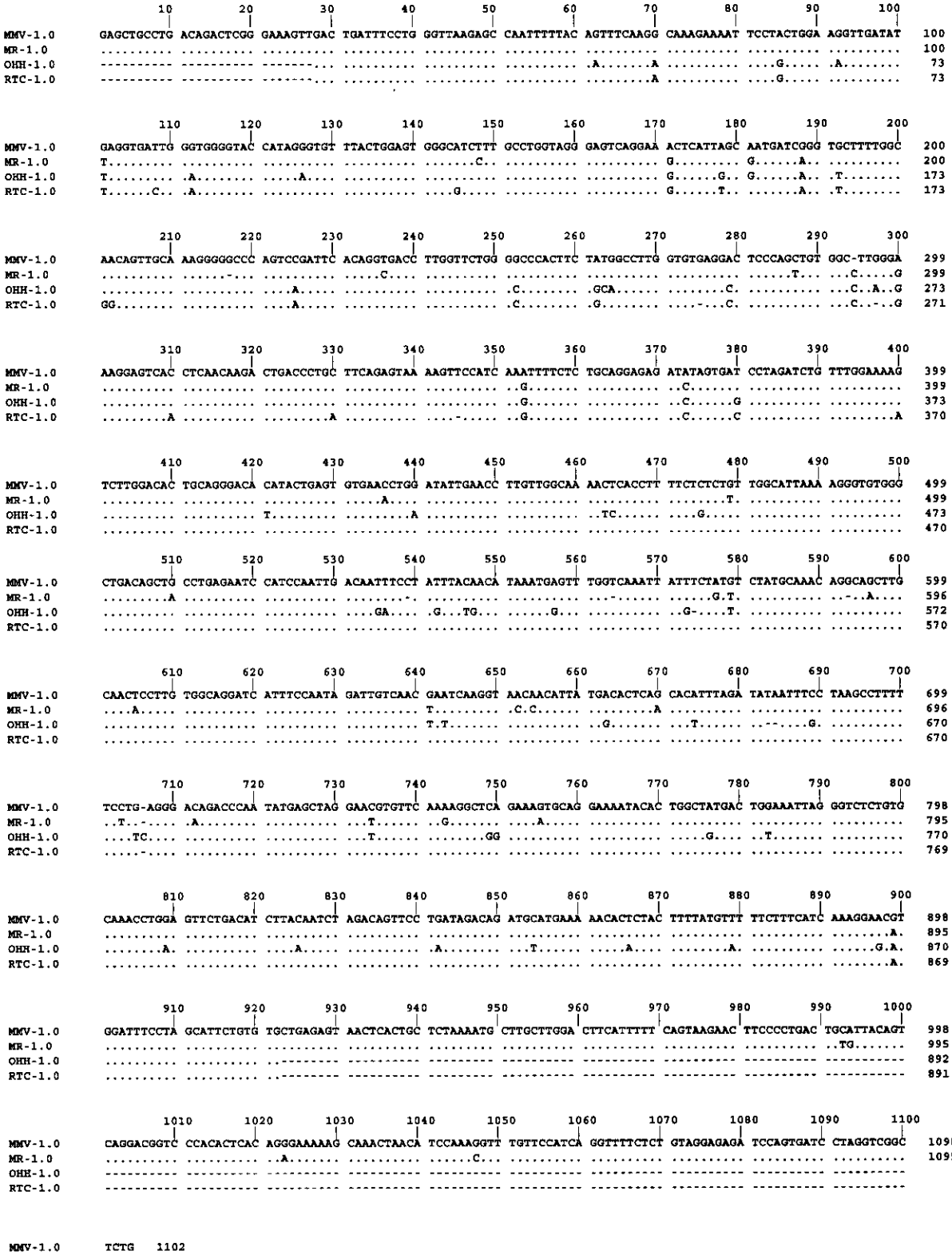
Fig. 2 Southern blot hybridization of Indian muntjac genomic DNA digested with restriction enzymes ApaI, EcoRI, NcoI, PstI, PvuII and NheI, fractionated by conventional gel electrophoresis, and probed with ^{32}P-labeled MMV-1.0 DNA. All digests show multiple hybridization bands of repeated DNA nature but the typical ladder pattern often seen with cervid satellite DNA is not observed in any digest. However, 1, 2, 3 and 4 kb hybridization bands are observed for the PstI digest, suggesting that a significant portion of the repeated sequence has a 1 kb register

striction endonucleases. Digested DNA samples were electrophoretically fractionated, transferred to a nylon membrane (Biodyne), and hybridized with [^{32}P]dCTP-labeled DNA probes. The conditions for hybridization, membrane washing and autoradiography were described previously (Li et al. 2000b). Polymerase chain reaction-Southern hybridizations were similarly performed with the exception that DNA samples used were PCR-amplified products from the four deer species. Hybridizations were carried out with either ^{32}P-labeled MMV-1.0 DNA probe or satellite II DNA probes. Slot-blot hybridization procedures for copy number estimations of the 1 kb repeated DNA elements in the genomes of four deer species were also described in detail earlier (Li et al. 2000b).

Fluorescence in situ hybridization and immunofluorescence microscopy

Chromosomal preparations of the Indian muntjac, Chinese muntjac and black tailed deer were obtained from established fibroblast cell lines following routine cytogenetic protocols. For single-color FISH experiments, the ~1 kb cloned DNA probes (MMV-1.0, MR-1.0 and OHH-1.0) were labeled with biotin and hybridized to metaphase chromosomes of each of the respective species. The hybridization signals were detected with Cy3-avidin. In dual-color FISH, the newly identified ~1 kb (MMV-1.0) DNA and a cervid satellite II DNA (Mmv-0.7) (Li et al. 2000b) were labeled with biotin and digoxigenin, respectively. The biotin-labeled probe was detected with Cy3-avidin whereas the digoxigenin-labeled probe was detected with fluorescein isothiocyanate (FITC)-conjugated antibodies. The protocol used for simultaneous immunofluorescence and FISH studies was described by Sullivan and Schwartz (1995) with a brief modification. The Indian muntjac cells were treated with Colcemid (0.1 μg/ml) for 3 h prior to harvesting. The harvested cells were incubated in 0.075 M KCl for 15 min at room temperature and subsequently cytospun onto slides. CREST sera, specific for anti-CENP-A, -B and -C antibodies (a gift from Dr. J.B. Rattner, University of Calgary, Canada) were used to identify kinetochore domains. The primary CENP antibodies were detected with FITC-conjugated rabbit anti-human IgG. After immunofluorescence experiments, cells on the slide were immediately fixed in 10% formalin/KCM buffer (120 mM KCl, 20 mM NaCl, 10 mM TRIS-HCl, pH 7.6, 0.1% Triton X-100) for 10 min and then in 3:1 methanol:acetic acid for another 15 min. Fluorescence in situ hybridization experiments were carried out in a manner similar to that described for single-color FISH experiments with biotin-labeled MMV-1.0 probes.

Fig. 3 Comparison of DNA sequences of ~1 kb PCR clones ▶ MMV-1.0, MR-1.0, OHH-1.0 and RTC-1.0. The sequence of MMV-1.0 is shown at the *top* of each set. Only unmatched bases (in comparison with the MMV-1.0 sequence) of MR-1.0, OHH-1.0 and RTC-1.0 are indicated and the matched nucleotides are indicated by *dots*. Occasional *gaps* (–) are introduced to improve the alignment. The percentages of sequence homology between any two of the four ~1 kb clones examined are also presented (*inset*)



	MR-1.0	OHH-1.0	RTC-1.0
MMV-1.0	96%	93%	96%
MR-1.0	-	91%	94%
OHH-1.0		-	93%

Results

Isolation and characterization of ~1 kb repetitive DNA elements in muntjacs

Primer pair sequences, designed to amplify cervid satellite II DNA, produced the three expected bands of 0.7 kb register with the Indian muntjac DNA template, and a 1.4 kb band with the Chinese muntjac DNA template. In addition, two unexpected PCR products (a prominent 1.1 kb band and a 1.9 kb band) were observed using both the Indian and Chinese muntjac DNA templates (Fig. 1). However, only bands with the expected size of 0.7 kb register were observed with template DNAs of black tailed deer and Canadian woodland caribou using the same pair of primers.

The ~1 kb PCR products of the two muntjac species were cloned, screened and designated as MMV-1.0 for the ~1 kb clone of Indian muntjac and MR-1.0 for that of Chinese muntjac. Southern blot hybridization with the MMV-1.0 probe to Indian muntjac genomic DNA produced multiple hybridization bands with irregular patterns (Fig. 2). Although, no typical type A-like pattern was found in any particular digestion, a 1 kb register can be detected in PstI-digested DNA. Copy number estimations indicated that about 0.08% of the Indian muntjac genome and 0.16% of the Chinese muntjac genome contain this repeated DNA sequence.

The MMV-1.0 clone of Indian muntjac and the MR-1.0 clone of the Chinese muntjac were sequenced (MMV-1.0 is 1102 bp and MR-1.0 is 1099 bp in length). Both clones are 60% AT-rich and share an extremely high (96%) sequence similarity to each other (Fig. 3). No significant internal subrepeat was detected by the single base shifts, self-comparison method of Pluciennicczak et al. (1982). These two clones also do not share significant sequence similarity with any DNA sequences currently deposited in GenBank, including known cervid or bovine satellite DNA families. These findings suggest that the 1 kb repeated DNA sequence could represent a new cervid satellite DNA family, which we now refer to as cervid satellite IV.

Polymerase chain reaction-Southern hybridization and isolation of cervid satellite IV DNA in black tailed deer and Canadian woodland caribou

Polymerase chain reaction-Southern blot hybridization was performed to verify further the co-amplification of both the ~1 kb repetitive DNA and the satellite II DNA from genomic DNA templates of the four deer species using a single primer pair derived from a satellite II clone (OvDII). A Southern blot of the PCR products was obtained after hybridization with ^{32}P-labeled MMV-1.0 (Fig. 4a, left panel). This probe hybridized very strongly with the corresponding 1.1 kb PCR product (Fig. 4a, right panel) in the two tested muntjac species. Two additional faint bands were also observed in the Chinese muntjac. However, these bands were not detected in the corresponding PCR products (Fig. 4a, right panel). The MMV-1.0 probe also detected a faint 1.1 kb band in both black tailed deer and Canadian woodland caribou that was not found in the corresponding gel (Fig. 4a, left panel). Another Southern blot was obtained after hybridization with satellite II DNA probes (a mixture of Mmv-0.7, Mr-0.7, Ohh-0.7 and Rt-0.7 clone DNA). Hybridization profiles showing bands of 0.7 kb register were observed in all four tested deer species (Fig. 4b, left panel). Together, these observations showed that both the ~1 kb repeated DNA and cervid satellite II DNA could be co-am-

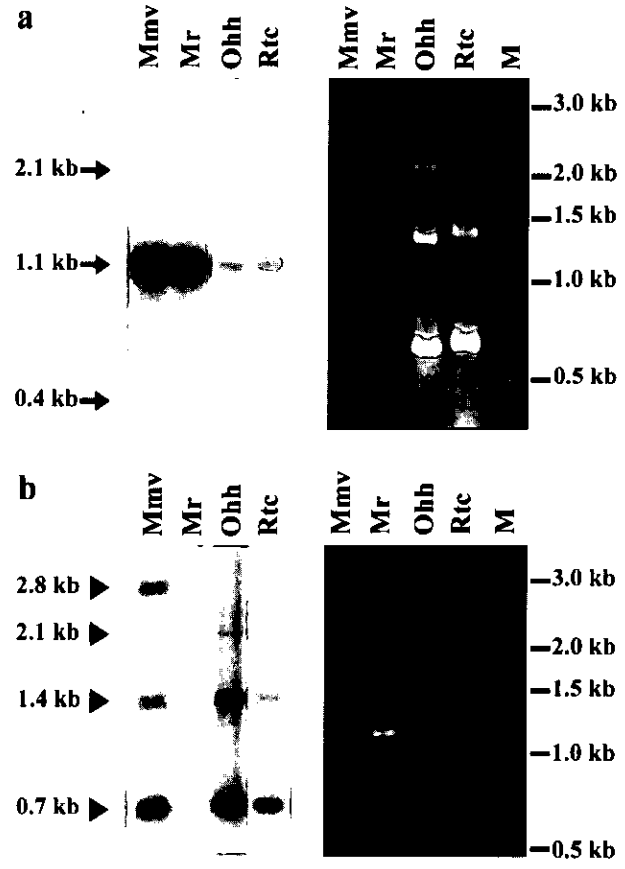


Fig. 4a, b Polymerase chain reaction-Southern blot analysis. **a** The electrophoretically separated PCR products (10 ng each for *Mmv* and *Mr*, and 150 ng each for *Ohh* and *Rtc*) shown in the *right panel* were transferred and hybridized with the ^{32}P-labeled MMV-1.0 DNA probe. The corresponding Southern blot is shown in the *left panel*. A prominent 1.1 kb hybridization band is observed in the *Mmv* lane. A very strong 1.1 kb hybridization band plus two faint bands, of 0.4 and 2.1 kb, can be seen in the *Mr* lane. A hybridization band of 1.1 kb can also be observed in both the *Ohh* and *Rtc* lanes. This confirms the presence of ~1 kb repeated DNA. **b** Electrophoretic separation of PCR products (*Mmv*, 60 ng; *Mr*, 40 ng; *Ohh*, 30 ng; *Rtc*, 20 ng) is shown in the *right panel*. The corresponding Southern blot shown in the *left panel* was obtained after hybridization with a mixture of Mmv-0.7, Mr-0.7, Ohh-0.7 and Rt-0.7 cervid satellite II DNAs as the probe. Mr-0.7 and Ohh-0.7 are cervid satellite II clones from Chinese muntjac and black tailed deer, respectively (Li et al., unpublished data). Four bands of 0.7 kb register (0.7, 1.4, 2.1 and 2.8 kb) are shown in the *Mmv* lane and a three band pattern of 0.7 kb register (0.7, 1.4 and 2.1 kb) is seen in the *Mr*, *Ohh* and *Rtc* lanes, confirming the presence of satellite II DNA products. *M* in the *right panels* denotes the standard DNA marker

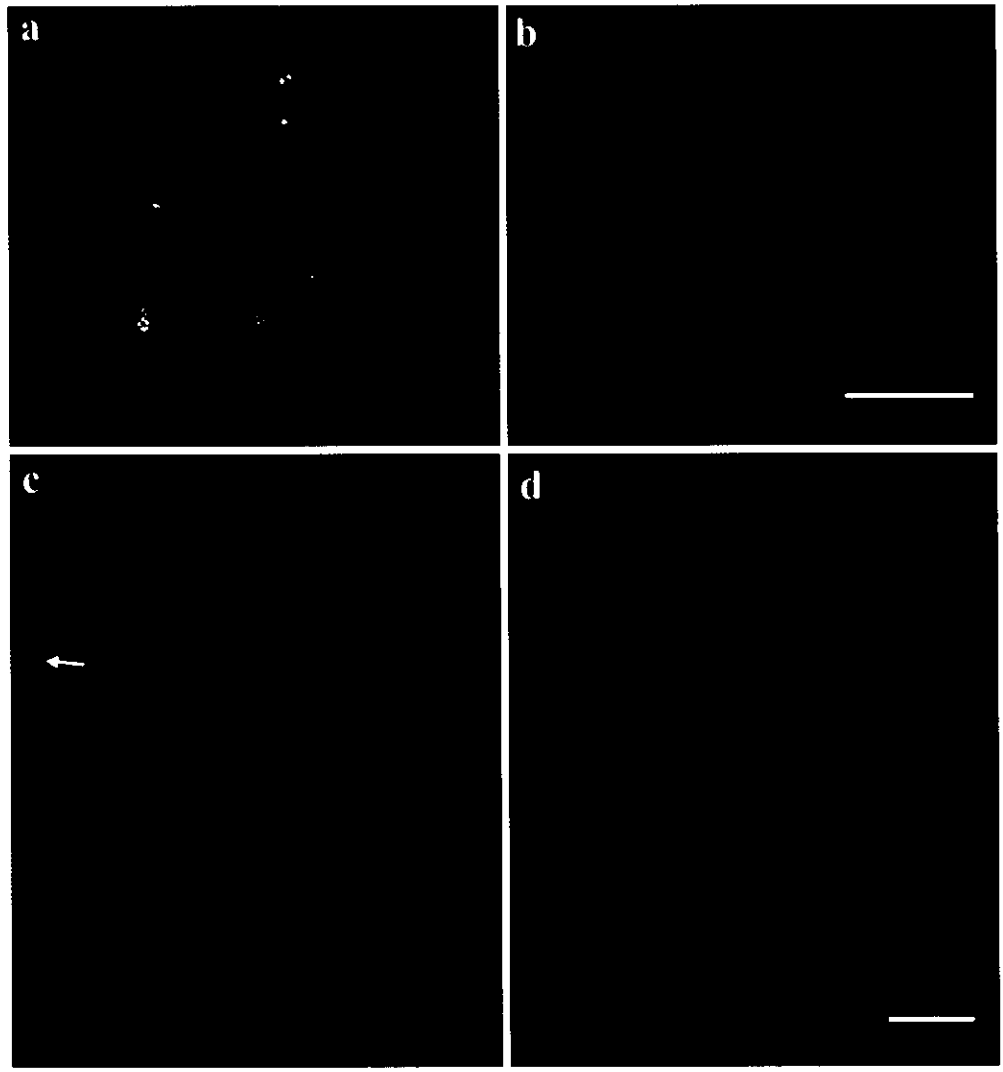


Fig. 5a–d Chromosomal localization of the satellite IV DNA clones and the satellite II DNA clone as well as immunofluorescence detection of kinetochores with human CREST serum. **a** Co-localization of satellite IV DNA clone MMV-0.1 (*red signals*) and satellite II DNA clone Mmv-0.7 (*green signals*) on centromeric regions of chromosomes from a male Indian muntjac. **b** Simultaneous detection of MMV-1.0 (*red signals*) and immunofluorescent signals of anti-centromere antibodies (*green signals*) on the centromeric region of chromosomes of the male Indian muntjac. Both fluorescent signals appear to be located together at the kinetochore sites. *Bar* in **b** represents 10 μm in **a** and **b**. **c** Localization of cervid satellite IV clone (MR-1.0) DNA to all the centromeric regions (except the Y chromosome, indicated by an *arrow*) and certain interstitial sites of chromosomes of the male Chinese muntjac. **d** Localization of cervid satellite IV clone (OHH-1.0) DNA to all the centromeric regions of female black tailed deer chromosomes. *Bar* in **d** represents 10 μm for **c** and **d**

plified using the same satellite II primer pairs from genomic DNA templates of the four deer species studied.

Since only limited amounts (Fig. 4a, right panel) of the 1 kb PCR products were obtained from genomic DNA templates of black tailed deer and woodland caribou with the initial OvDII-based primers, another primer pair was designed based on the DNA sequence of the newly identified MMV-1.0 sequence (see Materials and methods). Sufficiently more 1 kb PCR products were obtained from the black tailed deer and the woodland caribou with the new primer pair. The resulting clones were designated OHH-1.0 and RTC-1.0, respectively. DNA sequencing showed that the OHH-1.0 (892 bp) and the RTC-1.0 (891 bp) clones are also 60% AT rich, are devoid of any apparent internal subrepeats and share a very high degree of sequence homology between each other (93%) and between the 1 kb clones obtained from the muntjacs (Fig. 3). Therefore, all four 1 kb clones appear to belong to the same cervid satellite IV DNA family. Copy number estimation showed that about 0.06% of the black tailed genome and about 0.03% of the caribou genome contains this satellite DNA sequence.

Chromosomal mapping of cervid satellite IV DNA sequences

The FISH experiment initially localized cervid satellite IV (MMV-1.0) DNA exclusively at the centromeres of all chromosomes in the complement of Indian muntjac (data not shown). A subsequent FISH study was conducted by simultaneously hybridizing digoxigenin-labeled cervid satellite II (Mmv-0.7) probe with biotinylated cervid satellite IV (MMV-1.0) probe onto metaphase chromosomes of the male Indian muntjac. The satellite II DNA signals (green fluorescence) were located at the centromeric regions and at specific interstitial regions of the chromosomal arms as reported by Li et al. (2000b). The hybridization signals of satellite IV (red fluorescence) co-localized with the satellite II signals at the centromeric regions of all chromosomes (Fig. 5a). In addition, sequential CENP immunofluorescence and FISH experiments with satellite IV DNA probe (MMV-1.0) were performed. The satellite IV DNA was found to be closely associated with the CENPs (green fluorescence) in the kinetochore domains (Fig. 5b). Furthermore, FISH

studies localized the cervid satellite IV clone (MR-1.0) of Chinese muntjac to the centromeric regions of all Chinese muntjac chromosomes with the exception of the Y chromosome, which appeared to lack detectable hybridization signals (Fig. 5c). In the black tailed deer, the hybridization signals of biotin-labeled black tailed deer satellite IV clone (OHH-1.0) were also exclusively observed at the centromeric regions of all chromosomes of the species (Fig. 5d).

Discussion

In the course of amplifying cervid satellite II DNA from the genome of two muntjac species using a pair of primer sequences derived from white tailed deer satellite clone (OvDII) (Qureshi and Blake 1995), a prominent ~1 kb product was obtained in addition to the expected 0.7 kb satellite II DNA elements from the muntjac species as well as from black tailed deer and Canadian woodland caribou. Southern blot analysis, DNA sequencing and FISH studies indicated that the ~1 kb elements belong to a new family of cervid centromeric satellite DNA that is 60% AT-rich in contrast to the three previously identified cervid satellite DNAs (I, II and III), which are GC-rich. To our knowledge, it has not previously been reported that two different centromeric satellite DNA sequences could be simultaneously amplified with a single primer pair. Extremely high sequence homology (over 90%) was found among those ~1 kb clones isolated from the four deer species studied. Centromeric DNA consists of satellite repeats, which are thought to evolve rapidly among eukaryotic genomes (Henikoff et al. 2001). However, sequence homology of certain satellite DNA clones can be observed among related species. For example, cervid satellite I DNA clones with monomeric sequences sharing reasonably high sequence similarity were found among a number of deer species (Lee et al. 1997a). These cervid satellite I DNA monomers were shown to contain internal subrepeats organized in a higher-order fashion (Lee and Lin 1996). In the present study, no internal subrepeats were observed in the sequenced ~1 kb satellite DNA clones.

Although a regular A-type banding pattern often characteristic of satellite DNA (Horz and Zachau 1977) was not observed in the Southern blot hybridization of ~1 kb satellite DNA probe (MMV-1.0) to differentially digested Indian muntjac genomic DNA, the PstI digested sample did exhibit a 1 kb register pattern, along with other hybridization bands. This suggests that a portion of this repeated DNA family is organized as 1 kb tandem repeats in the genome of this deer species. Since the repeated sequence is localized at the centromeres of all deer species examined, it should be considered as a new cervid centromeric satellite DNA that was probably not detected previously due to its comparatively reduced abundance in these genomes (0.03%–0.16%). In general, some centromeric satellite DNA families may predominate in a given species because of dynamic evolution processes involving amplification and homogenization mechanisms (e.g., human α-satellite DNA, bovine satellite II and mouse major satellite) (Willard and Waye 1987; Nijman and Lenstra 2001; Wong et al. 1990). In the mean time, less prominent, or newly created satellite DNAs may coexist with the more prevalent and older satellite DNAs with some of the new sequences potentially being derived from elements of pre-existing satellite sequences (Buntjer et al. 1998). Whether this newfound repetitive DNA is derived from elements of other cervid satellite DNA sequences, particularly the satellite II family, remains to be explored. This newly discovered centromeric repeated DNA could be a relatively "new" satellite DNA only present in a small number of closely related species. Further studies should be carried out to examine whether this newly found repeated DNA is also present in the genomes of other deer species. On the other hand, FISH studies revealed that hybridization signals of this satellite DNA are co-localized with the CENPs in the centromeres and, due to the fact of its sequence conservation, one could speculate that the newfound repeated sequence could be a candidate for the functional "core sequence" of the cervid centromere.

Acknowledgements The authors would like to thank Dr. J.B. Rattner for supplying them with the human CREST sera for this study. These investigations were supported by research grants from the National Health Research Institute (NHRI-EX90-8933SL) and National Science Council (NSC90-2745-P040–200) of Taiwan, and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

Beridze T (1986) Satellite DNA. Springer, Berlin Heidelberg, New York

Bogenberger J, Schnell H, Fittler F (1982) Characterization of X-chromosome specific satellite DNA of *Muntiacus muntjak vaginalis*. Chromosoma 87:9–20

Bogenberger JM, Neumaier PS, Fittler F (1985) The muntjak satellite IA sequence is composed of 31-base-pair internal repeats that are highly homologous to the 31-base-pair subrepeats of the bovine satellite 1.715. Eur J Biochem 148:55–59

Bogenberger JM, Neitzel H, Fittler F (1987) A highly repetitive DNA component common to all cervidae: its organization and chromosomal distribution during evolution. Chromosoma 95:154–161

Buntjer JB, Nijman IJ, Zijlstra C, Lenstra JA (1998) A satellite DNA element specific for roe deer (*Capreolus capreolus*). Chromosoma 107:1–5

Choo KH (1997) The centromere. Oxford University Press, Oxford, New York, Tokyo

Choo KH (2000) Centromerization. Trends Cell Biol 10:182–188

Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293: 1098–1102

Horz W, Zachau GH (1977) Characterization of distinct segments in mouse satellite DNA by restriction nucleases. Eur J Biochem 73:383–392

Lee C, Lin CC (1996) Conservation of a 31 bp bovine subrepeat in centromeric satellite DNA monomers of *Cervus elaphus* and other cervid species. Chromosome Res 4:427–435

Lee C, Ritchie DBC, Lin CC (1994) A tandemly repetitive, centromeric DNA sequence from the Canadian woodland caribou (*Rangifer tarandus caribou*): its conservation and evolution in several deer species. Chromosome Res 2:293–306

Lee C, Court DR, Cho C, Haslett JL, Lin CC (1997a) Higher-order organization of subrepeats and the evolution of cervid satellite I DNA. J Mol Evol 44:327–335

Lee C, Wevrick R, Fisher RB, Ferguson-Smith MA, Lin CC (1997b) Human centromeric DNA. Hum Genet 100:291–304

Li Y-C, Lee C, Hseu T-H, Li S-Y, Lin CC (2000a) Direct visualization of the genomic distribution and organization of two cervid centromeric satellite DNA families. Cytogenet Cell Genet 89:192–198

Li Y-C, Lee C, Sanoudou D, Hseu T-H, Li S-Y, Lin CC (2000b) Interstitial colocalization of two cervid satellite DNAs involved in the genesis of the Indian muntjac karyotype. Chromosome Res 8:363–373

Lin CC, Sasi R, Fan Y-S, Chen Z-Q (1991) New evidence for tandem chromosome fusions in the karyotypic evolution of Asian muntjacs. Chromosoma 101:19–24

Maniatis T, Fritsch EF, Sambrook J (1982) Cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Nijman IJ, Lenstra JA (2001) Mutation and recombination in cattle satellite DNA: a feedback model for the evolution of satellite DNA repeats. J Mol Evol 52:361–371

Plucienniczak A, Skowronski J, Jaworski J (1982) Nucleotide sequence of bovine 1.715 satellite DNA and its relation to other bovine satellite sequences. J Mol Biol 158:293–304

Qureshi SA, Blake RD (1995) Sequence characteristics of a cervid DNA repeat family. J Mol Evol 40:400–404

Scherthan H (1991) Characterization of a tandem repetitive sequence cloned from the deer *Capreolus capreolus* and its chromosomal localization in two muntjac speices. Hereditas 115: 43–49

Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF (2001) Genomic and genetic definition of a functional human centromere. Science 294:109–115

Sullivan BA, Schwartz S (1995) Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of function centromeres. Hum Mol Genet 4:2189–2197

Vafa O, Shelby RD, Sullivan KF (1999) CENP-A associated complex satellite DNA in the kinetochore of the Indian muntjac. Chromosoma 108:367–374

Willard HF, Way JS (1987) Hierarchical order in chromosome specific human alpha satellite DNA. Trends Genet 3:192–198

Wong AKC, Biddle FG, Rattner JB (1990) The chromosomal distribution of the major and minor satellite is not conserved in the genus *Mus*. Chromosoma 99:190–195

Chromosoma (2002) 111:184–191
DOI 10.1007/s00412-002-0197-1

ORIGINAL ARTICLE

Fukashi Shibata · Masahiro Hizume

The identification and analysis of the sequences that allow the detection of *Allium cepa* chromosomes by GISH in the allodiploid *A. wakegi*

Received: 6 December 2001 / Revised: 9 May 2002 / Accepted: 15 May 2002 / Published online: 14 June 2002

Abstract In *Allium wakegi*, which is an allodiploid species between *Allium cepa* and *Allium fistulosum*, each genome can be clearly distinguished using genomic in situ hybridization (GISH). Genomic DNA of *A. cepa* and *A. fistulosum* is differentiated both qualitatively and quantitatively. We wanted to isolate nucleotide sequences that give genome-specific signals on *A. cepa* chromosomes in GISH experiments in *A. wakegi*. We isolated 23 clones that show GISH-like signal patterns in fluorescence in situ hybridization (FISH) and analyzed their distribution in the *A. cepa*- and *A. fistulosum*-derived genomes of *A. wakegi*. There was considerable variation in the abundance and distribution of these cloned sequences on the chromosomes of the two species. The degree of *A. cepa* specificity varied among the clones. Twenty-two of the clones showed an even distribution over most chromosome arms with some clustering in the pericentromeric regions, but one clone showed very distinct terminal signals on some chromosomes. Whereas these sequences are not specific for *A. cepa*, changes in bases in nucleotide sequences and in their amount result in genome-specific characteristics in GISH experiments.

Introduction

All species have a unique genome, and this is often evident in their karyotype through differences in, for example, genomic size, chromosome number and chromosome shape. In higher plants, even closely related species can differ in these properties. In addition to obvious karyological differences, use of genomic in situ hybridization (GISH) has demonstrated that even in pairs of species of very similar karyotype it is possible to differentiate the two parental genomes in hybrids (Parokonny et al. 1992). Although GISH has now been applied to many hybrids, a major question remains as to what kinds of sequence are important in the differentiation of genomes by this technique. To reveal the sequences relating to genomic differentiation in plants of large genomic size it is useful to understand the mechanisms of genomic evolution that accompany plant speciation.

Allium is a very large genus of about 750 species (Hanelt et al. 1992; Mes et al. 1999) with large genomes that range from 15.2 to 149 pg per 2C nucleus (Bennett and Leitch; http://www.rbgkew.org.uk/cval/homepage.html). Genomic in situ hybridization experiments have been done in many *Allium* species and hybrids (Hizume 1994; Friesen et al. 1997; Peterka et al. 1997; Shigyo et al. 1998; Khrustaleva and Kik 2000), and even among species classified in the same section genomes can be successfully identified. *Allium cepa* and *Allium fistulosum* are classified in the subgenus *Rhizirideum*, section *Cepa*, and have similar karyotypes, yet, using GISH, it is possible to identify the parental genomes in their hybrids (Hizume 1994; Shigyo et al. 1998; Khrustaleva and Kik 2000). What kinds of sequences determine GISH signals? Elucidation of genomic differences between these two related species will give important clues to the understanding of the process of genomic evolution. *A. cepa* and *A. fistulosum* are useful for this kind of experiment as there is an allodiploid species, *Allium wakegi* Araki (*A. cepa*×*A. fistulosum*), in the cells of which the two genomes can clearly be distinguished by GISH (Hizume 1994). The aim of this study was to isolate genome-characterizing sequences producing GISH signals and analyze their chromosomal distribution to gain an understanding of genomic evolution.

Edited by: F. Ishikawa

F. Shibata (✉)
Biological Institute, Faculty of Science, Ehime University, Matsuyama 790-8577, Japan
e-mail: shibatafukashi@hotmail.com

M. Hizume
Biological Institute, Faculty of Education, Ehime University, Matsuyama 790-8577, Japan

Materials and methods

Plant materials

Bulbs of *A. cepa* L. and *A. wakegi* Araki, and seeds of *A. fistulosum* L. were purchased in Matsuyama, Ehime, Japan. Bulbs of *A. cepa*

European Journal of Human Genetics (2002) 10, 495–498

www.nature.com/ejhg

SHORT REPORT

Mutations of *Cx26* gene (*GJB2*) for prelingual deafness in Taiwan

Yi-Chun Wang[1], Chiu-Yun Kung[2], Mao-Chang Su[3], Ching-Chyuan Su[2,4], Hsiu-Mei Hsu[1], Chin-Chu Tsai[1], Chyi-Chyang Lin[1] and Shuan-Yow Li*[,1]

[1]*Department of Life Sciences and Genetics Laboratory, Chung Shan Medical University, Taichung, Taiwan, Republic of China;* [2]*Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan, Republic of China;* [3]*Department of Otorhinolaryngology-Head & Neck Surgery, Chung Shan Medical University Hospital, Taichung, Taiwan, Republic of China;* [4]*Tian-Sheng Memorial Hospital, Tong Kang, Pin-Tong, Taiwan, Republic of China*

Mutations in the *Cx26* (*GJB2*) gene have been shown to be responsible for a major part of autosomal recessive non-syndromic inherited prelingual deafness. We have sequenced the coding region of *GJB2* gene from 169 Taiwanese patients with prelingual deafness and 100 unrelated normal individuals. In the deaf patients, three mutations were found: two novel mutations, 551G→A, and 299-300delAT, and one previously described mutation, 235delC. Four previously reported polymorphisms, 79G→A, 109G→A, 341A→G, and 608T→C, were also found in both deaf patients and normal individuals and one new possible polymorphism, 558G→A, which was only found in a patient. Interestingly, we did not find the 35delG allele, which is commonly found in the Caucasian population, either in the patients or in normal individuals we examined. Our data also showed 235delC to be the most common type of mutation found in *Cx26* mutants (approximately 57%). Therefore, based on our findings, we have developed a simple molecular test for the 235delC mutation and it should be of considerable help to those families to understand the cause of their children having the prelingual deafness.
European Journal of Human Genetics (2002) **10**, 495–498. doi:10.1038/sj.ejhg.5200838

Keywords: connexin 26; *GJB2*; prelingual deafness; Taiwanese

Introduction

The most common inherited sensory disorder is severe hearing impairment, which affects about 1 in 1000 children.[1] Currently, over 40 chromosomal loci have been identified for non-syndromic hearing loss through linkage studies (http://www.uia.ac.be/dnalab/hhh/). However, only 22 cloned nuclear genes have been shown so far to cause non-syndromic deafness (http://www.uia.ac.be/dnalab/hhh/hhhgenes.html), and nearly one half of the cases of non-syndromic deafness among Caucasian and Japanese are reported to be caused by mutations of *Cx26* gene.[2–4] *Cx26* (*GJB2*) encodes connexin 26 protein which is a member of the connexin protein family and can interact with connexin 32, connexin 46, and connexin 50 to form a hexameric of homotypic (composed of six identical connexin subunits) or heterotypic (composed of more than one species of connexins) half channel (connexon) of gap junctions. Such gap junctions play an important role in the local circulation of potassium ions in the inner ear.[5,6] Connexin 26 has been reported to be expressed in the stria vascularis, basement membrane, limbus and spiral prominence of cochlea.[7] Mutations in *Cx26* have also been found to be involved in various mechanisms that lead to the loss of hearing, such as interference with the proper oligomerisation or intracellular transport of *Cx26*, impairment of interaction between connexons in opposing cells, or block the recycling of potassium ions back to the endolymph of the cochlear duct after stimulation of the sensory hair cells.[8]

There is a lot of evidence indicating that mutations of *Cx26* might be a major contributor to prelingual deafness

*Correspondence: Dr S-Y Li, Department of Life Sciences and Genetics Laboratory, Chung Shan Medical University, Taichung, Taiwan, Republic of China. Tel: 886 4 2473 0022, ext. 1800; Fax: 886 4 2475 7412; E-mail: syl@csmu.edu.tw
Received 1 March 2002; revised 2 May 2002; accepted 7 May 2002

among Caucasians[9–12] and Japanese.[4] However, there have been significant differences in the types and frequencies of mutation of *Cx26* found between the two populations. 35delG has been reported to be the most common form of *Cx26* mutation among Caucasians, but 235delC is reported to be the most frequently found form of *Cx26* found among Japanese. We analysed the coding region of *Cx26* of 169 Taiwanese schoolchildren with prelingual hearing loss. Of the 169 children, 19 were having mutations in this region. Of these 19, eight were found to be homozygous for 235delC, two heterozygous for 235delC, four compound heterozygous for 235delC/299-300delAT, two heterozygous for 299-300delAT, one heterozygous 551G→A, and two heterozygous for 558G→A. Our data identified 235delC to be the most common type of mutation found in *Cx26* mutants (~57%).

Family studies were also performed on nine probands with *Cx26* mutations, eight of whom were shown to have inherited mutant allele(s) from the parents/relatives. The results of this study further illustrate the contribution of *Cx26* mutants to the pathogenesis of prelingual hearing loss in children.

Materials and methods

The present investigation, a connexin 26 analysis, was performed on children from the National Taichung School for the Deaf. A complete medical history was obtained from each child to record the age of onset of deafness and to exclude the possibility of such environmental causes as maternofoetal infection, perinatal complications, meningitis, mumps, prenatal or postnatal drug ototoxicity, and acoustic trauma. One hundred and sixty-nine children (103 males, 66 females), between 4 and 18 years of age had mild to profound prelingually sensorineural deafness. Family studies were performed on nine probands with *Cx26* mutation. One hundred randomly selected normal individuals were screened to be used as control subjects.

Using a QIAamp DNA Blood Kit (Qiagen), we obtained DNA samples from 5 ml of peripheral blood from each individual. The quality and quantity of purified genomic DNA were determined by running a 0.8% agarose gel and spectrophotometry.

The primer pair used in polymerase chain reaction and two additional internal primers used in DNA sequencing were: Cx26-AU (5'-TCTTTTCCAGAGCAAACCGC-3'), forward primer for PCR and sequencing; Cx873-L (5'-CTGGGCAATGCGTTAAACTGG-3'), reverse primer for PCR and sequencing; Cx501F (5'-GTGGCCTACCGGAGACATGAG-3'), primer for internal forward sequencing; AL (5'-GGACACGAAGATCGCTGCAG-3'), primer for internal reverse sequencing. The coding region of *Cx26* was PCR amplified by denaturation of template DNA at 94 C for 5 min, followed by 30 cycles of denaturation at 94 C for 1 min, annealing at 60 C for 1 min, extension at 72 C for 1 min, and finally at 72 C for 5 min. The PCR products were purified using a PCR Purification Kit (Qiagen). The purified DNA were then subjected to PCR-directed DNA sequencing using a DNA Sequencing Kit (Perkin-Elmer) and ABI Prism 310 Genetic Analyzer (Perkin-Elmer).

To verify the 235delC mutation, we used *ApaI* to digest the PCR product in a final volume of 20 μl: 5 μl of PCR product, 5 units of *ApaI*, 2 μl of 10X buffer, and 0.2 μl of 100X BSA. The reaction were carried out at 25 C for 2 h, and then run on a 2% agarose gel.

Results

The *Cx26* mutations observed in Taiwanese patients with prelingual non-syndromic hearing loss were two deletions, 235delC and 299-300delAT, and one substitution, 551G→A. In addition to the 235delC mutation, two other mutations were found to be novel. The prevalence of *Cx26* mutations are summarised in Table 1, and the effects of each mutation as well as the mutant allele frequency are summarised in Table 2. Our data also identified that 235delC is the most common mutant allele found in *Cx26*. However, we did not find any heterozygotes for mutant alleles among the 100 controls. Of the four 235delC/299-300delAT compound heterozygote deaf schoolchildren, two were sisters (assigned 043 and 053). Further sequence analysis of this F043/053 family, revealed the father to be heterozygous for 299-300delAT/wt and the mother heterozygous for 235delC/wt. However, both parents had normal hearing levels, indicating that the novel 299-300delAT mutation acts in a recessive manner. It is worth noting that the 35delG allele, which is common in Caucasian population, was not found in either the deaf patients or normal controls. We also studied eight families (14 family members in total) to learn more about the inheritance of individual mutations. It was found that all of the deaf children in these family families inherited the recessive mutant alleles from both parents except for patient 129, who had only one *Cx26* mutant allele and whose alleged father had a wild-type homozygote for *Cx26*.

Four polymorphic sites of *Cx26* were determined in the 100 unrelated normal controls as well as in our patient population (data not shown). The 341A→G and 608T→GC polymorphisms have been previously described by Kudo *et al.*[4] and the other two polymorphisms, 79G→A and 109G→A, have been previously described by Kelley *et al.*[13] In addition to these known polymorphisms, one new possible polymorphism, 558G→A, was detected in one patient (see discussion below).

Discussion

We have found three mutations, 551G→A, 299-300delAT and 235delC, in the *Cx26* gene among Taiwanese patients with prelingual deafness. Among those the 235delC mutation is the allele with the highest frequency in the Taiwanese deaf population. Similarly, the 235delC mutation has been described previously by Kudo *et al.* as being the

Table 1 The prevalence of *Cx26* mutations in 169 deaf patients

Genotype	*Individual found*	*Percentage*
551G→A/wt	1	0.59
235delC/235delC	8	4.73
235delC/wt	2	1.18
235delC/299-300delAT	4	2.36
299-300delAT/wt	2	1.18
Total	17	10.06

Table 2 Mutations and allele frequencies of *Cx26* found in 169 hearing impaired schoolchildren

Nucleotide change	*Amino acid change*	*Predicted effect*	*Allele frequency (chromosome no)*
235delC	L79C	Frameshift, stop at codon 81	6.70 (22)
299-300delAT	H100R	Frameshift, stop at codon 113	1.82 (6)
551G→A	R184Q	Missense mutation	0.29 (1)

most frequent allele found in Japanese deaf people,[4] though it has not been reported in the Caucasian population. Of the 169 deaf schoolchildren we studied, two had 235delC/wt heterozygotes. Since this mutation has been characterised as recessive[4] (and our data), it is very likely that these patients may have had other mutation(s) in other gene(s) which are responsible for the clinical manifestation of deafness. The 299-300delAT causes a frameshift and produces a premature Cx26 that stops at position 113. Patients 043 and 053 are siblings with the same genotype of 299-300delAT/235delC. Their father is 299-300delAT/wt heterozygote and their mother is 235delC/wt heterozygote. Both parents showed no hearing impairment. According to the results we obtained by studying the F043/053 family, this 299-300delAT mutation exhibited the characteristics of autosomal recessive heredity. Thus, the two 299-300delAT/wt heterozygote deafness school children, like 235delC/wt carrier, may also have had other mutation(s) in other genes responsible for deafness. Patient 129 was heterozygous for 551G→A/wt. Although his father was also deaf, the father did not show any mutation in *Cx26*. The mother of patient 129 was a normal individual but we could not obtain her DNA for sequence analysis. In patient 129's family (F129), the hearing loss of the father was clearly not caused by the mutation in *Cx26* and the child may have either inherited the other hearing loss factor(s) from the father and demonstrated the phenotype or perhaps due to the *de novo* 551G→A allele which caused the disease. Because the 551G→A is a missense mutation that causes R184Q substitution in Cx26 and the amino acid is located at the second extracellular loop, this mutation might affect the selective compatibility between different species of connexin proteins to form heterotypic functional channels.[14] However, whether the 551G→A is sufficient cause for the development of deafness or not needs further study. Another substitution found in this study was 558G→A. This substitution does not change the amino acid composition of Cx26 (silent mutation). However, this substitution was found in patient 117 only, whose parents showed no hearing impairment. Unfortunately, we could not obtain his parents' DNA to confirm whether he was a case of *de novo* mutation or whether there was mutation in another gene responsible for this phenotype. From the data we obtained, this silent mutation may represent a polymorphism rather than a pathogenic mutation. However, there is increasing evidence that the silent mutation might cause phenotypic variability by influencing splicing accuracy or efficiency.[15] Thus, we can not rule out the possibility that 558G→A may has the pathogenic effect. Further study is currently underway to verify this view point.

It is very interesting that we did not find any 35delG mutation in either our patients or in our controls. The 35delG allele has been demonstrated to be the most frequently found mutation to cause non-syndromic recessive prelingual deafness in Caucasians.[9,16] Tekin *et al*[16] suggested that the high frequency of 35delG allele may possibly arise from a founder effect and further enhanced by assortative mating in the population. This postulation could be true, since the oriental populations, Japanese[4] and Taiwanese (this study), do not exhibit this allele and the most common mutation in both oriental populations is 235delC, indicating that different causes between the Cx26 mutations of Caucasian and those of oriental populations.

One specific mutation, 235delC, accounted for the majority (approximately 57%) of *Cx26* mutant alleles in our study. Adding our findings to those of the Japanese study, we can say that 235delC is one of the most frequent disease mutations identified among oriental populations to date. Genetic counselling for prelingual deafness has been so far considerably hampered by the difficulty in distinguishing genetic and non-genetic deafness in families presenting with a single child. Based on the results presented in this study, we used restriction enzyme *ApaI* to verify the 235delC mutation and for the quick screening for such mutation. Since 235delC eliminates an *ApaI* site, therefore, we can identify an individual who is a normal, heterozygote, or homozygote for this mutation by *ApaI* digestion of the PCR product.

Acknowledgements

We thank Kuei-Hsiang Hung, Chung-Hsiang Yang, Yu-Che Cheng, Shi-Hsin Huang, and Chien-Hsun Wu for their technical assistance and to National Health Research Institute for support to CCL (NHRI-Ex90-8933SL). This work is supported by National Science Council, Republic of China (NSC 89-2745-P-040-002, NSC 89-2320-B-040-061, and NSC 90-2320-B-040-032).

References

1 Morton NE: Genetic epidemiology of hearing impairment. *Ann NY Acad Sci* 1991; 630: 16–31.

2 Zelante L, Gasparini P, Estivill X *et al*: Connexin26 mutations associated with the most common form of non-syndromic neurosensory autosomal recessive deafness (DFNB1) in Mediterraneans. *Hum Mol Genet* 1997; 6: 1605–1609.

3 Cohn ES, Kelley PM: Clinical phenotype and mutations in connexin 26 (DFNB1/GJB2), the most common cause of childhood hearing loss. *Am J Med Genet* 1999; 89: 130–136.

4 Kudo T, Ikeda K, Kure S *et al*: Novel Mutations in the connexin 26 gene (GJB2) responsible for childhood deafness in the Japanese population. *Am J Med Genet* 2000; 90: 141–145.

5 Salt AN, Melichar I, Thalmann R: Mechanisms of endocochlear potential generation by stria vascularis. *Laryngoscope* 1987; 97: 984–991.

6 Salt AN, DeMott JE: Ionic and potential changes of the endolymphatic sac induced by endolymph volume changes. *Hear Res* 2000; 149: 46–54.

7 Kelsell DP, Dunlop J, Stevens HP *et al*: Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. *Nature* 1997; 387: 80–83.

8 White TW: Functional analysis of human Cx26 mutations associated with deafness. *Brain Res Rev* 2000; 32: 181–183.

9 Denoyelle F, Weil D, Maw MA *et al*: Prelingual deafness: high prevalence of a 30delG 30delG mutation in the connexin 26 gene. *Hum Mol Genet* 1997; 6: 2173–2177.

10 Gasparini P, Estivill X, Volpini V *et al*: Linkage of DFNB1 to non-syndromic neurosensory autosomal-recessive deafness in Mediterranean families. *Eur J Hum Genet* 1997; 5: 83–88.

11 Guilford P, Ben Arab S, Blanchard S *et al*: A non-syndromic form of neurosensory, recessive deafness maps to the pericentromeric region of chromosome 13q. *Nature Genet* 1994; 6: 24–28.

12 Maw MA, Allen-Powell DR, Goodey RJ *et al*: The contribution of the DFNB1 locus to neurosensory deafness in a Caucasian population. *Am J Hum Genet* 1995; 57: 629–635.

13 Kelley PM, Harris DJ, Comer BC *et al*: Novel mutations in the connexin 26 gene (GJB2) that cause autosomal recessive (DFNB1) hearing loss. *Am J Hum Genet* 1998; 62: 792–799.

14 Krutovskikh V, Yamasaki H: Connexin gene mutations in human genetic diseases. *Mut Res* 2000; 462: 197–207.

15 Cartegni L, Chew SL, Krainer AR: Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002; 3: 285–298.

16 Tekin M, Akar N, Cin S *et al*: Connexin 26 (*GJB2*) mutations in the Turkish population: implications for the origin and high frequency of the 35delG mutation in Caucasians. *Hum Genet* 2001; 108: 385–389.