

# 臺、緒論

## 一 基因體學與生物資訊學的興起

1990 年在美國國家衛生研究院與美國能源部的主導下，成立人類基因體組織(Human Genome Organization, HUGO)，進而推動人類基因體計劃(Human Genome Project, HGP)，該計劃為人類二十世紀三大計畫之一(1940 年-曼哈頓原子彈計劃；1960 年-阿波羅登月計劃)，由美國、英國、法國、德國、中國與日本等六國科學家所組成的龐大團隊，經費大部份來自美國聯邦政府的「國家衛生研究院」(NIH)與英國的「衛爾康基金會」(WellcomeTrust)。1998 年由 Venter 所創立的「賽雷拉公司」(Celera Genomics)，則以異軍突起之姿投入破解人類基因密碼的競技場，以最先進的技術加上許多超級電腦的協助，以一家公司之力得以與全世界的研究力量相抗衡。基因體計畫的存在，意味著人們想要運用所產生之序列資料的企圖，如何結合分子及結構生物學、醫學、物理、化學及資訊等領域之科學，解讀基因體所蘊含的遺傳訊息，瞭解各個基因之結構與功能，以及其在不同生理及病理狀態所扮演的角色，基因體學(Genomics)因以衍然而生，其目標是瞭解生命體的完整基因體資料，進而辨識每一個基因，比對每個基因與其所組成的蛋白質，以及辨識每個蛋白質的結構與功能。隱含其中的是，期待能根據這些知識，深入了解生命運作過程並精確地操控生命

體的能力。

人類基因體(human genome)是由大約三十億鹼基對所組成。1980年代末期由美國生物學者主導發起之人類基因體計劃大規模地對人體細胞內所有染色體 DNA 進行定序，短短數年內產生了大量序列及相關資訊，如果將這些資訊儲存成文字檔，則需仰賴電腦建立巨型資料庫來儲存。近二十年來，這些資料庫的資料量正以指數速率成長，同時，生物學文獻之數量也以指數速率在成長。資訊技術方面的專家們管理著特殊的生物資訊軟體，連結高速電腦網路、高速電腦伺服器 and 強大儲存設施的電腦，將有關生物資訊的資料儲存在公用資料庫中。生物資訊最早開始於生物資料庫的建立，其中最著名的即是 GenBank。GenBank 現在是由 NIH (National Institute of Health, USA) 底下的「美國國家生技資訊中心」(National Center for Biotechnology Information, NCBI)來管理。GenBank 每天與 EMBL(European Molecular Biology Laboratory)資料庫和 DDBJ(DNA DataBank of Japan)資料庫進行同步交換。NCBI 將生物資訊學(Bioinformatics)定義為一種將分子生物、電腦科學和資訊技術融合成一門獨立學問的科學。生物資訊學主要的重點在於如何呈現、儲存、傳輸、進而分析生物相關的大量資訊。嚴格來說，生物資訊學是更大範圍的計算生物學(computational biology)的子集合，是模擬生物系統時，量化分析的應用技術。(Gynthia

Gibos. and Per Jambeck , 2002)

生物資訊技術可極快速、精確的對生物序列進行分析與探討，在短時間內處理大量資料，不僅效率遠超越傳統實驗方法，其所能提供的整體觀(holistic view)探索方式也是分子生物學家從未達成的夢想。該技術的快速進展，導源於資訊科技的卓越發展，更因分子生物和高速平行篩選技術(high-throughput screening)的進步，火速激增的核酸序列資料，不斷刺激生物資訊運算法、新軟體工具及生命資訊資料庫之開發，供需之間跨領域的巧妙配合，徹底改變傳統生物科技的研究方式，藉由「生物資訊」這項有力的工具，過去需耗費數年之工作，幾乎可縮短在數週或數天內完成。其旨在於加速生命科學的新發現與創造全面性的理解生物原理。二十一世紀正式邁入生物資訊時代，可望透過這項新的工具技術，讓人類生物技術有重大迅速的突破。總而言之，生物資訊學之主要目標不是發展出精緻的演算法，或是完成神秘難解的分析，而是找出生命運作的秘密。

## 二 功能性基因體學與蛋白質體學之未來發展

生物個體的遺傳與功能資訊儲存在 DNA、RNA 及蛋白質中，這些大分子是由已經充分了解的化學物質所組成，它們由特定的字母所代表：DNA 是由四種去氧核糖核酸組成(腺嘌呤 adenine、胸腺嘧啶 thymine、胞嘧啶 cytosine 與鳥糞嘌呤 guanine)，RNA 是由四種核糖

核酸組成(腺嘌呤 adenine、尿嘧啶 uracil、胞嘧啶 cytosine 與鳥糞嘌呤 guanine)，蛋白質是由二十種胺基酸組成。它們能以序列符號(單一字母)來表示，透過此些序列的排比，找出功能或形態類似的分子之間的關聯性，或演算出「可能相當重要」的結果，藉以詮釋以往從未被發現的資料之機會。基因序列不僅僅只是序列，它代表著立體的結構、分子形狀與構形改變、活化區域(active site)、化學反應以及分子間交互作用等諸多細節。

生物資訊資料庫中儲存了大量的序列資料，但如何才知道序列的哪個部分，控制不同生命的化學反應？假若已經了解某些蛋白質的結構與功能，要如何辨別新的蛋白質的編碼，該如何在密碼中找到有意義的新「字」？人類的 DNA 序列中大概僅有 5% 的部份是能產生蛋白質的基因，因此須從人類基因體中辨認出有功能的基因。人類基因可簡單分為：啟動子(promoter)，5 端非轉譯區域(5' untranslated region, 5'UTR)，表現子(exons)，內含子(introns)，3 端非轉譯區域(3' UTR)，聚腺嘌呤化作用位置(polyadenylation site)等幾個部份。其中僅有表現子才攜帶產生蛋白質的訊息。因此辨認基因的主要任務即是從 DNA 序列中找出起始密碼子(start codon)、終止密碼子(stop codon)與接合位點(splicing site)，進而將同一基因所有的表現子拼湊出來，建立一完整之基因。目前預測基因的方法大致可分為兩種：一是根據機率與統

計的方法，另一是尋找相似性(similarity or homology)的方法。隨著已知基因的大量累積和生物資料庫的日趨完備，大多採用尋找相似性的方法。NCBI 提供一個方便易使用之整合型檢索系統(search and retrieval system)Entrez (<http://www.ncbi.nlm.nih.gov/Entrez/>)，以利調閱 GenBank 之資料。運用 NCBI 的 BLAST(Basic local Alignment Search Tool) Web 介面 (<http://www.ncbi.nlm.nih.gov/BLAST/>)搜尋比對，能快速進行生物序列的逐對(pair wise)排比、多重比對(multiple alignment)、親緣樹分析(phylogenetic analysis)、基元辨識(motif identificaton)等演算分析。

由於人類基因體計劃的提早完成，以及毛細管序列儀、生物晶片(Biochip)包括基因微陣列分析(microarray)和基因晶片(GeneChip)等多種基因體技術的進步，即正式宣告進入後基因體時代(post-genomic era)，預估此時的重點將轉移至功能性基因體學(functional genomics)、蛋白質體學(proteomics)和單一核苷酸多樣性(single nucleotide polymorphism, SNP)之上(NCBI Handbook chap.2, <http://www.ncbi.nlm.nih.gov/Sitemap/index.html#NCBIHandbook>)。

目前生物資訊經序列或構造相似性之分析後，仍然有諸多無法預測之基因功能與其扮演的角色，未來生物學主要之挑戰是發展確認或定義此類基因的方法。功能性基因體學即是對大規模基因體之

mRNA 和蛋白質表現做有效率有系統之整體性分析。藉由鑑定蛋白質之開放閱讀框架(open reading frame, ORF)、基因的系列分析(Serial Analysis of Gene Expression, SAGE)或微陣列分析(microarray), 以定義序列並探討其差異性基因表現。蛋白質體學則是研究生物系統中蛋白質之表現, 包含: 相對表現量、轉譯後修飾、穩定性與環境改變時細胞之反應, 試圖瞭解無法以資料庫相似性比對所確認之蛋白質功能。其可用以預測蛋白質之功能及蛋白質間之交互作用。

將生物資訊的資料作不同層次的應用, 則會發展出不同預測或模擬的結果, 例如: 根據新發現的基因比對序列, 進而發展出功能方面之假說; 或者將已知的蛋白質立體結構拆解, 以尋找足以預測蛋白質如何摺疊的序列樣式; 又或者模擬蛋白質與細胞中的代謝物, 如何協同工作讓細胞得以正常運作。

近百年來, 人類在應用生物學領域中所關切的許多重要問題, 例如: 如何預防感染、治療疾病? 在人口爆炸的狀況下, 如何生產足夠多的食物? 研發新藥、農藥、混種植物與開發環境整治的生物技術等正是驅動生物資訊學研究之主要動力。生物資訊學研究之工作範圍, 不僅是擷取、整理與呈現資料, 或是從全然不同的領域中得到靈感, 而是想藉由不同領域之科學與生物學之間找尋其關聯性, 以期從序列資料及蛋白質結構分析, 到代謝模型、族群與生態的大規模資料分

析，達到瞭解生命的歷史與進程之目的。

### 三 本研究動機與方向

近年來，由於人類基因體計劃的推展造成生物資訊資料庫(如：NCBI)之 EST 基因庫(Expression Sequence Tags Library)中核酸序列資料的大量累積，加以電腦運算能力日益強大及網際網路連結之速率加快，使得生物資訊資料庫之整合型檢索系統(Entrez)的使用，益加簡便。因此於人類基因中，發現許多懸而未知其功能的新穎蛋白質。對於這些未被定義或未解其生物角色之新穎蛋白質，激起諸多研究生命科學的學者，極欲探索其間奧秘之渴望。然而此類新穎的蛋白種類之繁，如何擇其要而為之，則是此項研究的重點之一。有鑑於此，係選擇具有  $FAD^+$ 、 $NAD^+$  或  $NADP^+$  等輔酶(coenzyme)的結合結構區之新穎氧化還原酵素，利用大腸桿菌表達系統表達與純化蛋白質，藉由 NCBI 網站進行核酸或胺基酸序列之比對，只能初步推斷出該類蛋白質可能具有的結構區，以此推論及探討其可能之催化反應，對其進行酵素動力學的分析，以確認酵素進行催化反應時在受質的親和力及酵素的代謝迴轉上所扮演的角色。至於其真正的功能與在人體生理上扮演的角色，則與中研院生化所黃銓珍副研究員之實驗室合作，以去除基因或過量表現基因的方式來作進一步的探討。

在探討未知酵素催化的代謝反應時，可將所欲探討的未知酵素分

成兩類。第一類酵素是某種已知酵素的同功酶(isozyme)，即該酵素催化的反應和已知的酵素可能相同，但是可能是因為其催化上的特性(例如  $K_m$ ,  $K_{cat}$ )或者是分佈的位置、調控的方法與已知的酵素不同所以在生理上可能有不同的功能。對於其催化反應的受質之確認可依據其同功酵素的受質為參考。第二種酵素則是透過與已知的結構區(domain)或基元(motif)比對後，經由蛋白一級構造之相似性推論得知酵素可能進行的催化反應。該類酵素往往可能是催化雙受質或兩個受質以上之酵素，其受質則必須由酵素本身的結構區中尋找線索，此類推演的過程往往要運用到酵素所在之位置，以及該位置可能具有的受質及生理之功能加以進行判斷。

若得知該酵素催化的代謝反應後可經由該酵素在全身各器官及組織的分佈推測出該酵素在生理上所扮演的角色。反之，若將酵素在生物體中去除(或增強)後對生物體生理上所造成的影響，以反向的方式推論該酵素可能在生理上所扮演的角色，進而推論該酵素可能催化的代謝反應。因此確認酵素催化的代謝反應，以及去除(或增強)該酵素並觀察生物體發生的變化是可以同時進行、相互印證，並對酵素之生理功能作進一步的探討。

欲了解一基因之功能，最直接的方法就是將該基因轉殖至活體細胞內，研究分析該基因的表現狀況。目前基因轉殖方法有：脂質體



(Liposomes) (Schaefer-Ridder *et al.*, 1982 ;Mannino and Gonld-Fogerite, 1988), 磷酸鈣-DNA 沉澱法 (Calcium phosphate-DNA coprecipitation) (Graham and Van der Eb, 1973), 病毒載體法(Viral vectors), 電破法 (Electroporation) (Xie *et al.*, 1990)等。魚類基因轉殖方面, 1985年Zhu等人, 首先以顯微注射法(Microinjection)將外來基因轉殖到金魚, 之後在魚類基因轉殖的方法又有許多改進, 例如電破法 (Electroporation), 精子載體法(Sperm mediator)(Orban *et al.*, 1992; Sin *et al.*, 1993; Symonds *et al.*, 1994)及粒子槍法(Particle gun)(Klein *et al.*, 1992)。由於中研院生化所黃銓珍副研究員自1995年即建立斑馬魚系統, 藉其斑馬魚基因轉殖方式進行功能分析, 利用顯微注射的方式在斑馬魚的受精卵中引入適當的遺傳物質以抑制(RNA interference, RNAi)(Summerton, 1999; Oates *et al.*, 2000)或者增強該酵素之表現, 透過對該斑馬魚發生之觀察, 可對該酵素之生理功能作初步的判斷。

為了以上述斑馬魚的模式, 進行酵素分析, 所以對欲探討的酵素, 除了其催化特性是未知的之外, 該酵素在人類及斑馬魚中應均有表現, 最終並將從斑馬魚模式中觀察的結果應用在人類。因此必須先在NCBI的資料庫中經過蛋白質相似性之相互比對, 找到人類及斑馬魚中相互對應之酵素。該酵素之cDNA可能涵蓋全長或部份的編碼區(coding region), 若資料庫中的序列僅包含部份的編碼區, 則可藉由5'

或/及 3' RACE PCR 得到其全長之序列，接著可以 PCR 將人類及斑馬魚全長 cDNA 放大以進行蛋白質之表現，並利用其進行完整的酵素動力學分析。此外已知的 cDNA 序列亦可用以製作探針調查該酵素在生物體中的分佈情形，或製作特定形式之 RNA 以抑制或增強斑馬魚模式中酵素的表現，在觀察斑馬魚的生理上作準備工作。

#### 四 本研究之概況與研究目的

於 NCBI 資料庫中經搜尋比對後，尋找到一對在人類和斑馬魚中皆有表現，且相互對應新穎的氧化還原酵素(為假設性的酒精去氫酶)。其 NCBI accession number 分別為 AI477544 (zebrafish)及 NP\_115679 (Human)。AI477544 (zebrafish)的 cDNA，涵蓋部分編碼區的 632 個核酸序列(參見附錄 1)。NP\_115679(Human)的 cDNA，涵蓋全長的 3003 個核酸及 345 個胺基酸序列(參見附錄 2)。由 NCBI 之 PubMed (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>) 搜尋，NP\_115679(Human)無其相關之文獻報導；而 AI477544(zebrafish)則比對到一斑馬魚之第三類酒精去氫酶(Alcohol dehydrogenase ; ADH3)(Asok K. Dasmahapatra,2001)。將 AI477544(zebrafish)與文獻中之 ADH3 (zebrafish)經核酸及胺基酸序列比對後發現，二者並不相同(參見附錄 3)。

將 AI477544(zebrafish)與 NP\_115679(Human)和人類已知的酒精去氫酶(7 個同功酶)胺基酸序列比對(參見圖五)後發現，此推定的(斑馬魚及人類) Alcohol dehydrogenase，相較於已知的人類或斑馬魚 Alcohol dehydrogenase 的胺基酸序列僅有部分的胺基酸相似，但不完全相同。經結合結構區(domain)及基元(motif)的比對搜尋，發現不論斑馬魚或人類皆具有二個相同的結構區，分別為 adh\_short 及 SCP2。

adh\_short(short-chain alcohol dehydrogenases)已知是  $\text{NAD}^+$  或  $\text{NADP}^+$  依賴的氧化還原酵素。此類酵素，至少具有兩個結合結構區，第一個通常為輔酶(coenzyme)/ $\text{NAD}^+$  的結合結構區(Benyajati C *et al.*,1981)，第二個結合結構區則是受質結合位置，此受質的結合乃是決定此酵素蛋白質的特異性(specificity)及催化作用(catalysis)。

SCP-2(sterol carrier protein 2)屬於固醇類轉移家族(sterol transfer family)，涉及固醇類的結合與細胞內的膽固醇(cholesterol)及許多脂質(lipids)的運輸(Szyperski T *et al.*,1993)。此結合區發現於 SCP2 protein，其 C 端序列相同於 estradiol 17-beta-dehydrogenase [EC:1.1.1.62]。

由於胺基酸序列之初步比對發現，此酵素不同於人類已知的酒精去氫酶，推測可能為一新穎酵素，所以可由此二結合結構區來找尋可

能之催化反應受質，藉以推論其在酵素反應上扮演之角色。由於此人類之新穎酵素，其編碼區完整的核酸序列已公佈，所以可直接經 PCR 放大人類的 cDNA 基因庫而得之。然而斑馬魚之編碼區經搜尋比對後仍無完整之核酸序列，因此利用 RACE PCR 之方式取得其全長之序列。

於 GeneCards™ (<http://bioinfo.weizmann.ac.il/cards/index.html>) 網站中搜尋得知，NP\_115679(Human)其基因之 genomic location: 位於 chromosome 9q32，Electronic Northern 中顯示此基因在骨髓(Bone marrow)、脾臟(Spleen)、腦(Brain)、心臟(Heart)、骨骼肌(Skeletal muscle)、肝臟(Liver)、胰腺(Pancreas)、前列腺(Prostate)、腎臟(kidney)和肺臟(Lung)皆有存在。其具有十一個表現子(Exons)，轉錄長度(transcript length)為 3345bp，轉譯長度(translation length)為 418 residues (參見附錄 4)。

根據上述在 NCBI 之資料庫中，經過搜尋比對後，尋找到一對感興趣之新穎的氧化還原酵素(一假設性之 17- $\beta$  Hydroxysteroid dehydrogenase )，該新穎酵素在斑馬魚和人類中都有表現且相互對應，其 NCBI 編號分別為 AI477544 (Zebrafish)與 NP\_115679(Human)。為了確認該新穎酵素之功能與探討其功能在生理上所扮演的角色，現

階段本研究之目的在於利用 NCBI 網站中之生物資訊資料庫的功能及資料，以一連串的分析比對和搜尋之動作，確認一可能為人類和斑馬魚之新穎蛋白質(酵素)，並藉其聯結分析序列之功能，獲得可能之完整開放閱讀框架(蛋白質編碼區域)序列。接下來以此完整序列建構該新穎酵素之全長 cDNA，並表現與純化此酵素之蛋白質，找出可能參與酵素反應之受質，藉由酵素動力學分析該酵素之特性，推測此新穎酵素可能之酵素反應及此反應的重要性。

## 貳、材料與方法

本研究方向主要分為二個部分進行，一是利用網際網路中與生物技術相關之工具網站的資料庫(data base)和一些應用程式，作搜尋、比對、分析和資料整合的工作，找出新穎之酵素。另一方面利用重組 DNA 的技術和方法來表達此未知之蛋白質，並分析其在生物體中所扮演之功能與角色。

### 一 NCBI 生物資訊網路搜尋

#### (一) Entrez

以 zebrafish putative(novel/hypothetical/new/unknow)dehydrogenase 為關鍵字詢問核苷酸(nucleotide)序列資料庫(參見附錄 10-網頁 1)，對有興趣之序列進行檢索。

選定一個 NCBI accession number : AI477544 的斑馬魚 (zebrafish) cDNA clone 其核酸序列的 5'端相似於 WP : C48B11.3 CE05427 Alcohol dehydrogenase，此序列為片段序列(partial sequence)，缺少 PolyA Tail(參見附錄 1)，其組織來源為胚胎 26 體節(26 somite embryos)、成魚肝臟和胚胎盾狀期(shield stage embryos)。

#### (二) BLAST(Basic Local Alignment Search Tool)：比對搜尋相似序列

(1)斑馬魚片段序列 blast to nr(non-redundant database)(參見附錄

10-網頁 2)

利用 NCBI 中 BLAST 網頁的 Translated BLAST Searches 項目中之 Nucleotide query-Protein db[blastx](參見附錄 10-網頁 3)，搜尋比對此斑馬魚(Danio rerio)片段序列與人類(Homo sapiens)比對到之所有相似序列，並比較其相似度(identity)，將比對到之人類的相似序列以 BLAST 網頁的 Pairwise BLAST 項目中之 BLAST 2 Sequences (參見附錄 10-網頁 3)來區別其間之差異，篩選與輸入之片段序列最相似者。

(2)擷取人類的開放編閱架構(Open Reading Frame, ORF)之核苷酸或胺基酸(參見附錄 2)序列 blast to est other (參見附錄 10-網頁 4) 利用 NCBI 中 BLAST 網頁的 Translated BLAST Searches 項目中之 Nucleotide query-Translated db[tblastx] 或是 Protein query-Translated db[tblastn] (參見附錄 10-網頁 3)，以人類與該斑馬魚片段序列相似度最大者之核酸或胺基酸序列，搜尋比對相似之斑馬魚序列，利用 BLAST 網頁的 Pairwise BLAST 項目中之 BLAST 2 Sequences (參見附錄 10-網頁 3)來區分其間之差異，並檢索是否有比對回原始斑馬魚序列。

(3)搜尋斑馬魚序列是否延長或具有完整序列

a.分別擷取斑馬魚(AI477544)前端(5')和尾端(3')序列之六十個鹼

基，各自 blast to est or est other

以 NCBI 中 BLAST 網頁的 Translated BLAST Searches 項目中之 Nucleotide query-Translated db[tblastx](參見附錄 10-網頁 3)，搜尋比對。選取比對到之序列相似性為 100% 之斑馬魚序列。將此些序列進行序列組裝及比對之工作，以檢視其原始序列是否延長。

b.將斑馬魚片段序列輸入斑馬魚序列資料庫中 blast to EST/ EST Traces (參見附錄 10-網頁 5)

於 NCBI 中 BLAST 網頁的 Genomic BLAST pages 項目中之 Zebrafish(參見附錄 10-網頁 3)，搜尋相似性 100% 之序列，並進行序列組裝，檢視其原始序列是否延長或為完整序列。

### (三) 生物資訊應用軟體之分析/處理

#### (1) Alignment Homo sapiens and Danio rerio cDNA(and/or amino acid) sequence

a.利用一些網站提供之服務(軟體)將斑馬魚之核酸序列轉譯為胺基酸序列，分別比對人類和斑馬魚之核酸及/或胺基酸序列，藉由兩者之比對試圖找出可能之起始密碼子(AUG)及終止密碼子(參見附錄 5)。

b.設計實驗所需之 RACE PCR 引子(primer)

對於蛋白質結構和表達方面的研究，獲得一全長的 cDNA 是非常重要的。由於反轉錄酵素無法由一個全部的基因序列來讀



出一股 cDNA 的 5'端序列，則其序列就可能被遺漏，問題特別出現在極大的基因轉錄物件。一個未知序列的 5'或 3'的 mRNA (messenger RNA)，可由 mRNA 為模板來放大核酸序列之方法已發展出來，稱之為 RACE(Rapid Amplification of cDNA Ends) PCR。由於此選定之斑馬魚基因缺乏 3'端之完整序列因此將以 3' RACE PCR 的方式來初步獲得其 3'端的序列。分別設計二段 C 端的引子(參見表一)，藉由兩階段之 3' RACE PCR 進行放大作用。

#### (四) 其他工具網站之應用

將 RACE PCR 放大之核酸序列與原來網路選定之斑馬魚片段序列(accession number AI477544)相互比對，藉由序列重疊區域(overlap)進行組裝，初步構築成可能之原始完整序列(open reading frame)。依據此序列設計正反兩股之引子，內含有特定之內切酵素位點( *Sal* I / *Hind* III ) (參見表一)，以放大合成所欲得之全長 cDNA。

由 PCR 放大之 cDNA，經序列分析後所得確定之斑馬魚新穎基因的全長核酸序列與序列相似比對所對應之人類(accession number NP\_115679)全長核酸序列，利用以下之各項比對分析網站與應用程式進行分析：

(1)斑馬魚全長序列 blast to nr

於 NCBI 中 BLAST 網頁的 Translated BLAST Searches 項目中之 Nucleotide query-Protein db[blastx]，將此斑馬魚的完整核酸序列及胺基酸序列與基因資料庫中的序列進行搜尋比對。篩選比對到之人類的相似序列，找尋相似性最高者，並再次確認是否比對到相同之序列。

(2)構成酵素活性之演化保留區域搜尋(conserved domain search)

在 NCBI 中 BLAST 網頁之 Protein BLAST 項目中 PSI-and PHI-BLAST 或 Search for served domains 項目中 Search the Conserved Domain Database using RPS-BLAST(參見附錄 10-網頁

3)，分別輸入斑馬魚與人類胺基酸序列搜尋。

(3)特殊結合結構區或基元相似性搜尋和比對

利用 NCBI 之 Entrez 檢索系統下的資料庫(ex: Domain/PuMed)找尋具有相同結合結構區或基元之已知蛋白質(或酵素)，將找到之各酵素胺基酸序列以生物資訊應用軟體進行比對，分別就編碼區與基元來比對，以判斷其相似程度。

(4)查尋人類該新穎蛋白質基因所在位置及分布

於 Gene Cards™網站(<http://bioinformatics.weizmann.ac.il/cards/>)，參見附錄 10-網頁 6)或至 NCBI 之 Uni-Gene 資料庫詢問(參見附

錄 6)，輸入人類該新穎酵素 NCBI 編號 NP\_115679/NM\_032303，點選 Display 進入(參見附錄 10-網頁 7)，即可得知此蛋白之基因位置(chromosome 上)和組織分佈，並可藉由此網站和其它相關網站聯結，取得其轉錄、轉譯上之資訊(參見附錄 4)。

#### (5) 查尋可能之代謝反應路徑

參考所查到之演化保留區域(特殊基元)推測此酵素之反應，就其可能之酵素名稱或 EC (Enzyme Commission) numbers，利用 KEGG 網站(<http://www.genome.ad.jp/kegg/>，參見附錄 10-網頁 8)之 KEGG-Table of contents(參見附錄 10-網頁 9)點選 Enzyme: EC number classification 或 Enzyme 選擇酵素類別或輸入酵素名稱搜尋(參見附錄 10-網頁 10)，藉尋找到之代謝反應路徑，預測及篩選可能之酵素反應受質及參與之代謝反應路徑。

## 二 材料與實驗方法

實驗部分主要分三個階段來進行，第一階段是由找尋到之可能為新穎酵素的現有斑馬魚片段序列，經由實驗方法將完整之序列拼湊組裝。接下來藉組裝後可能為此斑馬魚之完整序列與 NCBI 上比對到之人類全長序列，以 PCR 方式建構並轉形到合適的表達質體。最後再以其表現系統純化出蛋白質，並分析其功能與特性。

### (一) 材料

菌種(Bacteria Strain)：*Escherichia coli* M15 strain、pDM1.1 plasmid、*Escherichia coli* JM109 strain 及質體(Plasmid)：pQE-9、pQE-31 plasmid (Villarejo and Zabin, 1974; Liu, 1994)，由中山醫學院營養科學研究所劉德中博士提供。prestained SDS-PAGE standards, low range 購自 Bio-Rad；restriction endonuclease(*Bam*H I/*Hind* III/*Eco*R I)、*T4* DNA ligase 購自 New England；imidazole、isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG)、Estriol、Estrone、5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol、16 $\alpha$ - Hydroxyestrone、 $\beta$ -Estradiol 與 Hydrocortisone 購自 Sigma；*Taq* DNA polymerase 購自 TAKARA；QIAprep spin miniprep kit 購自 QIAGEN；ABI PRISM<sup>TM</sup> dye terminator cycle sequencing ready reaction kit 購自 PERKIN-ELMER Applied Biosystems；Gel-M<sup>TM</sup> Gel Extraction System Kit 購自 VIOGEN；pGEM<sup>®</sup>-T Easy Vector、X-Gal 購自 Promega；Nucleo Spin Plus Miniprep Kit 購自 CLONTECH；4-Androstene-3,17-dione 購自 Aldrich。

## (二) 延長及放大原始 Z-HSD 完整序列

### (1) RACE (Rapid Amplification of cDNA Ends) PCR

由於經比對後發現此斑馬魚缺乏 3'端之編碼區，故利用 3'RACE PCR 來取得 3'端之 cDNA 序列。此聚合酶連鎖反應(PCR)分為兩

階段式進行。

a. First (1st) PCR

首先第一次之 3'RACE PCR 的 50 $\mu$ l 反應體積中內含 10 $\times$ Ex Taq DNA polymerase reaction buffer 5 $\mu$ l ; 2.5mM dNTP 混合物 4 $\mu$ l ; 3-1 RACE primer (0.5 $\mu$ g/ $\mu$ l)及 oligo(dT) primer (0.5 $\mu$ g/ $\mu$ l) (Marathon kit , CLONTECH)(參見表一)各 1 $\mu$ l ; 模板(template;0~7 day 斑馬魚胚胎 cDNA) 2 $\mu$ l ; Ex Taq DNA polymerase (5U/ $\mu$ l)(TAKARA) 0.5 $\mu$ l ; d2H<sub>2</sub>O (0.22 $\mu$ m Filter Unit) 36.5 $\mu$ l , 將此反應混合物置於 GeneAmp PCR System 2400 (PERKIN-ELMER Applied Biosystems), 進行 PCR 反應, 其條件為: 高溫 96 $^{\circ}$ C 下反應 30 秒; 在 65 $^{\circ}$ C 下反應 30 秒, 進行黏合作用(annealing); 在 72 $^{\circ}$ C 下反應 2 分鐘, 以行擴展作用(extension), 以上述之條件充分進行 45 cycles ; 在 72 $^{\circ}$ C 下反應 15 分鐘, 最後儲存於低溫 4 $^{\circ}$ C 中。

b. nest PCR

將 1st PCR 的產物稀釋 50 倍之後作為模板(template)。在第二階段的 3'RACE PCR 之反應體積 100 $\mu$ l 中內含 10 $\times$ Ex Taq DNA polymerase reaction buffer 10 $\mu$ l ; 2.5mM dNTP 8 $\mu$ l ; 3-2 primer (0.5 $\mu$ g/ $\mu$ l)及 oligo(dT) primer (0.5 $\mu$ g/ $\mu$ l) (參見表一)各 1 $\mu$ l ; 模板(1st PCR product 50 $\times$ dilution ) 2 $\mu$ l ; Ex Taq DNA polymerase (5U/ $\mu$ l) 0.5 $\mu$ l ; d2H<sub>2</sub>O 77.5 $\mu$ l , 進行 PCR 反應, 其條件為: 在高溫 96 $^{\circ}$ C 下反應 30 秒; 在 65 $^{\circ}$ C 下反應 30 秒; 在 72 $^{\circ}$ C 下反應 2 分鐘, 以

上述之條件充分進行 40 cycles；在 72°C 下反應 15 分鐘，最後儲存於低溫 4°C 中。

3'RACE PCR 產物預期的大小則由 1% 電泳瓊膠(agarose-LE，Amersham LIFE SCIENCE;USB)經由電泳分析(electrophoresis)，並將其切下以 Gel-M™ Gel Extraction System Kit (VIOGEN)純化。

## (2)次選殖(subcloning)

### a. 聯結反應(ligation)

將經由 3'RACE PCR 放大之 DNA 片段與載體 pGEM<sup>®</sup>-T Easy Vector (Promega)，利用 T4 DNA ligase 在 16°C 下進行聯結反應至少 4 小時至 12 小時。其反應總體積為 10μl，嵌入物(insert; 3'RACE nest PCR 產物) 5μl；T-vector 2μl；10×T4 DNA ligase reaction buffer 1μl；T4 DNA ligase (3U/μl) (New England) 1μl；d2H<sub>2</sub>O 1μl。

### b. 轉形作用(transformation)

將經聯結反應之後的表達質體經轉形作用送入 *E. coli JM109* strain 中。首先製備 *JM109* 勝任細胞(competent cells)，挑選單一之菌落在 LB 培養液(Luria-Bertani broth，USB)中進行 37°C 隔夜培養。隔天取 1/100 體積培養之菌液至新鮮的 LB 培養液(1×體積)，於 37°C 培養 3 小時(OD<sub>600</sub>=0.3~0.4)。之後將菌液在 4°C 下離

心(2500 r.p.m.or 1000×g) 10 分鐘，去上清液後加入培養菌液 1/2 體積的 1×TSS wash (by vortex)，接下來在 4°C 下離心(2500 r.p.m. 或 1000×g) 10 分鐘，去除上層液後加 1/10 體積之 Ice-cold TSS 再懸浮(resuspend)，並各取 80µl 分裝於 eppendrof 中保存在- 70 °C~- 80°C 備用。

表達質體(10µl)和 *E. coli JM109* strain 80µl 之 competent cell 混合均勻後，冰浴 30 分鐘，42°C 熱休克(heat shock) 1.5 分鐘，再置於冰上隨後加上 200µl 之 LB 培養液，於 37°C 培養 1 小時。最後加入 IPTG (0.2M) (Sigma) 15µl 及 X-Gal (50mg/ml)(Promega) 15µl 混合均勻，取 100~160µl 塗於含有 Ampicillin (100µg/ml)(Sigma) 之瓊膠培養基(plate)上，在 37°C 培養 12~16 小時。

#### c. PCR check

點選單一的白色菌落(clone)做為模板，總反應體積 15µl 中含有：  
10×Pro Taq DNA polymerase reaction buffer 1.5µl；2.5mM dNTP 0.33µl；Mg<sup>2+</sup>(15mM) 1.5µl；T-vector primer T7 & SP6 (0.5µg/µl) (參見表一)各 0.1µl；Pro Taq DNA polymerase (2U/µl) 0.12µl；d2H<sub>2</sub>O 11.35µl，進行 PCR 反應，其條件為：在高溫 94°C 下反應 1 分鐘；在 52°C 下反應 1 分鐘；在 72°C 下反應 1 分鐘，以上述之條件充分進行 35 cycles；在 72°C 下反應 7 分鐘，最後儲存於低溫 4°C 中。  
挑選放大出原始預測序列片段大小之菌落。

d. 少量質體 DNA 製備與消化作用篩檢(miniprep / digestion check)

以 Nucleo Spin Plus Miniprep Kit (CLONTECH)純化出該 clone 之質體 DNA (plasmid DNA)，將抽出之質體 DNA 2 $\mu$ l；10 $\times$ restriction enzyme (*EcoR* I) buffer 1 $\mu$ l；restriction enzyme(*EcoR* I)(New England) 0.5 $\mu$ l；d2H<sub>2</sub>O 6.5 $\mu$ l，總體積為 10  $\mu$ l 於 37 $^{\circ}$ C 下反應 2 小時。以 1% 電泳瓊膠(agarose-LE)經由電泳分析。

(3)去氧核糖核酸序列分析(NDA sequencing)

首先將初步確定之 DNA，以 ABI PRISM™ Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit(Perkin Elmer Applied Biosystems)來進行定序反應。

a. 定序反應之樣本製備

該方法係以欲分析之 DNA(2 $\mu$ l)為模板，分別用 sequencing primer: T7、SP6(25ng/ $\mu$ l)(參見表一) 2 $\mu$ l；10 $\times$ Big Dye buffer (dNTPs，MgCl<sub>2</sub>) 1.5 $\mu$ l；Big Dye Mixture (含 AmpliTaq™ DNA Polymerase，ddNTPs) 2 $\mu$ l；ddH<sub>2</sub>O 7.5 $\mu$ l，總體積為 15 $\mu$ l 以 PCR 方式，在 96 $^{\circ}$ C 30 秒，50 $^{\circ}$ C 30 秒，60 $^{\circ}$ C 4 分鐘，進行 deoxynucleotide termination extention reaction，反應 30 cycles。ddNTPs 混合物乃是由不同螢光標定之 terminator (Big Dye™ Terminators)，A(adenine)-dye 以 dichloro [R6G]標定綠色表示之；C(cytosine)-dye 以 dichloro [TAMAR]標定藍色表示之；G(guanine)-dye 以 dichloro[R110]標定



黑色表示之；T(thymine)-dye 以 dichloro[ROX]標定紅色表示之。  
定序反應之產物以酒精沉澱之方法純化再以 10 $\mu$ l 的 TSR  
(Template Suppression Reagent )10 $\mu$ l 再懸浮(resuspend)，之後於  
95 $^{\circ}$ C 復性(denatured) 2~3 分鐘後快速置於冰上。

#### b. 分析核酸序列

以 ABI PRISM<sup>®</sup> 310 Genetic Analyzer (Perkin Elmer Applied Biosystems)利用毛細管電泳(capillary electrophoresis)藉自動雷射誘發螢光(automated laser-induced fluorescence DNA sequencing)分析系統來分析 DNA 之序列。

### (三) 建構 H-HSD 與 Z-HSD 全長核酸序列

#### (1) 全長序列聚合酶連鎖反應

##### a. zebrafish (Z-HSD)

以 RACE PCR 放大組裝並經序列分析所得之 Z-HSD 全長序列，設計 N 端(含內切限制酵素位點：*Sal* I)及 C 端(含內切限制酵素位點：*Hind* III)兩段引子(參見表一)，PCR 反應總體積 100 $\mu$ l 中內含 10 $\times$ Ex Taq DNA polymerase reaction buffer 10 $\mu$ l；2.5mM dNTP 混合物 8 $\mu$ l；正反兩股(forward and reverse)oligonucleotide primer (0.5 $\mu$ g/ $\mu$ l) (參見表一)各 1 $\mu$ l；模板(0~7day 斑馬魚胚胎之 dsDNA) 2 $\mu$ l；Ex Taq DNA polymerase (5U/ $\mu$ l) 0.5 $\mu$ l；d<sub>2</sub>H<sub>2</sub>O 77.5 $\mu$ l，進行 PCR 反應，其條件為：在高溫 96 $^{\circ}$ C 下反應 30 秒；在 50 $^{\circ}$ C 下反應

30 秒；在 72°C 下反應 2 分鐘，以上述之條件充分進行 40 cycles；

在 72°C 下反應 15 分鐘，最後儲存於低溫 4°C 中。

PCR 產物大小則由 1% 電泳瓊膠(agarose-LE)經由電泳分析(electrophoresis)，並將其切下以 Gel-M™ Gel Extraction System Kit (VIOGEN)純化。

#### b. Human (H-HSD)

由於 H-HSD 的完整序列(ORF)已然公佈，因此以其發佈之序列為依據設計 N 端(含內切限制酵素位點：*BamH* I)和 C 端(含內切限制酵素位點：*Hind* III)兩段引子(參見表一)，藉由 PCR 放大取得其 cDNA 全長序列。反應總體積 100µl 中包含 10×Ex Taq DNA polymerase reaction buffer 10µl；2.5mM dNTP mixture 8µl；forward and reverse oligonucleotide primer (0.5µg/µl) (參見表一)各 1µl；模板(HepaG2 dsDNA) 2µl；Ex Taq DNA polymerase (5U/µl) 0.5µl；d2H<sub>2</sub>O 77.5µl，進行 PCR 反應，其條件為：在高溫 96°C 下反應 30 秒；在 50°C 下反應 30 秒；在 72°C 下反應 2 分鐘，以上述之條件充分進行 40 cycles；在 72°C 下反應 15 分鐘，最後儲存於低溫 4°C 中。

PCR 產物大小則由 1% 電泳瓊膠(agarose-LE)經由電泳分析(electrophoresis)，並將其切下以 Gel-M™ Gel Extraction System Kit (VIOGEN)純化。

## (2) 次選殖(subcloning)

### a. ligation

分別取 Z-HSD 及 H-HSD 純化之 PCR 放大產物 5 $\mu$ l 為 insert，pGEM<sup>®</sup>-T Easy Vector 1 $\mu$ l，10 $\times$ T4 DNA ligase reaction buffer 2 $\mu$ l；T4 DNA ligase(3U/ $\mu$ l) 1 $\mu$ l；d2H<sub>2</sub>O 11 $\mu$ l，其反應總體積為 20 $\mu$ l，在 16 $^{\circ}$ C 下進行聯結反應至少 4 小時至 12 小時。

### b. transformation

將經聯結反應之後的表達質體經轉形作用送入 *E. coli JM109* strain 中。表達質體(20 $\mu$ l)和 *E. coli JM109* strain 80 $\mu$ l 之 competent cell 混合均勻後，冰浴 30 分鐘，42 $^{\circ}$ C 熱休克 1.5 分鐘，再置於冰上隨後加上 200 $\mu$ l 之 LB 培養液，於 37 $^{\circ}$ C 培養 1 小時。最後加入 IPTG (0.2M) 15 $\mu$ l 及 X-Gal (50mg/ml) 15 $\mu$ l 混合均勻，取 100 ~ 160 $\mu$ l 塗於含有 Ampicillin (100 $\mu$ g/ml)之 plate 上，在 37 $^{\circ}$ C 培養 12~16 小時。

### c. PCR check

各別點選 Z-HSD 與 H-HSD 之單一白色菌落(clone)做為模板，總反應體積 15 $\mu$ l 同前述之作法，挑選放大出原始全長序列片段大小之菌落。

### d. miniprep/digestion check

以 Nucleo Spin Plus Miniprep Kit (CLONTECH)純化出該 clone 之 plasmid DNA，將抽出之質體 DNA 以前述之方法進行 digestion

check，最後以 1% 電泳瓊膠經由電泳分析。

### (3) 去氧核糖核酸序列分析(NDA sequencing)

將初步確定之質體 DNA，以 ABI PRISM™ Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit(Perkin Elmer Applied Biosystems)來進行定序反應。以 ABI PRISM® 310 Genetic Analyzer (Perkin Elmer Applied Biosystems)分析系統來分析 DNA 之序列。

## (四) 建構蛋白質表達質體

### (1) 消化作用及聯結反應(digestion / ligation )

經由核酸定序分析確認的質體 DNA，分別就預先設計之內切限制酵素位點，Z-HSD 以 *Sal* I 及 *Hind* III 二個內切限制酵素進行消化作用；H-HSD 則以 *Bam*H I 及 *Hind* III 二個內切限制酵素進行消化作用，於 37°C 至少 4 小時以上，以作為 insert。然後將 QIAexpress system(QIAGEN)之 pQE-31 和 pQE-9 之表達質體亦分別用 *Sal* I / *Hind* III (pQE-31)及 *Bam*H I / *Hind* III (pQE-9)進行消化作用，作為載體(vector)。經過消化作用後 Z-HSD 和 H-HSD 之 insert 與載體，各別利用 T4 DNA ligase 在 16°C 下至少培養 4 小時進行聯結反應，最後接合成 pQE-31(zebrafish)和 pQE-9(human)蛋白質表達質體。

### (2) 轉形作用(transformation)

經聯結作用後之蛋白質表達質體 pQE-31(zebrafish)、pQE-9 (human)以  $\text{CaCl}_2$  transformation 方法送入 *E. coli JM109* strain (不具任何抗藥性)中。首先製備 *E. coli JM109* strain 之 competent cell，挑選單一 *JM109* 菌落，在 2 ml LB 培養液中進行 37°C 隔夜培養(約 15 小時)，接著將菌液轉至 50 ml 新鮮的 LB 培養液 37°C 培養約 2.5 小時( $\text{OD}_{600\text{ nm}}=0.7$ )，將菌液於 4°C 之下，用 1500g，離心 8 分鐘後，去除上清液再加入已滅菌之 25 ml  $\text{CaCl}_2$  (100 mM )，懸浮均勻後冰浴 30 分鐘；之後，重複離心步驟(4°C、1500g、8 分鐘)，再加入 1 ml  $\text{CaCl}_2$  (100 mM )，輕輕的懸浮後再冰浴 30 分鐘，然後分別加入上述經聯結之蛋白質表達質體充分均勻混合後，在冰浴中靜置 30 分鐘，以 42°C 熱休克 3 分鐘，快速置於冰中 5 分鐘，採用熱漲冷縮原理先利用  $\text{CaCl}_2$  在細胞膜表面上打洞讓表達質體進入細胞中，之後快速冷卻使細胞膜上洞口緊縮，並將經轉形作用之 *E. coli JM109* 塗在含有抗生素 Ampicillin (100  $\mu\text{g/ml}$ )的 LB plate 中，在 37°C 恆溫隔夜培養以進行微篩選。

### (3)微篩選(mini screening)

經隔夜培養的 LB plate 篩選存活的單一菌落，加入 2ml LB 培養液(內含抗生素 ampicillin)中 37°C 恆溫隔夜培養。抽出隔夜菌液

50  $\mu$ l，同樣加入 phenol : chloroform = 1 : 1 之混合液 50  $\mu$ l，接著震盪破細胞。首先先震盪 45 秒，靜置 1 分鐘，再震盪 45 秒，然後離心 5 分鐘，取 15  $\mu$ l 上清液以瓊膠(1% agarose gel)進行電泳分析。分別利用 pQE-9 或 pQE-31 質體做為標記物(marker)進行微篩選。以 QIAprep spin miniprep kit(QIAGEN)純化聯結成功的突變蛋白質表達質體，再進一步以先前所設計的二種限制內切酶進行消化作用(digestion)來確認聯結結果是否為蛋白質表達質體當嵌入物(insert)，與 pQE-9(or pQE-31)質體當載體。

#### (4) 去氧核糖核酸序列分析(DNA sequencing)

將上述構建成功內含表達質體之 H-HSD 與 Z-HSD 的 DNA，利用 ABI PRISM™ 377 DNA Sequencer 及 ABI PRISM™ dye terminator cycle sequencing ready reaction kit(Perkin Elmer Applied Biosystems)進行定序分析作進一步的確認。

#### (五) 蛋白質大量表達、純化與定量分析

以下方法參照 J.Sambrook 等人在” Molecular Cloning: a Laboratory Manual ”中所描述之方法進行。

##### (1) 蛋白質質體大量表達(overexpression)與純化(purification)

將經 DNA 序列分析確認，正確序列的新穎酵素之表達質體，利用 CaCl<sub>2</sub> transformation 的方法送入含有 pDM1.1 質體的 *E. coli* M15 strain(具有 kanamycin 抗藥性)宿主細胞中表達蛋白質。製備

*E. coli* M15 strain 之 competent cell 條件方法如上述轉形作用 (transformation)。

將轉形至 *E. coli* M15 strain 中具有表達質體的單一菌株培養在含有 ampicillin(100 µg/ml)及 kanamycin(50 µg/ml)(Sigma)的 2 ml LB 培養液中 37°C 隔夜培養。將隔夜菌液取 1/100 體積接種到新鮮的 LB 培養液(1×體積)37°C 培養 2~3 小時( $OD_{600\text{ nm}}=0.3\sim0.4$ )，接下來將菌液加到 200ml 之 LB 培養液於 37 °C 培養 3~4 小時( $OD_{600\text{ nm}}=0.6\sim0.7$ )，之後加入 IPTG(200 µg/ml)後，轉移到 30 °C 下繼續培養 3~4 小時。因在 *E. coli* M15 strain 宿主細胞中含有攜帶 *lacI<sup>q</sup>* gene 之質體 pDM1.1(具有 kanamycin 抗藥性)會產生 *lac* 抑制子 (repressor)，與質體 pQE-9(or pQE-31)上之 phage T5 啟動子 (promoter)結合，進而抑制 pQE-9(or pQE-31)質體蛋白質的表達；為使表達質體順利啟動蛋白質大量表達，故加入 IPTG (isopropyl  $\beta$ -D-thiogalactopyranoside)與 *lac* repressor 結合，使 *lac* repressor 離開 phage T5 promoter 後，RNA 聚合酵素得以順利結合到 phage T5 promoter 上，表現基因得以順利進行轉錄作用(transcription)合成 mRNA，再經 mRNA 轉譯作用(translation)合成蛋白質。之後將菌液離心(8,000×g，10 min，4°C)後收集沈澱細胞(cell pellets)，將其懸浮於冰的 PE 緩衝溶液(Dialysis buffer：50 mM

NaH<sub>2</sub>PO<sub>4</sub>，0.25 mM EDTA，pH =7.4)中，最後終體積 20 ml；接著以 French<sup>®</sup> Pressure Cells and Press 在 4°C 下高壓破碎細胞(1000 psi；三~四次)；經超高速離心(4°C，50,000g，15 min)後收集上清液，注入預先以 PEN buffer(Enzyme extract buffer：50 mM NaH<sub>2</sub>PO<sub>4</sub>，0.25 mM EDTA，50 mM NaCl，pH =7.4)平衡之 Ni-NT (nickel- nitrilotriacetic-agarose)管柱，再以二倍管柱體積之 PEN buffer 清洗管柱，去除非專一性(non-specific)結合蛋白質。因在 pQE-9 載體之 N 端接有六個組胺酸(histidine)密碼子，所以表達的蛋白質 N 端的六個 His-peptide 可以與 Ni 進行親和性的結合，而使重組的蛋白質結合至 Ni -NTA 樹脂之管柱，再利用與組胺酸構造相似之咪唑(imidazole)(Sigma)濃度梯度(5-100 mM)與蛋白質競爭樹脂上的結合部位，將蛋白質溶析出來(參見附錄 7)。將收集純化的蛋白質在透析液(Dialysis buffer：50 mM NaH<sub>2</sub>PO<sub>4</sub>，0.25 mM EDTA) 4°C 下進行隔夜透析，藉以去除大部分的咪唑。

## (2)蛋白質濃度(protein concentration)測定

將純化出來的蛋白質採用 Bradford 之方法測其蛋白質濃度，該方法利用 BSA(bovine serum albumin)作為定量的標準溶液，設定五組 BSA 標準濃度(standard concentration)，其濃度分別為 0 µg/ml、5 µg/ml、10 µg/ml、15 µg/ml 及 20 µg/ml，以 PE 緩衝溶液定量體積至 800 µl，各標準溶液分別再加入 200 µl 的 Bradford 緩衝溶液，使最終體積為 1000 µl 後，在室溫下進行反應 5 分鐘



後置入透光石英比色管中，以可見光/紫外光光譜分析儀(Hitachi U-3000)波長 595nm 下偵測其吸光值，建立標準 BSA 濃度曲線。測試樣品之濃度經範圍內標準濃度曲線內差法計算後即可得之。

### (3)蛋白質純度測定(SDS-PAGE)

SDS-PAGE 之濃度為 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis，其製備方法如下：首先將 4 ml 下層膠(resolving gel)：acrylamide-bisacrylamide (29 : 1) 1.66 ml；1.5 M Tris-HCl (pH=8.8) 1 ml；10 % SDS 0.04 ml；1.5 % Ammonium persulfate 0.5 ml；dH<sub>2</sub>O 0.79 ml；TEMED 2  $\mu$ l，進行聚合(polymer)凝膠反應；接著將 2 ml 上層膠(stackig gel)：acrylamide- bisacrylamide (29 : 1) 0.25 ml；0.5 M Tris-HCl (pH=6.8) 0.5 ml；10 % SDS 0.02 ml；1.5 %Ammonium persulfate 0.1 ml；dH<sub>2</sub>O 1.03 ml；TEMED 2  $\mu$ l，同樣著進行聚合 (polymer)凝膠反應，之後再將 prestained SDS-PAGE Standards protein marker/Low range(分子量分別為：113kDa ---Phosphorylase B；92kDa---Bovine serum albumin；52.9kDa---Ovalbumin；35.4kDa---Carbonic anhydrase；29kDa---Soybean trypsin inhibitor；21.5kDa---Lysozyme，BIO-RAD)、純化之蛋白質(10  $\mu$ g)分別與 2X sample buffer[總體積 10 ml：0.5 M Tris-HCl (pH=6.8) 1 ml；2 % SDS 2 ml；1.25 M DTT 0.5 ml(1/20)；10 % glycerol 1ml；0.4 % bromophenol blue 0.5

ml(0.02 %) ; dH<sub>2</sub>O 定量至 10 ml]以 1 比 1 的比例混合均勻後，在沸水中煮沸 5 分鐘使蛋白質變性(denature)，進行 SDS-PAGE 電泳。其 SDS-PAGE electrophoresis buffer : Tris base 3 g ; SDS 1 g ; Glycine 14.4 g ; 加 dH<sub>2</sub>O to 1000 ml ; 電泳上層膠之電壓為 100 V，當 protein band 至下層膠電壓改為 120 V，利用蛋白質本為負電荷原理，讓電泳由負電極帶向正電極，約電泳 2 小時。之後，SDS gel 以 coomassie blue staining solution(0.1 % coomassie blue R250 208 ml ; methanol 208 ml ; glacial acetic acid 84 ml)隔夜染色，再經 destain solution(acetic acid 50 ml ; methanol 150 ml ; dH<sub>2</sub>O 300 ml)退染後，觀察標的蛋白質單體(monodimer)分子量與蛋白質的純度。

#### (4) 酵素動力學分析(Enzyme kinetic assay)

酵素反應是以 NAD<sup>+</sup> (或 NADP<sup>+</sup>)、4-Androstene-3,17-dione (Aldrich) 、 5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol 、 Estriol 、 16 $\alpha$ -Hydroxyestrone、Estrone、 $\beta$ -Estradiol 與 Hydrocortisone (Sigma) 為受質，利用 NADH 在波長 340 nm 有最大吸光之特性，進行酵素動力學分析，監測產物 NADH(or NADPH)的生成速率，來測量酵素活性，觀察酵素反應機制、酵素對受質之親和力及代謝迴轉率。單位酵素活性(U)之定義是指每毫克蛋白質在每分鐘所產生(或消失)的產物(ex: NADH)含量( $\mu$ mole)。首先將反應中一個 NAD<sup>+</sup>(or NADP<sup>+</sup>)濃度固定，則改變另一受質濃度進行反應，即可得一組 5 個反應速率之數據，再利用五種 NAD<sup>+</sup>濃度重複上述

步驟，結果可得五組(5×5=25 個)數據。在每個反應濃度下分別求得酵素反應初速度，以 OD/min 之單位代入公式計算，將所得數據以 Erithacus Software 公司所設計的 GraFit(Version 3.0)軟體進行雙受質酵素分析。推測此新穎酵素之反應有兩個受質，故其反應機制有兩種可能性：

1、乒乓機制(Ping-Pong mechanism)：

$$V=V_{\max}[A][B]/(K_mB[A]+K_mA[B]+[A][B])$$

2、序列機制(Sequential mechanism)：

$$V=V_{\max}[A][B]/(K_{ia}K_mB+K_mB[A]+K_mA[B]+[A][B])$$

經由電腦校正後，即可得酵素反應之初速度( $V_{\max}$ ；是以 OD/min 值表示)，及  $K_m$ (單位為 mM)，此代表酵素與其受質之結合能力。以初速度( $V_{\max}$ )經換算可得到  $K_{cat}$ ，即酵素反應之代謝迴轉速率。其換算公式為：

$$V_{\max} = OD/6.22 \times 10^{-3} / \text{min/mg/ml} = \mu\text{mole /min/mg}$$

$k_{cat}$  即  $V_{\max} / \text{sec/mg} = V_{\max} / 60 \times 100000^{-5}$  (單位是 1/sec)。比較蛋白質之酵素反應的各個  $K_m$  及  $K_{cat}$  後，由此初步確定蛋白質在酵素反應中扮演之角色。

## 參、結果

### 一 生物資訊資料庫(NCBI)搜尋

本研究以新穎之斑馬魚去氫酶(zebrafish novel dehydrogenase)為關鍵字，詢問 NCBI 之核酸序列資料庫，進行檢索。最終選定一片段序列(NCBI accession number AI477544)，該序列之 5'端核酸序列相似於 WP:C45B11.3 CE05427Alcohol dehydrogenase (參見附錄 1)。運用 NCBI 提供之 BLAST 服務，搜尋比對到一個與此斑馬魚片段序列相互對應之人類的相似序列(NCBI accession number NP\_115679)，該人類之 hypothetical protein 已具有完整之蛋白質編碼區，但其實際之功能仍未被確認(參見附錄 2)。此人類之蛋白質與斑馬魚之序列相似性(identity)為 71% (參見附錄 8)，由於斑馬魚缺乏 3'端編碼區序列，不論於 EST 或 Zebrafish genome 資料庫皆無法延長其 3' 端序列，因此預定藉由 3'RACE PCR 之方法取得其完整之蛋白質編碼區之 cDNA 序列。將此段放大之 cDNA 片段與原 NCBI 上發佈之參考序列進行比對組裝，依據序列重疊部分將此片段序列延伸，並找出可能之終止密碼子，由 NCBI 上之斑馬魚及人類參考序列與此 3' RACE PCR 放大之序列進行比對(參見圖二)。以人類之起始密碼子及終止密碼子為參考依據，推測出組裝延長後之斑馬魚全長序列之可能起始密碼子及終止密碼子，並以之設計放大全長 cDNA 之引子(參見表一)。

經全長 PCR 放大且定序確認無誤之斑馬魚及人類的序列，分別將之胺基酸序列進行排比，可發現其 N 端和 C 端之序列有高度相似性(參見圖三)。以 NCBI 之下 BLAST 的保留結合結構區(conserved domain)搜尋比對功能，詢問斑馬魚及人類之胺基酸構成酵素活性之演化保留區域為何？結果顯示，兩者皆具有相同之二個結構區，分別為 N 端之 adh\_short 與 C 端之 SCP2(參見圖四)。adh\_short domain (short-chain dehydrogenases / reductases family) 已知是  $\text{NAD}^+$  或  $\text{NADP}^+$  依賴的氧化還原酵素之結合結構區(Benyajati C *et al.*,1981)，SCP2 (sterol carrier protein 2) 涉及固醇類的結合與細胞內的膽固醇(cholesterol)及許多脂質(lipids)的運輸(Szyperski T *et al.*,1993)。將人類及斑馬魚之胺基酸序列，分別和已知的酒精去氫酵素與固醇類攜帶蛋白家族進行多重比對，可觀察出不論人類或斑馬魚，和七種人類已知的酒精去氫酵素之蛋白質序列僅有少部份相似，然而與固醇類攜帶蛋白家族之蛋白質序列 C 端部分之相似性較高(參見圖五、六)。接著在 NCBI 之 Uni-Gene 資料庫中探問該人類(NP\_115679/NM\_032303)基因之相關資訊，查尋結果發現此新穎人類蛋白分別與人類(*H.sapiens*)、大鼠(*R.norvegicus*)、小鼠(*M.musculus*)和線蟲(*C.elegans*)等具有同源相似(參見圖七與附錄 6)，其中最相似於人類之 estradiol 17-beta dehydrogenase [EC:1.1.1.62] type 4。由以上之結果顯示此二個新穎之

斑馬魚及其對應之人類酵素的功能，可能是參與固醇類有關物質的氧化還原作用之酵素(NAD<sup>+</sup>或 NADP<sup>+</sup>依賴性)。人類之 estradiol 17-beta dehydrogenase [EC:1.1.1.62] type 4，用其查尋 KEGG 網站之資料庫，詢問之結果顯示，此類酵素又稱之為 17-beta hydroxysteriod dehydrogenase，屬於 androgen 和 estrogen 之代謝路徑中的作用酵素(參見附錄 9)。據此酵素反應分析進行時，可能之反應受質係以此為依據，推測及篩選構造相似之物質作為酵素反應之受質，目前可購得之固醇類化合物有七種，分別為 4-Androstene-3,17-dione、5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol、Estriol、16 $\alpha$ - Hydroxyestrone、Estrone、 $\beta$ -Estradiol 與 Hydrocortisone (參見圖八)。

## 二 斑馬魚(Z-HSD)全長序列之選殖與定序

利用 3'RACE PCR 之放大作用，延長斑馬魚片段序列所缺少之 C 端編碼區域，結果找到一個 clone 內含約 872 個核酸序列以及 Poly A Tail，經 DNA 序列分析結果(參見圖九、十)經初步組裝，以此假定斑馬魚全長序列為藍本(參見圖二)，用以設計放大全長 cDNA 之正反二股引子，該引子 5'端及 3'端分別另行加入設定之內切酵素位點 *Sal* I / *Hind* III(參見表一)。以斑馬魚 0~7 天魚胚胎 cDNA 為模板，PCR 放大約 1.3 kpb 之產物，經由聯結作用將此放大合成之 cDNA 片段接合 pGEM<sup>®</sup>-T Easy Vector，以轉形作用將此質體送入 *E. coli*. JM109 strain

進行次選殖，篩選正確之 clones，並於小量質體 DNA 製備後，進行核酸序列分析，實驗流程參見(圖一)。經 DNA 序列分析結果，其序列分析圖(參見圖十一、十二、十三)，此新穎斑馬魚之全長序列確認之後，其開放閱讀框架(open reading frame)為 1469 個核酸，序列之 72~1319 為蛋白質編碼區(coding region)，推定出 416 個胺基酸(參見圖十四)。

### 三 人類(H-HSD)全長序列之選殖與定序

依據 NCBI 上所提供人類之參考序列設計放大全長 cDNA 之正反二股引子，該引子 5'端及 3'端分別另行加入設定之內切酵素位點 *Bam*HI / *Hind* III(參見表一)。以 HepaG2 之 cDNA 為模板，PCR 放大取得人類全長 cDNA，同樣經由聯結作用將此放大合成之 cDNA 片段接合 pGEM<sup>®</sup>-T Easy Vector 以轉形作用將此質體送入 *E. coli*. JM109 strain 進行次選殖，篩選正確之 clones，並於小量質體 DNA 製備後，進行核酸序列分析，經 DNA 序列分析結果，其序列分析圖(參見圖十五、十六)，此新穎人類之全長序列確認之後，得其序列為 1035 個核酸，與 NCBI 上所提供人類之參考序列相同。開放閱讀框架(open reading frame)為 3303 個核酸，序列之 130~1164 為蛋白質編碼區(coding region)，345 個胺基酸(參見圖十七)。

### 四 人類與斑馬魚之(HSD)蛋白質表達質體之建構

之後就預先設計之內切限制酵素位點，分別將人類之質體 DNA 和 pQE9 表達質體以 *BamH* I / *Hind* III 作用；斑馬魚之質體 DNA 和 pQE31 表達質體以 *Sal* I / *Hind* III 作用後，經由聯結作用進行表達載體與 H-HSD 及 Z-HSD 的質體 DNA 之接合作用，然後轉形至 *E coli*. JM109 strain，接續以微篩選及小量質體 DNA 製備方式，取得正確之質體 DNA，用以進行蛋白質表達，實驗流程(參見圖一)。

## 五 新穎蛋白質之表達與純化

將上述內含 pQE 表現載體系統之質體 DNA(人類及斑馬魚)，轉形至 *E coli*. M15 strain(內含 pDM1.1 質體)。分別在室溫、30°C 和 37°C 下以 IPTG 誘發其過量表現蛋白質，結果發現皆可表現蛋白質，但其量隨溫度不同逐漸增加，以 SDS-PAGE 來觀察分析(參見圖十八)。此人類與斑馬魚之蛋白質表達質體，經 IPTG 誘發蛋白質大量表達三小時，破碎細胞及超高速離心後，取得細胞粗萃物上清液，經 Ni-NTA 樹脂之管柱層析(nickel column chromatography)，利用咪唑(imidazole)與組胺酸競爭結合作用，以不同之濃度梯度將標的蛋白溶析出來，藉由 SDS-PAGE 進行純度分析，其膠體經 coomassie blue 染色後，可觀察出純化的斑馬魚和人類之標的蛋白質，結果其蛋白質分子量和利用核酸序列所預測之蛋白質分子量大小相符，斑馬魚之蛋白質分子量約為 45.546kDa；人類之蛋白質分子量約為 37.320kDa(參見圖十九、二



十)。人類與斑馬魚之蛋白質純化表現分析，結果顯示斑馬魚對 Estriol 之 recovery activity 為 42.9%，對 5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol 之 recovery activity 為 95.58%；人類 Estriol 之 recovery activity 為 22.38%，對 5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol 之 recovery activity 為 16.99%(參見表二、三)。

## 六 酵素反應分析

根據前述之搜尋資料來推測，此酵素可能為雙受質反應，因此酵素反應分析(enzyme kinetic assay)是以 NAD<sup>+</sup>(或 NADP<sup>+</sup>)為反應之第一受質，另一受質則以 Estradiol 17-beta dehydrogenase 反應參與之反應受質及其構造相似物質：4-Androstene-3,17-dione、Estriol、Estrone、5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol、16 $\alpha$ -Hydroxyestrone 與 Hydrocortisone 等為受質，利用波長 340 nm 之吸光下測定產物 NADH(或 NADPH)的生成速率。單位酵素活性(U)之定義是每毫克蛋白質在每分鐘所產生的 NADH (或 NADPH)的微莫耳數的量。酵素反應是在 pH=7.4 的磷酸鹽緩衝溶液(50mM)中進行，結果僅有 Estriol 和 5 - Androstene- 3 $\beta$ , 16 $\alpha$ , 17 $\beta$ -triol 在波長 340 nm 吸光下可觀察到 NADH(或 NADPH)的生成，此酵素則無法對其他的構造相似物質產生氧化作用。經由波長 340 nm 吸光下測得之吸光值，可計算出斑馬魚與人類此新穎之 17-beta 羥基固醇去氫酶(17-beta hydroxysteroid dehydrogenase )分別對

Estriol 和 5- Androstene -3 $\beta$ , 16 $\alpha$ , 17 $\beta$ -triol 這二種受質之比活性 (specific activity)。斑馬魚對 Estriol/NAD<sup>+</sup> 之 specific activity 為 3.16(unit/mg); 對於 Estriol/NADP<sup>+</sup> 之 specific activity 為 3.71(unit/mg); 對於 5-Androstene/NAD<sup>+</sup> 之 specific activity 為 1.75(unit/mg); 對於 5-Androstene/NADP<sup>+</sup> 之 specific activity 為 1.41(unit/mg)。人類對於 Estriol/NAD<sup>+</sup> 之 specific activity 為 10.74(unit/mg); 對於 Estriol/NADP<sup>+</sup> 之 specific activity 為 10.73(unit/mg); 對於 5-Androstene/NAD<sup>+</sup> 之 specific activity 為 12.81(unit/mg); 對於 5-Androstene/NADP<sup>+</sup> 之 specific activity 為 13.62(unit/mg)(參見表二、三)。

## 肆、討論

在本實驗中，藉由生物資訊資料的大量生物資料，協助我們可快速與多方面地取得相關的訊息，並且能針對特定功能或作用之新穎蛋白進行研究。此次實驗藉由 NCBI 搜尋到之斑馬魚片段序列 (AI477544)，以 RACE PCR 的方式來建構其完整之核酸序列，現今由於斑馬魚 genome 資料庫和其他物種之 genome 資料庫已陸續建立完成，則可直接由這些基因資料庫進行搜尋比對，將其可能之完整核酸序列尋找出來，可省略繁複地實驗步驟，並且利用電腦及網路之強大功能，亦可快速且便宜地完成研究之部分工作。而且許多不同生物資訊資料庫的多方連結，相互交換資料庫中之資訊，讓我們能夠不必逐一地到不同的網站去搜尋不同的資料。簡言之，有了生物資訊資料庫這項有力的工具，使得生命科學的研究歷程更事半功倍。

此次由 NCBI 上所搜尋到之人類和斑馬魚新穎 17-beta 羥基固醇去氫酶(17-beta hydroxysteroid dehydrogenase, 17 $\beta$ -HSD)，在比對的過程中，如何去篩選與確認其相互對應之關係，乃是此研究主題之一大課題。在比對時需就研究的目的來選擇不同的資料庫(database)種類，藉以充分發揮其輔助的功用。因此當判斷其是否為一對一之對應關係，則需仔細審慎的依據其所提供之資料(例如:功能定義、來源、特殊基元和相似性等)，並以多重比對方法來快速篩選與排除非相關

之對像，這些比對過程可能會影響之後研究方向的判斷。因此我們利用了 Uni-Gene、PuMed、Domain 和 Protein/Nucleotide 等資料庫以及一些應用軟體來作為參考的依據。

於斑馬魚和人類之新穎 HSD 的轉殖時，發現當 H-HSD 轉接到 pQE30 之表達質體時，無法大量表現蛋白質產生(資料未附)，但轉接至 pQE9 的表達質體時，則能夠被誘發其大量表現。然而經由破碎細胞及高速離心，直至透過 Ni-NTA 管柱層析純化 H-HSD 之蛋白，相同的條件之下，與 Z-HSD 比較後可發現，人類的 HSD 純化出之蛋白質質量明顯比斑馬魚之 HSD 減少，一公升之培養菌液僅能純化出約 0.12 毫克之蛋白質(參見表二、三)。為了確認可能之因素，再次將純化過程各階段之 fraction(supernatant、pellet and flow through)以 12.5%之 SDS-PAGE 分析(資料未附)，發現人類之 HSD 有極大部分之蛋白質產生於 insoluble fraction，因此相對來說，其純化之蛋白質則大大地減少。有鑑於此，考慮再利用不同的表現載體(例如：Glutathione S-transferase gene fusion system)或以其他方式(例如：真核細胞)來表達與純化此新穎之人類 HSD 蛋白質。目前將此人類之新穎的 HSD 質體 DNA 轉接到 GST-vector 以表達蛋白質之工作已在進行之中。

酵素反應之分析，選擇了七種與羥基固醇去氫酵素反應有關之化合物(參見圖八)，其中只有 Estriol 和 5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol 可

被此斑馬魚和人類之 HSD 所氧化，因此我們比較這七種受質之結構，可以歸納出此酵素對同時在 C3、C16 和 C17 的位置上有 hydroxyl group 之固醇類能推動其酵素反應，而單獨或只在此三個位置的其中二個上有 hydroxyl group，特別是在 C3/C16 和 C3/C17 之位置，則無法推動反應。另一個觀察是當其中二個位置上(C3/C17: 4-Androstene-3,17-dione)為 keto group，亦無反應產生，所以我們認為此酵素可能針對 C16 及 C17 上之 hydroxyl group 進行氧化作用。然而目前我們所購得的受質種類有限，而且亦無對產物進行分析，所以還不足以下定論此酵素會對那一類之受質的酵素反應最佳，此後還需找更多此固醇類相似構造的受質來作進一步的測試。

進行酵素動力學之分析時，我們所遭遇一主要的問題乃是，固醇類受質溶解度不佳，導致其加入酵素反應的 solution 時會再析出，尤其當受質濃度提高時，使得整個反應的 solution 呈混濁狀，其波長 340nm 之吸光則受影響，最終酵素動力學之分析的數值無法正確計算出其酵素的  $V_{max}$ 、 $K_m$  與  $K_{cat}$  值，因以亦無法推論其酵素的反應機制。因此我們所要解決的問題是，固醇類化合物如何提高溶解度的方法，在此之前曾經以 DMSO (Dimethyl Sulfoxide)、EtOH (ethanol)、PG (Propylene Glycol) 來作為受質之溶劑，其中 DMSO 和 PG 可提高受質之溶解度，但當受質濃度越高，則其產生混濁況狀越嚴重，而 EtOH

之溶解度較差，但其造成混濁況狀較輕微，因此無法提高受質的使用濃度。之後嘗試加入 Tris、SDS、Tritone-100 (1~0.1%)等來改善混濁況狀，使用之後雖然可有不同程度的改善，可是卻會使酵素失去活性。

綜合上述的情形，此斑馬魚和人類之新穎 17-beta 羥基固醇去氫酶(17 $\beta$ -HSD)，除了再找尋其他可能之受質之外，還要再解決溶解度之問題。受質部分或許可以合成的方式來設定特定位置的官能基，亦可思考將受質合成為水溶性的形式，來解決溶解度之問題。所以目前僅將斑馬魚和人類之新穎 17-beta 羥基固醇去氫酶(17 $\beta$ -HSD)表現及純化出來，其酵素反應對 Estriol 和 5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol 之比活性：斑馬魚對 Estriol / NAD<sup>+</sup>之 specific activity 為 3.16(unit/mg)；對於 Estriol/NADP<sup>+</sup>之 specific activity 為 3.71(unit/mg)；對於 5-Androstene /NAD<sup>+</sup>之 specific activity 為 1.75(unit/mg)；對於 5-Androstene/NADP<sup>+</sup>之 specific activity 為 1.41(unit/mg)。人類對於 Estriol/NAD<sup>+</sup>之 specific activity 為 10.74(unit/mg)；對於 Estriol/NADP<sup>+</sup>之 specific activity 為 10.73(unit/mg)；對於 5-Androstene/NAD<sup>+</sup>之 specific activity 為 12.81 (unit/mg)；對於 5-Androstene/NADP<sup>+</sup>之 specific activity 為 13.62 (unit/mg)(參見表二、三)，但其酵素動力學分析還未完成，還待再做分析，找出最適合此新穎酵素之反應受質，以探究其參與之催化反應和此酵素之重要性。

## 伍、參考文獻

**Arnold, FH.** (1991). Metal-affinity separations: a new dimension in protein processing. *Bio/Technology*. 9:151-156.

**Benyajati C, Place AR, Powers DA, Sofer W.** (1981). Alcohol dehydrogenase gene of *Drosophila melanogaster*: relationship of intervening sequences to functional domains in the protein. *Proc Natl Acad Sci U S A*. May; 78:2717-2721.

**Blomquist CH, Bonenfant M, McGinley DM, Posalaky Z, Lakatua DJ, Tuli-Puri S, Bealka DG, Tremblay Y.** (2002). Androgenic and estrogenic 17beta- hydroxysteroid dehydrogenase/17-ketosteroid reductase in human ovarian epithelial tumors: evidence for the type 1, 2 and 5 isoforms. *J Steroid Biochem Mol Biol*. Aug; 81:343-351.

**Blomquist CH.** (1995). Kinetic analysis of enzymic activities: prediction of multiple forms of 17beta- hydroxysteroid dehydrogenase. *J Steroid Biochem Mol Biol*. Dec; 55:515-524.

**Brown WM, Metzger LE, Barlow JP, Hunsaker LA, Deck LM, Royer RE, Vander Jagt DL.** (2003). 17-beta-Hydroxysteroid dehydrogenase type1: computational design of active site inhibitors targeted to the Rossmann fold. *Chem Biol Interact*. Feb 1; 143-144:481-491.

**Chai Z, Brereton P, Suzuki T, Sasano H, Obeyesekere V, Escher G, Saffery R, Fuller P, Enriquez C, Krozowski Z.** (2003). 17 beta-hydroxysteroid dehydrogenase type XI localizes to human steroidogenic cells. *Endocrinology*. May; 144:2084-2091.

**Dasmahapatra AK, Doucet HL, Bhattacharyya C, Carvan MJ 3rd.**

(2001). Developmental expression of alcohol dehydrogenase(ADH3) in zebrafish (*Danio rerio*). *Biochem Biophys Res Commun.* Sep 7; 286:1082-1086.

**Gangloff A, Shi R, Nahoum V, Lin SX.** (2003). Pseudo-symmetry of C19 steroids, alternative binding orientations, and multispecificity in human estrogenic 17beta-hydroxysteroid dehydrogenase. *FASEB J.* Feb; 17:274-276.

**Gangloff A, Garneau A, Huang YW, Yang F, Lin SX.** (2001). Human oestrogenic 17beta- hydroxysteroid dehydrogenase specificity: enzyme regulation through an NADPH-dependent substrate inhibition towards the highly specific oestrone reduction. *Biochem J.* May 15; 356: 269-276.

**Gould-Fogerite S, Mazurkiewicz JE, Raska K Jr, Voelkerding K, Lehman JM, Mannino RJ.** (1989). Chimerasome-mediated gene transfer in vitro and in vivo. *Gene.* Dec 14; 84:429-438.

**Graham FL, van der Eb, AJ.** (1973). Transformation of rat cells by DNA of human adenovirus 5. *Virology.* 54:536-539.

**Graham FL, van der Eb, AJ.** (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology.* 52:456-467.

**Gynthia Gibos, Per Jambeck.** (2002). Developing Bioinformatics Computer Skills.

**He XY, Merz G, Yang YZ, Mehta P, Schulz H, Yang SY.** (2001). Characterization and localization of human type10 17beta-hydroxysteroid dehydrogenase. *Eur J Biochem.* Dec; 268:6637.



**He XY, Schulz H, Yang SY.** (1998). A human brain L-3-hydroxyacyl-coenzyme A dehydrogenase is identical to an amyloid beta-peptide-binding protein involved in Alzheimer's disease. *J Biol Chem.* Apr 24; 273:10741-10746.

**Klein TM, Arentzen R, Lewis PA, Fitzpatrick-McElligott S.** (1992). Transformation of microbes, plants and animals by particle bombardment. *Bio/Technology.* 10:286-291.

**Knutson JC, Yee D.** (1987). Electroporation: parameters affecting transfer of DNA into mammalian cells. *Anal Biochem.* Jul; 164:44-52.

**Koh E, Noda T, Kanaya J, Namiki M.** (2002). Differential expression of 17beta-hydroxysteroid dehydrogenase isozyme genes in prostate cancer and noncancer tissues. *Prostate.* Oct 1; 53:154-159.

**Kristan K, Rizner TL, Stojan J, Gerber JK, Kremmer E, Adamski J.** (2003). Significance of individual amino acid residues for coenzyme and substrate specificity of 17beta-hydroxysteroid dehydrogenase from the fungus *Cochliobolus lunatus*. *Chem Biol Interact.* Feb 1; 143-144:493-501.

**LE Lain R, Barrell KJ, Saeed GS, Nicholls PJ, Simons C, Kirby A, Smith HJ.** (2002). Some coumarins and triphenylethene derivatives as inhibitors of human testes microsomal 17beta-hydroxysteroid dehydrogenase (17beta-HSD type 3): further studies with tamoxifen on the rat testes microsomal enzyme. *J Enzyme Inhib Med Chem.* Apr; 17:93-100.

**Li KX, Smith RE, Krozowski ZS.** (1998). Cloning and expression of a novel tissue specific 17beta-hydroxysteroid dehydrogenase. *Endocr Res.*

Aug-Nov; 24:663-667.

**Lu ML, Huang YW, Lin SX.** (2002). Purification, reconstitution, and steady-state kinetics of the trans-membrane 17 beta-hydroxysteroid dehydrogenase 2. *J Biol Chem.* Jun 21; 277:22123-22130.

**Luu-The V.** (2001). Analysis and characteristics of multiple types of human 17beta-hydroxysteroid dehydrogenase. *J Steroid Biochem Mol Biol.* Jan-Mar; 76:143-151.

**Mannino, RJ, Gould-Fogerite, S.** (1988). Liposome mediated gene transfer. *BioTechniques.* 6:682-690.

**McKeever BM, Hawkins BK, Geissler WM, Wu L, Sheridan RP, Mosley RT, Andersson S.** (2002). Amino acid substitution of arginine 80 in 17beta-hydroxysteroid dehydrogenase type 3 and its effect on NADPH cofactor binding and oxidation/reduction kinetics. *Biochim Biophys Acta.* Nov 19; 1601:29-37.

**Neumann, E, Schaefer-Ridder, M, Wang, Y, Hofschneider, PH.** (1982). Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* 1:841-845.

**Oates, AC, Bruce, AE, Ho RK.** (2000). Too much interference: injection of double-stranded RNA has nonspecific effects in the zebrafish embryo. *Dev Biol.* 224:20-28.

**Provost PR, Blomquist CH, Drolet R, Flamand N, Tremblay Y.** (2002). Androgen inactivation in human lung fibroblasts: variations in levels of 17 beta-hydroxysteroid dehydrogenase type 2 and 5 alpha-reductase activity compatible with androgen inactivation. *J Clin Endocrinol Metab.*

Aug; 87:3883-3892.

**Qiu W, Campbell RL, Gangloff A, Dupuis P, Boivin RP, Tremblay MR, Poirier D, Lin SX.** (2002). A concerted, rational design of type 1 17beta-hydroxysteroid dehydrogenase inhibitors: estradiol-adenosine hybrids with high affinity. *FASEB J.* Nov; 16:1829-1831.

**Rizner TL, Adamski J, Stojan J.** (2000). 17Beta-hydroxysteroid dehydrogenase from *Cochliobolus lunatus*: model structure and substrate specificity. *Arch Biochem Biophys.* Dec 15; 384:255-262.

**Sano T, Hirasawa G, Takeyama J, Darnel AD, Suzuki T, Moriya T, Kato K, Sekine H, Ohara S, Shimosegawa T, Nakamura J, Yoshihama M, Harada N, Sasano H.** (2001). 17 beta-Hydroxysteroid dehydrogenase type 2 expression and enzyme activity in the human gastrointestinal tract. *Clin Sci (Lond).* Nov; 101:485-491.

**Schaefer-Ridder M, Wang Y, Hofschneider PH.** (1982). Liposomes as gene carriers: efficient transformation of mouse L cells by thymidine kinase gene. *Science.* Jan 8; 215:166-168.

**Steven A Farber, Shiu-Ying Ho, Michael Pack.** (2001). Genetic analysis of Digestive Physiology Using Fluorescent Phospholipid Reporters. *Science.* 292:1385-1388.

**Summerton, J.** (1999). Morpholino antisense oligomers: the case for an RNase H-independent structural tupe. *Biochim Biophys Acta.* 1489:141-158.

**Szyperski T, Scheek S, Johansson J, Assmann G, Seedorf U, Wuthrich K.** (1993). NMR determination of the secondary structure and the

three-dimensional polypeptide backbone fold of the human sterol carrier protein 2. *FEBS Lett.* Nov 29; 335:18-26.

**Torn S, Nokelainen P, Kurkela R, Pulkka A, Menjivar M, Ghosh S, Coca-Prados M, Peltoketo H, Isomaa V, Vihko P.** (2003). Production, purification, and functional analysis of recombinant human and mouse 17beta-hydroxysteroid dehydrogenase type 7. *Biochem Biophys Res Commun.* May 23; 305:37-45.

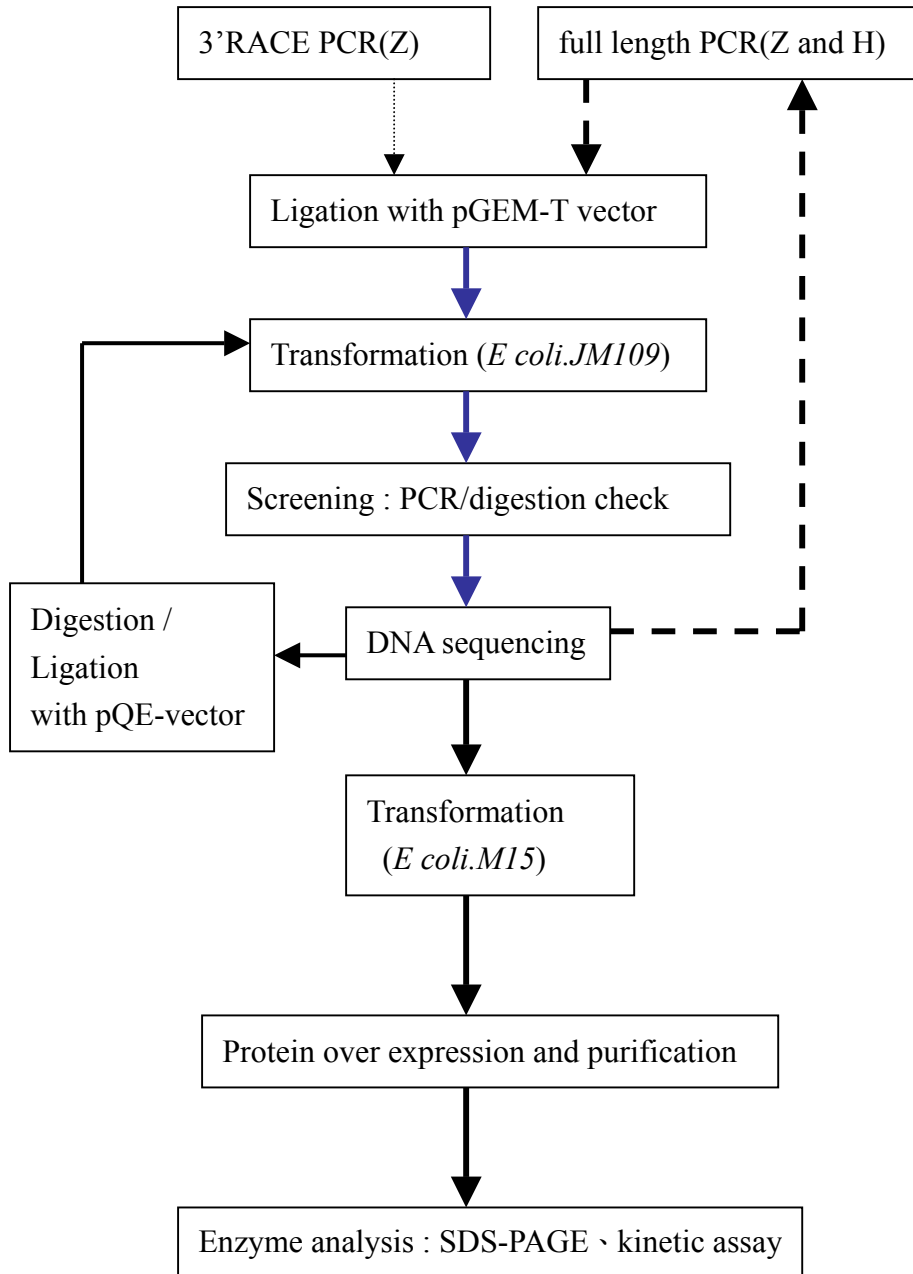
**Vihko P, Harkonen P, Oduwole O, Torn S, Kurkela R, Porvari K, Pulkka A, Isomaa V.** (2002). 17 beta-hydroxysteroid dehydrogenases and cancers. *J Steroid Biochem Mol Biol.* Dec; 83:119-122.

**Xie H, Ram-Mohan LR, Wolff PA.** (1990). Electron-hole recombination in narrow-band-gap Hg<sub>1-x</sub>Cd<sub>x</sub>Te and stimulated emission of LO phonons. *PHYSICAL REVIEW B CONDENSED MATTER.* Aug 15; 42:3620-3627.

**Xie TD, Tsong TY.** (1990). Study of mechanisms of electric field-induced DNA transfection. II. Transfection by low-amplitude, low-frequency alternating electric fields. *Biophys J.* Oct; 58:897-903.

**Zhu B, Liu X.** (1985). Influences of naloxone and DA microinjected in the head of the caudate nucleus on the effect of electroacupuncture in the rat. *Zhen Ci Yan Jiu.* 10:194-198.

## 陸、圖表



圖一：實驗設計流程圖。實驗順序由細虛線箭號依序至粗線箭號。



Human (1) GATAAATGCGGAGGGACGGTCCAGCTTTAGCTCTCTGCTCGCCGCCGCCG

Zebrafish (1) -TTTAATGTGTACCAGGAGCCC-CTCTGGATCATTGCGTTCCCA-----

Human (51) CTGTGCGCCGACCTCCTCTGATCTACGAAAGTC**ATG**TTACCCAACACCG

Zebrafish (44) ---TATCAGACATCATCCAGGAGACACAATA---**ATG**CTGCAGAACACAG

Human (101) GGAGGCTGGCAGGATGTACAGTTTTTATCACAGGTGCAAGCCGTGGCATT

Zebrafish (88) GGAAGCTGGCAGGATGCACCATTTTTCATAACCGGAGCAAGTCGAGGTATT

Human (151) GGCAAAGCTATTGCATTGAAAGCAGCAAAGGATGGAGCAAATATTGTTAT

Zebrafish (138) GGCAAGGCCATTGCTCTGAAAGCTGCGCAGGATGGAGCCAATGTAGTCAT

Human (201) TGCTGCAAAGACCGCCAGCCACATCCAAAACCTTCTAGGCACAATCTATA

Zebrafish (188) CGCTGCCAAAACAGCTGATCCACATCCAAAACCTCCCGGCACCATCTACA

Human (251) CTGCTGCTGAAGAAATTGAAGCAGTTGGAGGAAAGGCCTTGCCATGTATT

Zebrafish (238) CAGCTGCAGCAGAGATTGAAGCAGCTGGAGGAAAGCACTGCCGTGTATT

Human (301) GTTGATGTGAGAGATGAACAGCAGATCAGTGCTGCAGTGGAGAAAGCCAT

Zebrafish (288) GTGGACGTCCGTGATGAGAAGCAGATCAATGATGCTGTTGAACAGGCTGT

Human (351) CAAGAAATTTGGAGCTTATACCATTGCTAAGTATGGTATGTCTATGTATG

Zebrafish (338) GGAGAAGTTTGGAGGG-----ATTGACA--TATTGGTCAACAATGCCAG

Human (401) TGCTTGAATGGCAGAAGAATTTAAAGGTGAAATTGCAGTCAATGCATTA

Zebrafish (380) TGCC-----ATCAATTTAACAGGGACTCTTCAGAC--TCCA--A

Human (451) TGGCCTAAAACAGCCATACACACTGCTGCT-ATGGATATGCTGGGAGG--

Zebrafish (415) TG-----AAAAAGGCTGACCTCATGCTGGGCATCAATCTCAGGGGAACAT

Human (498) ACCTGGTATCGAAAGCCAGTGTAGAAAAGTTGATATCAT-TGCAGATGCA

Zebrafish (460) ACCTGACGTCTAAA--CTGTGCAT-----TCCACATCTTCTGAAGA-GCA

Human (547) GCATATTCCATTTTCCAAAAGCCAAAAGTTTTACTGGCAACTTTGTCAT

Zebrafish (502) AAAACCTCACATACTAAACCTCA----GCCACCTCTCAACCTCAACCC

Human (597) TGATGAAAATATCTTAAAAGAAGGAATAGAAAATTTTGACGTTTATG

Zebrafish (548) CATTTGGTTTA----AAAAC**CACACGGCTTACACCATTTGCAAGTATG-G**

3' RACE **G-G**

Human (647) CAATTAACCAGGTCATCCTTTGCAA--CCAGATTTCTTCTTAGATGAAT

Zebrafish (593) **CATGTN---CATGTGTGTACTGGGAATGGCAGAAGAGTTCAGA**-----

3' RACE (593) **CATGTC---CATGTGTGTACTGGGAATGGCAGAAGAGTTCAGAGGATCCA**

Human (695) ACCCAGAAGCAGTTAGCAAGAAAAGTGGAAATCAACTGGTGTGTTCCAGAA

3' RACE (640) **TTGCCGTTAATGCCTTATGGCCAAAGACAGCCATTCAGACGGCTGCCATG**

Human (745) TTCAAAGAAGAGAAAAGTGCAGCTGCAACCAAACCACGTTCTGGAGCTGT

3' RACE (690) **GACATGTTGGGCGGTTCTGAGGTTGGCAAGCAGTGCAGGAAGGTGGAGAT**

Human (795) GGAAGAAACATTTAGAATTGTTAAGGACTCTCTCAGTGATGATGTTGTTA

3' RACE (740) **CATGGCTGATGCAGCATACGCCATCTTCAAACAACCCACCAGCTCACCG**

Human (845) AAGCCACTCAAGCAATCTATCTGTTTGAAGTCTCCGGTGAAGATGGTGGC

3' RACE (790) **GACAGTTTGTTATTGATGAGGACATTCTCAAAAAGGAGGGCATTAAAGAT**

Human	(895)	ACGTGGTTTCTTGATCTGAAAAGCAAGGGTGGGAATGTCGGATATGGAGA
3' RACE	(839)	TTTGATGTTTATGC-TGTTGAGCCAGGTCATCCATTGCTTCTGACTTTTT
Human	(945)	GCCTTCTGATCAGGCAGATGTGGTGATGAGTATGACTACTGATGACTTTG
3' RACE	(888)	CTTGGACGGCCAGCCTGAGGATCTAGTCAAGCATATGGAGGCACATGGTG
Human	(995)	TAAAAATGTTTTTCAGGGAAACTAAAACCAACAATGGCATTTCATGTCAGGG
3' RACE	(936)	CCACTCCGGCGTTCAACAAC-TGCAAAAGCAGATCCCCTTGCCGCAGGACC
Human	(1045)	AAATTGAAGATTAAAGGTAACATGGCCCTAGCAATCAAATTGGAGAAGCT
3' RACE	(985)	AGTTTCTGAGATG-TCAATACAATCAGAGGAATTATCAGTCCAGAGATGG
Human	(1095)	AATGAATCAGATGAATGCCAGACTGTGAAGGAAAATATAAAAAAAGTC
3' RACE	(1035)	TGAAAACCAACAAGGAGTGTCAAA-TTAACTTGGCAGGCGAACATGCTG
Human	(1145)	GACTGCTATGCTCAAAAAGTAAAAAAGCTCAACAGTTAAAATCTAATGT
3' RACE	(1085)	GAG-CTGGNATCTTG-CCTGAAGAACGAT-CTGGAA-TGCTGGAAATGGG
Human	(1195)	TTGTTTTCTTTCTGTTATATTATAAGGATATGCACGTTTTGTTCTGGAAA
3' RACE	(1135)	GAACCTTCTGTNAA--CTGATGTTG-CATGTCGATGGCCA-TGAGGA-TT
Human	(1245)	AGATAGAATTTGTCTCTAAAAGACTTGAAATTGTAATTAATGGCAAGC
3' RACE	(1185)	TGNCAANATGNTTGGANGGA TGTCAANATGTTTGNAGGGAAATTAANCCANCCATGNCCTTCATGTCTG
Human	(1295)	TAATCAAACATAAGCTTCATTAAGTGGGATTCTAAGACAGTCTGTGTTTT
3' RACE	(1235)	GAAAGCTGACTATTAAGGGTGACATGGCCCTTGTTCATCAAAGTGGAGAAG
Human	(1345)	TATATTTCAAGGGTTAACCCTTTGAGCCTTACATCTCATTCACTGTCTT
3' RACE	(1285)	ATGATGGCCATGATGAAGTCTAAACTGTGAACNCATCACGCAGTNTCCTC
Human	(1395)	TCTCCAAGAAAAGTATTTTGGGCGGACAGTCAGATCAAGCAGTAAAATTA
3' RACE	(1335)	CCTTACAANATGTACATNATGCCCTCTTTTCTANATCAGTGTTTTTTGTAT
Human	(1445)	GCTCTTTCAAATCTTCTTGTTCATGTAATAAGGCTAGTCTGTTTTAAAA
3' RACE	(1385)	TAAACTAATTGTTGAGATTGTACTAATTTTAATAAAGAACTATCAGGCAA
Human	(1495)	TTTTTAGTTTTGGATTGTATACTAATGAAAATCTTAATGATGTTTTTGTAT
3'RACE	(1435)	AAAAAAAAAAAAAAAAAAAAAAAAAAAA

圖二：Z-HSD 核酸組裝序列與 H-HSD 之核酸序列比對。將 Z-HSD 由

3' RACE PCR 所延長之序列與原 Z-HSD 片段序列組裝，並與 H-HSD 之核酸序列比對，參考 H-HSD 之起始密碼子和終止密碼子(紅字框格區域)，以判斷 Z-HSD 之蛋白質編碼區(底線標示處為可能之始密碼子和終止密碼子)。Z-HSD 之核酸序列第 564~632 為 3' RACE 延長序列之重疊部分(黃色部分)，藍字部分為 3' RACE 延長之 DNA 序列。灰色陰影處表示其核酸序列相同。

```

          *           20           *           40           *           6
Human      : MLPNTGRLAGCTVFITGASRGIGKAIALKAARDGANIVIAAKTAQHPKLLGTIYTAAE : 59
Zebrafish : MLQNTGKLAGCTIFITGASRGIGKAIALKAARDGANVVIAAKTADHPKLPGTIYTAAA : 59
          ML NTG4LAGCT6FITGASRGIGKAIALKAA DGAN6VIAAKTA PHPKL GTIYTAA

          0           *           80           *           100           *           1
Human      : EIEAVGGKALPCIVDVRDEQQISAAVEKAIKKFG-----AYT----- : 96
Zebrafish : EIEAAGKALPCIVDVRDEKQINDAVEQAVEKFGGIDILVNNASAINLTGTLQTPMEKA : 118
          EIEA GKGALPCIVDVRDE QI AVE A6 KFG T

          20           *           140           *           160           *
Human      : -----IAKYG---MSMYVLGMAEEFKGEIAVN-----ALWPK--TATH----- : 129
Zebrafish : DLMLGINLRGTYLTSKLCTPHLLSKNPHILNLSPPLNLHPWPKNHTGLHHCKVWHVQ : 177
          I G S 6 K 6N 6W K T 6H

          180           *           200           *           220           *
Human      : -----TAAMD---M-LGG--PG--IESQCRKVDIADAA--YS--IFQ-----KP : 162
Zebrafish : CVTWEWQKSSQDPLPLPYGQRQPFRLPWTCWAVLRLASSAGRWRSWLMQHTPSSNNE : 236
          D M G P 6 C V 6A A 5 6 Q P

          240           *           260           *           280           *
Human      : KSFTGNFVIDENILKEEGIENFDVYAIKPGHPLQPDFFLDEYPEAVSKKVESTGAMPEE : 221
Zebrafish : PASPDSLLLMRTFSKRRALKILMFMLSQVIHCFPDFFLDGQPEDLVKHMEAHGATPAF : 295
          66 K 6 6 PDFFLD PE 6 K 6E GA P F

          300           *           320           *           340           *
Human      : KEEKLQLQPKPRSGAVETERIVKDSLSDDVVKATQAIYLELSGEDGGTWFLDLKSKG : 280
Zebrafish : TTAKAD--PVA-AGPWSEMENTIRGISPEMVKTTQGVYKENLAGEHAGVWYLDLKNDA : 351
          K P G V E F 64 6S 6VK TQ 6Y F L GE G W5LDLK

          360           *           380           *           400           *
Human      : GNVGYGEPSDQADVVMSMTTDDFVKMFSGKLKPTMAFMSGKLKIKGNMALAIKLEKLMN : 339
Zebrafish : GSACNGEPPVKADVVMSMDSEDFVKMFSGKLKPTMAFMSGKLTIKGDMALVIKLEKMMA : 410
          G G GEP ADVVMSM 3 DFVKMF GKLKPTMAFMSGKL IKG1MAL IKLEK6M

Human      : QMNARL : 345
Zebrafish : MMKSKL : 416
          M 4L

```

圖三：Z-HSD 與 H-HSD 之胺基酸序列比對。黑色區塊表示胺基酸序列相同。



## Zebrafish domain search



## Human (NP\_115679) domain search

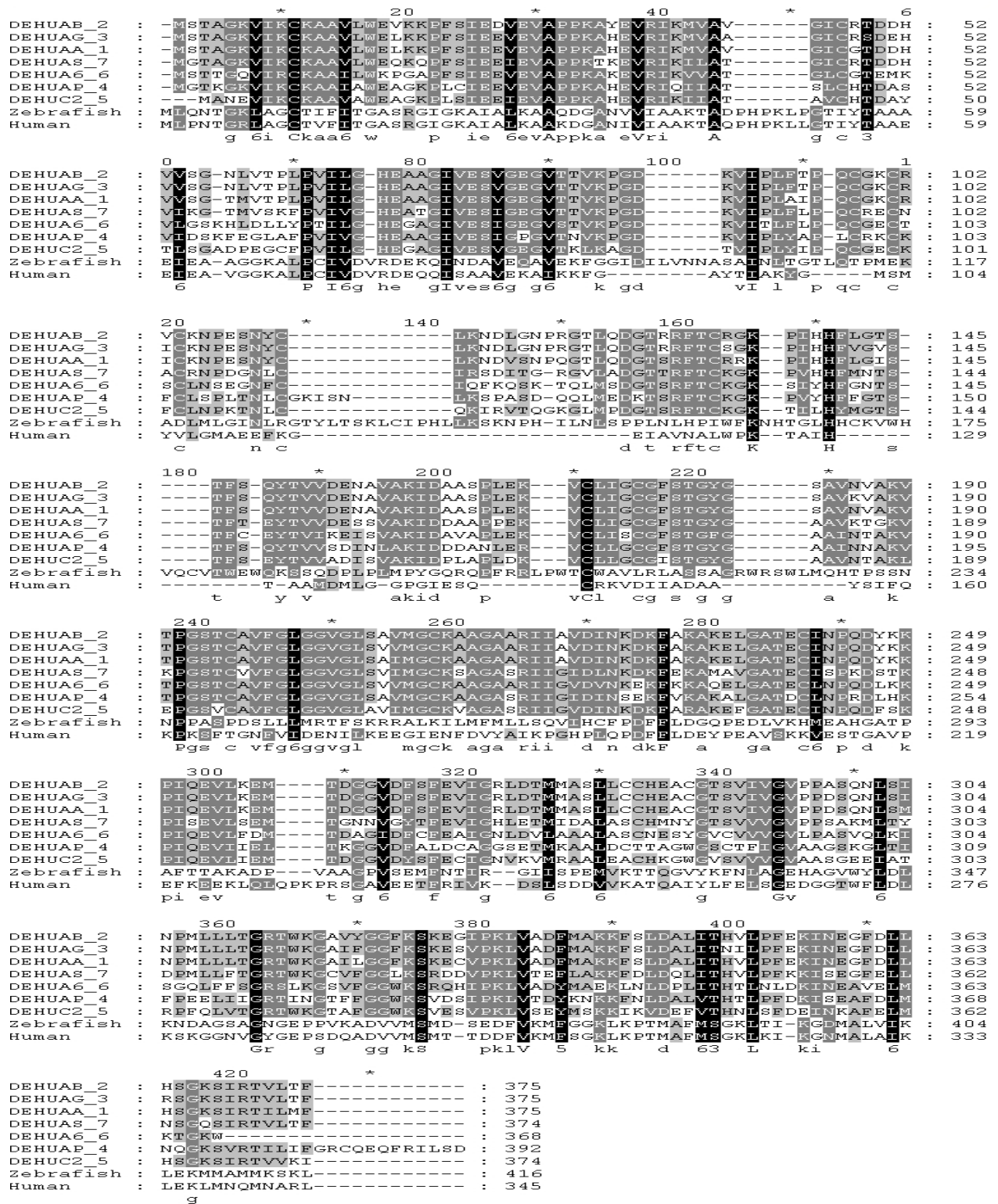


(摘錄自 NCBI 網站)

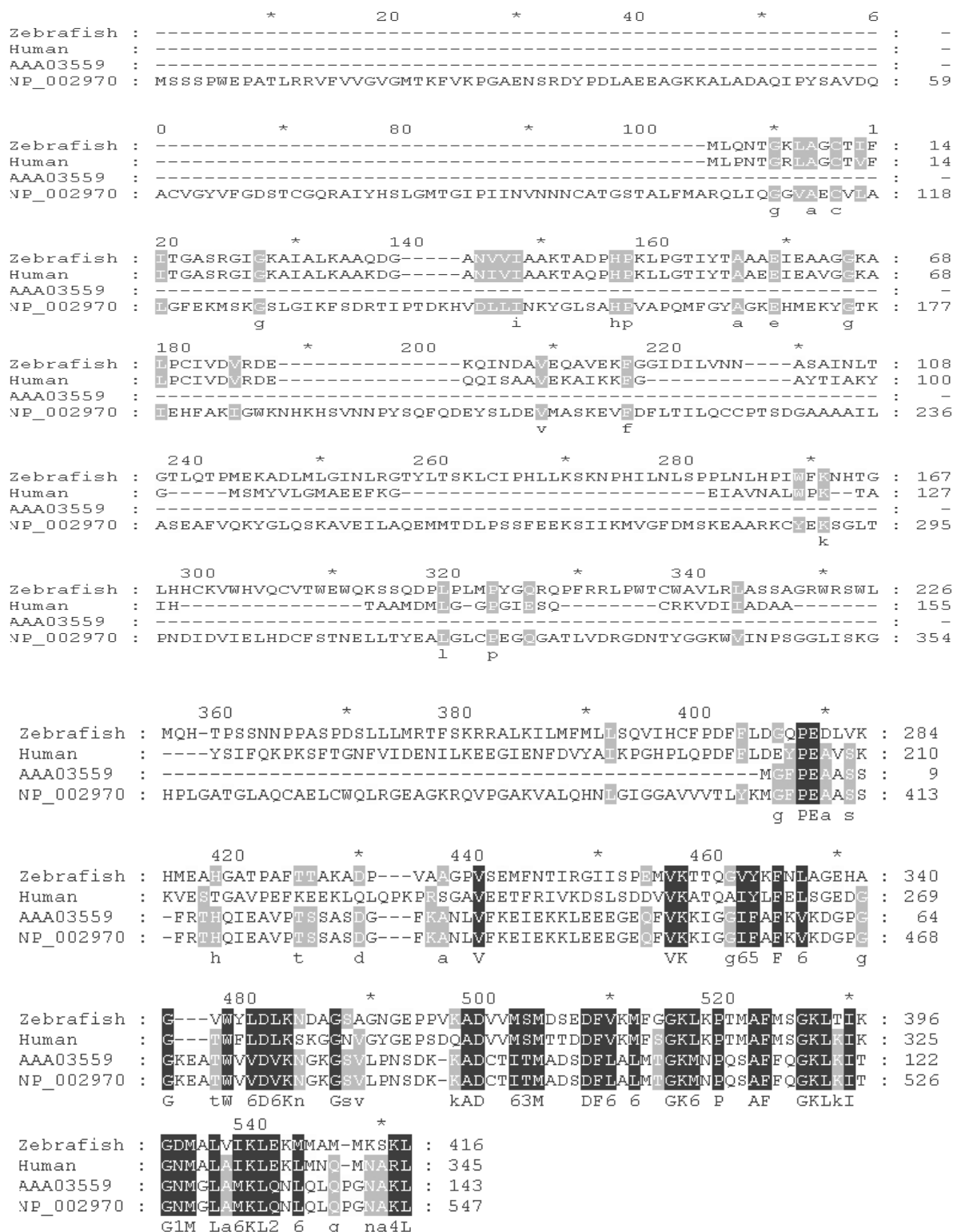
### 圖四：Z-HSD 和 H-HSD 之新穎酵素的結合結構區。NCBI 之內 RPS

BLAST 搜尋保留結合結構區(conserved domain)資料庫之結果，兩者皆具有相同之二個結構區，分別為 N 端之 adh\_short 與 C 端之 SCP2, FabG / COG4221 與 adh\_short 同義；COG3255 與 SCP2 同義。





圖五：Z-HSD 及 H-HSD 與酒精去氫酵素之胺基酸序列多重比對。上列之胺基酸序列分別為人類已知之酒精去氫酵素 type1~type7 和 H-HSD 及 Z-HSD 之胺基酸序列，黑色區塊表示胺基酸序列相同，灰色陰影處為胺基酸之屬性相似。



圖六：Z-HSD 及 H-HSD 與固醇類攜帶蛋白之胺基酸序列多重比對。

上列之胺基酸序列為人類已知的固醇類攜帶蛋白(AAA03559/ NP\_002970)與 H-HSD 及 Z-HSD 之胺基酸序列，黑色區塊表示胺基酸序列相同，灰色陰影處為胺基酸之屬性相似。

```

*          20          *          40          *
M.musculus : ----- : -
R.norvegic : ----- : -
H.sapiens  : ----- : -
human      : ----- : -
C.elegans  : MSRRVEFDLSTEDHSRRRTNTFSSDEEDGVENEVADYLVVYFSRMVDEQNVP EILTLYD : 58
zebrafish  : ----- : -

60          *          80          *          100          *
M.musculus : ----- : -
R.norvegic : ----- : -
H.sapiens  : ----- : -
human      : ----- : -
C.elegans  : QAF PDLTEREF RDRMWPDENVVERIIGPGNKLFIIILYKELYRQLYARNTRG PLLVHR : 116
zebrafish  : ----- : -

120         *          140         *          160         *
M.musculus : ----- : -
R.norvegic : ----- : -
H.sapiens  : ----- : -
human      : ----- : -
C.elegans  : YESFMNYQELFSELLSSKDPIPLSLPNVWLWDIIDEFVYQFQAFCLYKANPGKRNADE : 174
zebrafish  : ----- : -

180         *          200         *          220         *
M.musculus : ----- MASPLRFDGR : 10
R.norvegic : ----- MASPLRFDGR : 10
H.sapiens  : ----- MGSPLRFDGR : 10
human      : ----- : -
C.elegans  : VEDLINIEENQNAWNIYPVLNIIYLSLLSKSQIVEQLKALKEKRNPDSSVADEFGQSDLY : 232
zebrafish  : ----- : -

240         *          260         *          280         *
M.musculus : VVLVTGPGGGIGRAYALAF AERGA LVI VNDLGGDFKGVG-----KGS SAADKVVA : 60
R.norvegic : VVLVTGAGGGLGRAYALAF AERGA LVVNDLGGDFKGVG-----KGS SAADKVVE : 60
H.sapiens  : VVLVTGAGAGLGRAYALAF AERGA LVVNDLGGDFKGVG-----KGS LAADKVVE : 60
human      : ----- : -
C.elegans  : FKLGYPALIGLLRTHVLLGDYHQALKTVOYVDIDPKGIYNTVPTCLVTLHYFVGF SHL : 290
zebrafish  : ----- : -

300         *          320         *          340
M.musculus : EIRRRGGKAVANYDSVEAGEKLVK TALDTPGRIDVVVNNAGILRDRSEFRISDEDWDI : 118
R.norvegic : EIRRRGGKAVANYDSVEAGEKLVK TALDTPGRIDVVVNNAGILRDRSEFRISDEDWDI : 118
H.sapiens  : EIRRRGGKAVANYDSVEAGEKLVK TALDAGRIDVVVNNAGILRDRSEFRISDEDWDI : 118
human      : ----- : -
C.elegans  : MMRNYGEATKMFVNCLLYIQRTKSVQNOQPSKKNFQYDVI GKTWDQLEHLLAICLAIQ : 348
zebrafish  : ----- : -

*          360          *          380          *          400
M.musculus : IHRVHLRGSFQVTRA AWDHMKKQNYGRILMTSSASGIYGNFGQANYSAAKLG----- : 170
R.norvegic : IQRVHLRGSFQVTRA AWDHMKKQNYGRILMTASASGIYGNFGQANYSAAKLG----- : 170
H.sapiens  : IHRVHLRGSFQVTRA AWEHMKKQKYGRILMTSSASGIYGNFGQANYSAAKLG----- : 170
human      : ----- : -
C.elegans  : PQRIDESIASQLSERCGERMHMANGNIDEFRNAFATGCPKFLSPTTVVYEGVNSKE : 406
zebrafish  : ----- : -

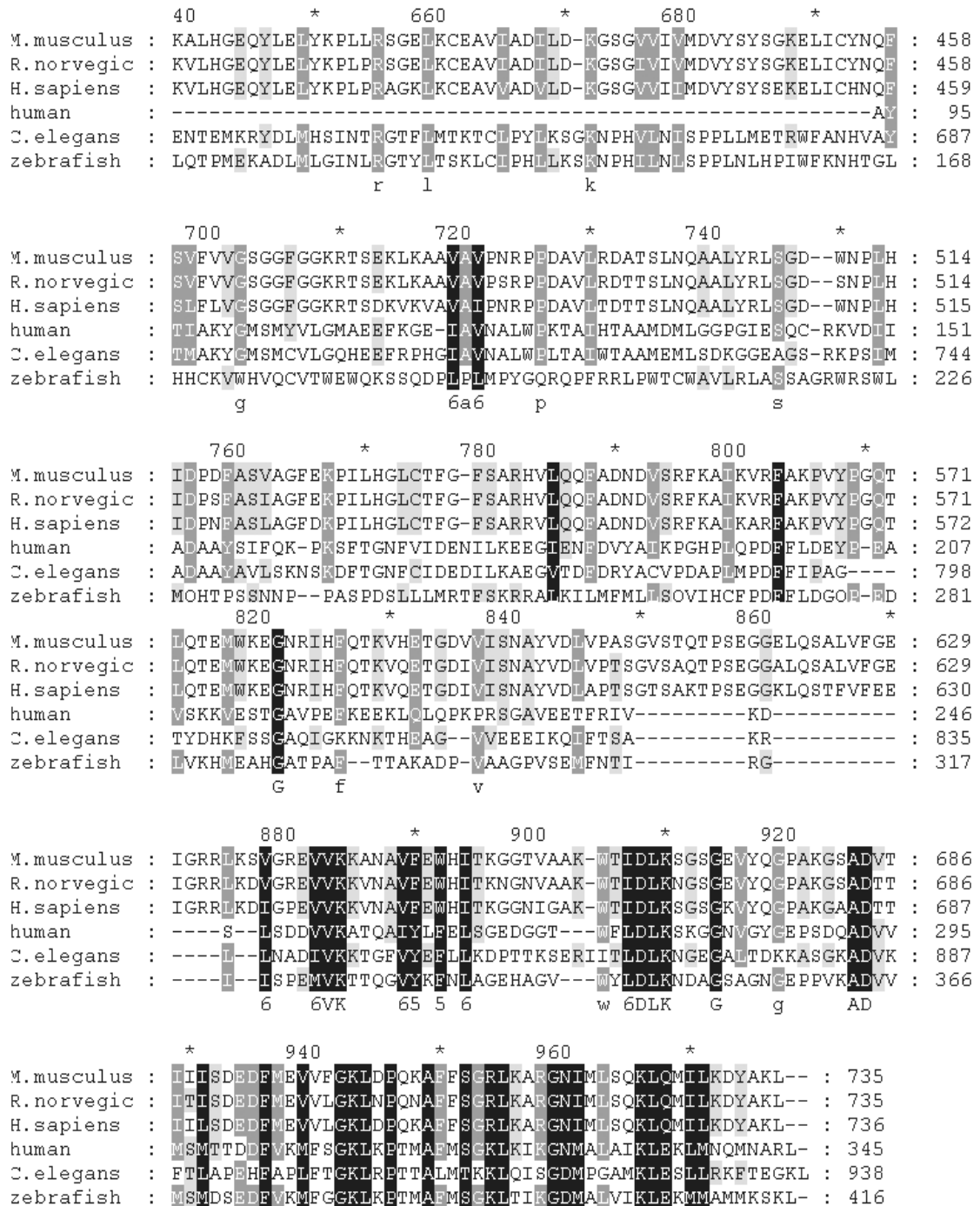
*          420          *          440          *          460
M.musculus : ILGLCNTLAIIEGRKNNIHCNTIAPNAGSRMTETVLPEDLVEALKPEYVAPLVLWLICHE : 228
R.norvegic : LLGLANTLVIEGRKNNIHCNTIAPNAGSRMTETVMPEDLVEALKPEYVAPLVLWLICHE : 228
H.sapiens  : LLGLANSLAIEGRKNNIHCNTIAPNAGSRMTQTVMPEDLVEALKPEYVAPLVLWLICHE : 228
human      : ----- : -
C.elegans  : PLLRQTQSFLLEGIESQMALPVLRGYLKLYTTLPTKKLASFMDVDEHYDSF I GKLITY : 464
zebrafish  : ----- : -

*          480          *          500          *          520
M.musculus : SCEENGLFEV GAGWIGKLRWERTLGAI VRKRNQPM TPEAVRDNWEKICDFSNASKPQ : 286
R.norvegic : SCEENGLFEV GAGWIGKLRWERTLGAI VRKRNQPM TPEAVRDNWKICDFSNASKPK : 286
H.sapiens  : SCEENGLFEV GAGWIGKLRWERTLGAI VRQKNHPMTPEAVKANWKKICDFENASKPQ : 286
human      : -----MLP : 3
C.elegans  : KMIVNELGKEAGPSSADDEPQTDIDFYVDRDMINIADTKVARHVGCQOTTRYPETMI : 522
zebrafish  : -----MLQ : 3

*          540          *          560          *          580
M.musculus : TIQESTGGIIEVLHKVDSEG--ISP NRTSHAAPAATSCFVGVAVGHRKLPFSSSYTEIQS : 343
R.norvegic : SIQESTGGIIEVLHKIDSEG--ISQNH TQGVASADASC FAVGVGHRKLPFSSSYTEIQC : 343
H.sapiens  : SIQESTSIIIEVLSKIDSEGGV SANHTSRATSTATSCFAGAI GQKLPFSSSYAYTEEA : 344
human      : NTGRLAGCTVFITGASRGIG-----KAI ALKA AKDCANIVVIAAKT---AQPHPKLLG : 52
C.elegans  : LKKKFVQRTVLTITGASRGIG-----KEI ALKLAKDCANIVVIAAKT---ATAHPKLPG : 571
zebrafish  : NTGKLACTVFITGASRGIG-----KAI ALKAAQDCANVVIAAKT---ADPHKLPG : 52
          G 6 6          G          a a G 6 K          L

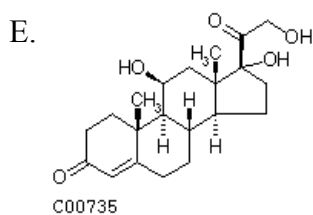
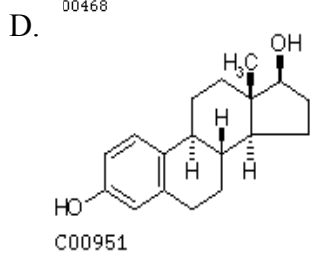
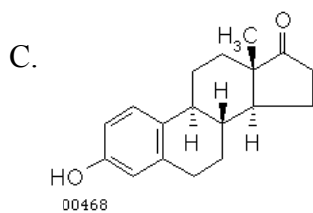
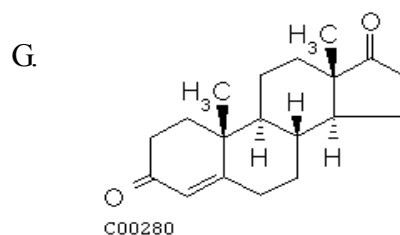
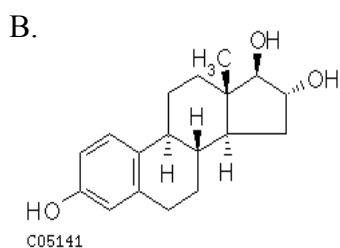
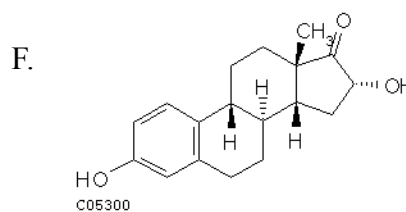
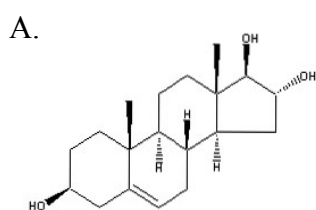
*          600          *          620          *          6
M.musculus : IMYALGVGASVKNPKDKFVYEGSADF SCLPTFGVIVAQKSMNGGTA EVPGLSINFA : 401
R.norvegic : IMYALGVGASVKNPKDKFVYEGSADF SCLPTFGVIVAQKSLMSGGTA EVPGLSINFA : 401
H.sapiens  : IMYALGVGASIKDPKDKFVYEGSADF SCLPTFGVITGQKSMGGGTA EI PGLSINFA : 402
human      : TIYTA AEEIEAVGGKALPCIVDVRDEQI SAAVEKATKKFG----- : 93
C.elegans  : TIYSA AEEIEKAGGKALPCIVDVRDEASVKASVEEAVKKEGGIDILNNASAI SLTDT : 629
zebrafish  : TIYTA A AEEIEAAGKALPCIVDVRDEKQINDAVEQAVKKEGGIDILNNASAI NLGT : 110

```



圖七：物種間 Z-HSD 及 H-HSD 胺基酸序列同源相似性之多重比

對。查尋 NCBI 之 UniGene 資料庫得知斑馬魚與其對應之新穎 H-HSD 蛋白，分別與人類 (*H.sapiens*)、大鼠 (*R.norvegicus*)、小鼠 (*M.musculus*) 和線蟲 (*C.elegans*) 等具有同源相似。黑色區塊表示胺基酸序列相同，灰色陰影處為胺基酸之屬性相似。



	Molecular formula	molecular weight
A. 5-androstene-3 $\beta$ , 16 $\alpha$ , 17 $\beta$ -triol	C <sub>19</sub> H <sub>30</sub> O <sub>3</sub>	306.4
B. Estradiol	C <sub>18</sub> H <sub>24</sub> O <sub>3</sub>	288.4
C. Estrone	C <sub>18</sub> H <sub>22</sub> O <sub>2</sub>	270.4
D. Estradiol	C <sub>18</sub> H <sub>24</sub> O <sub>2</sub>	272.4
E. Hydrocortisone	C <sub>21</sub> H <sub>30</sub> O <sub>5</sub>	362.5
F.16alpha-Hydroxyestrone	C <sub>18</sub> H <sub>22</sub> O <sub>3</sub>	286.4
G. 4-Androstene	C <sub>19</sub> H <sub>26</sub> O <sub>2</sub>	286.4

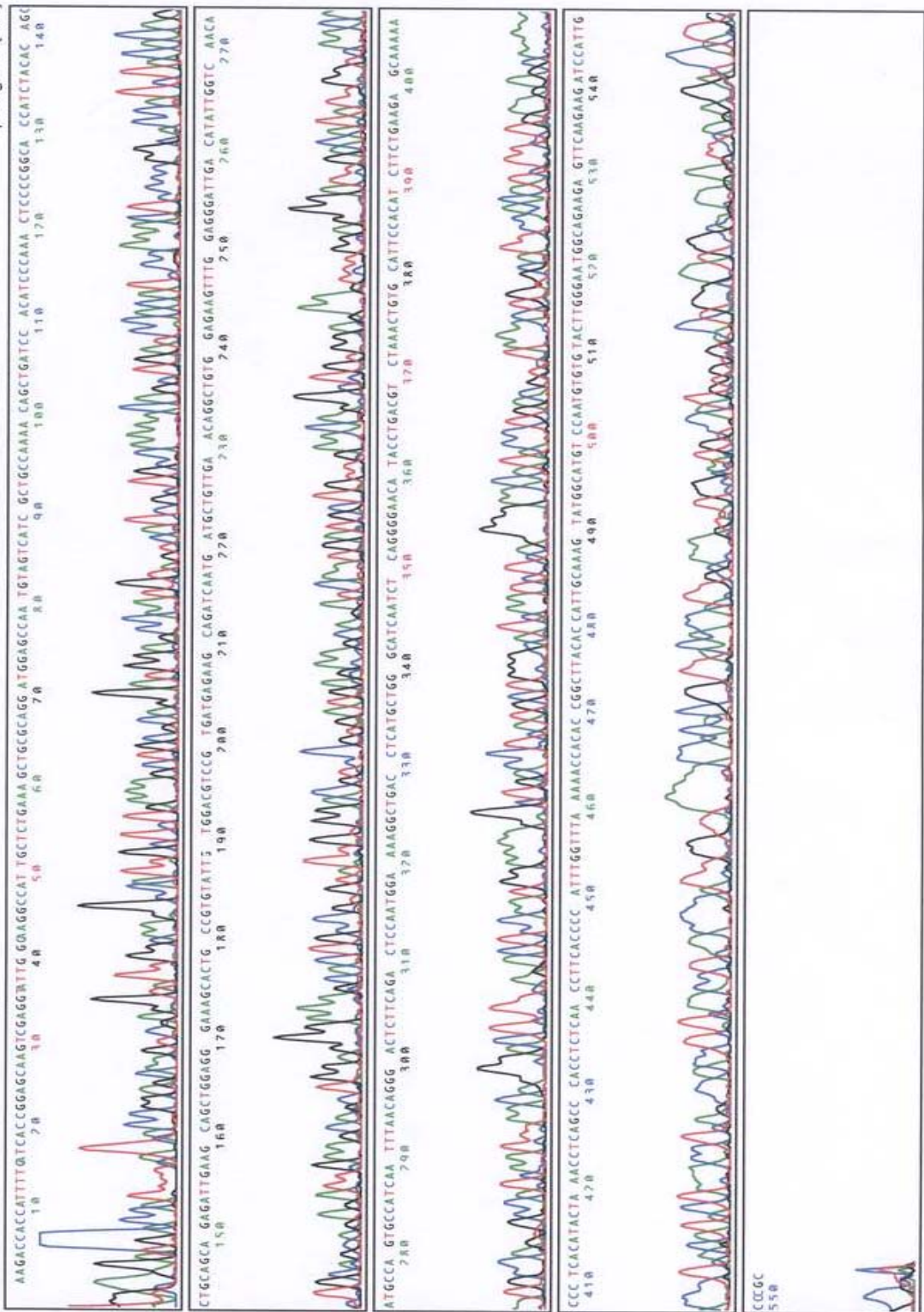
圖八：酵素反應候選受質之造構。



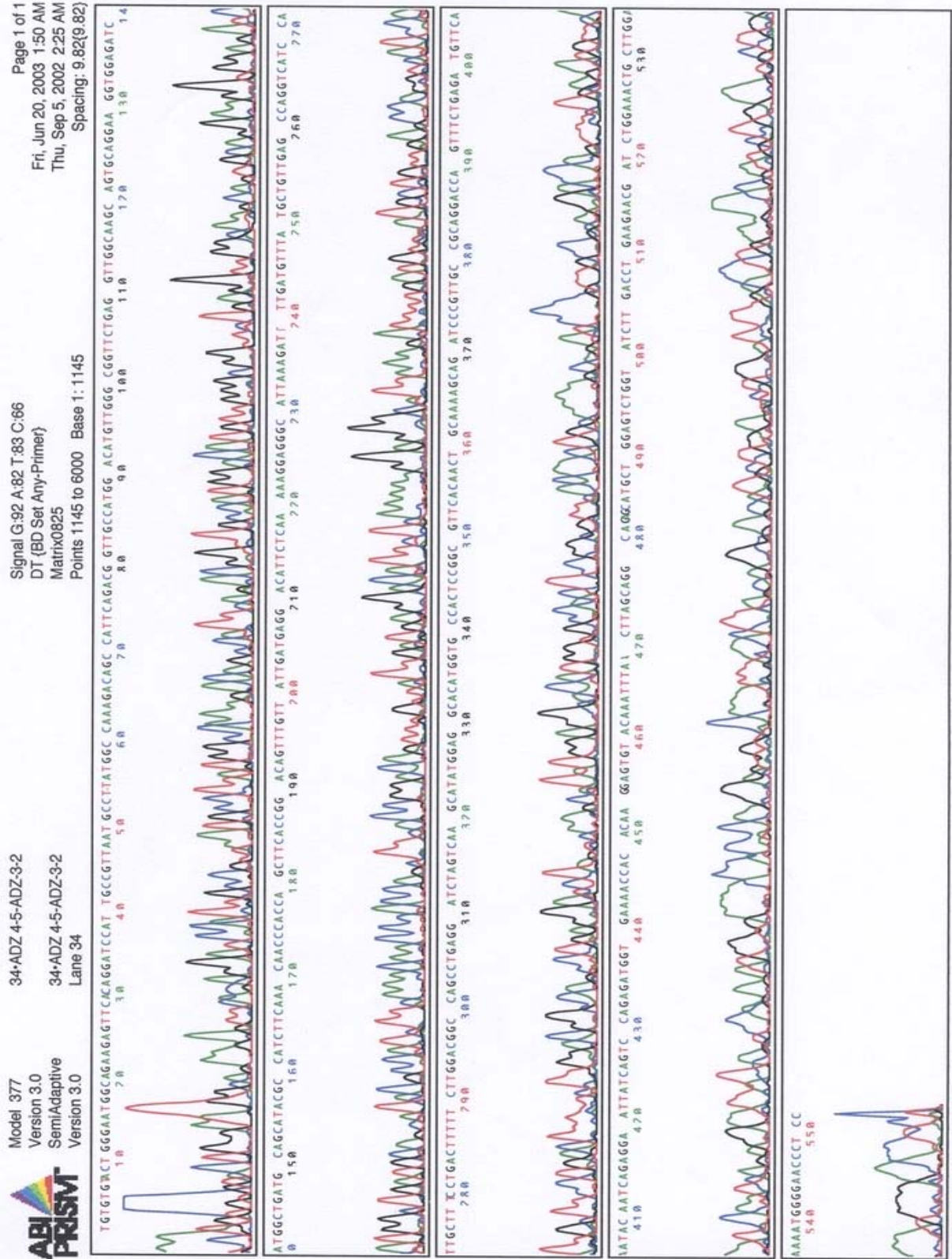






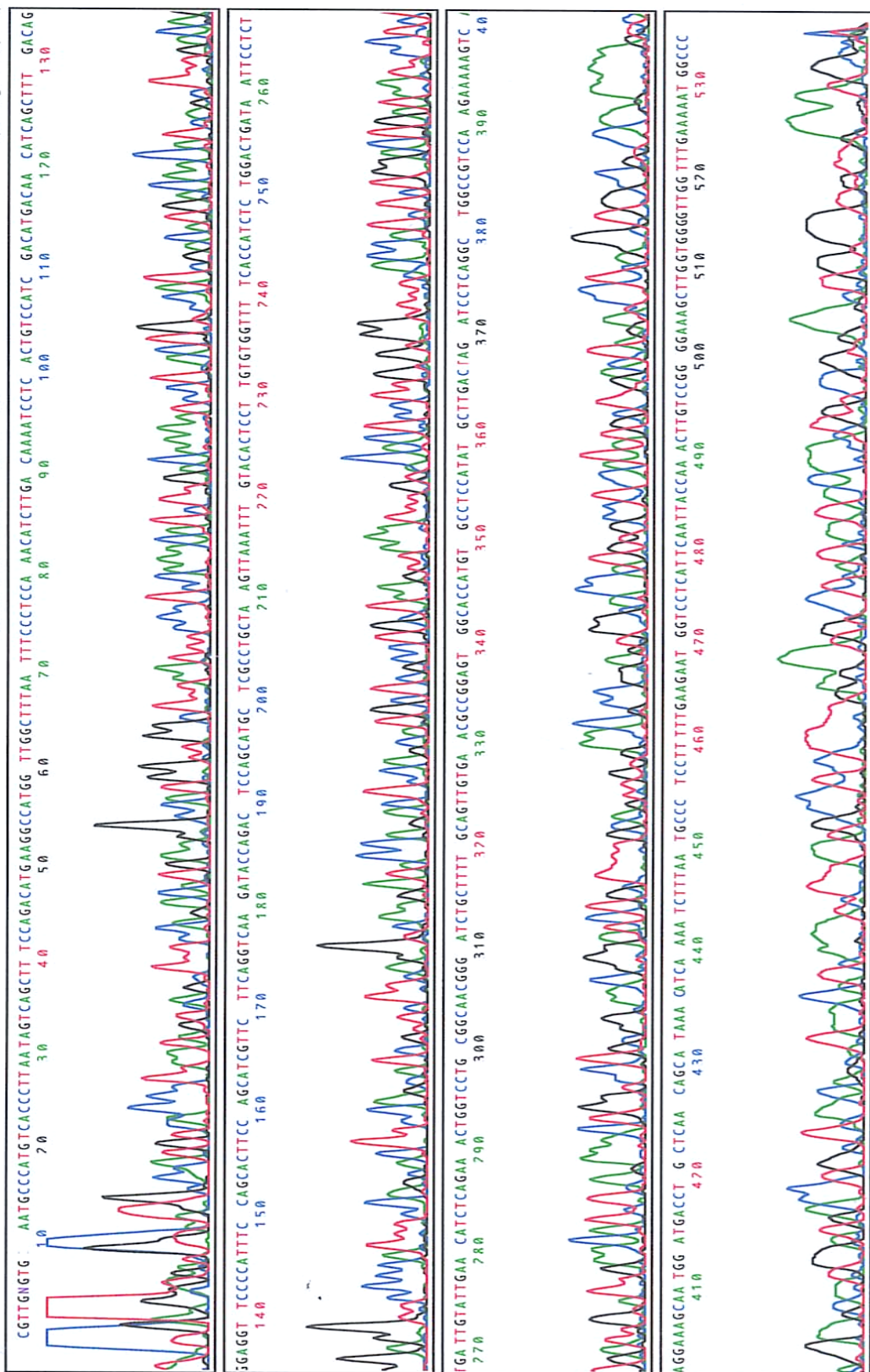


圖十一：Z-HSD 全長 DNA 定序圖 I。



圖十二：Z-HSD 全長 DNA 定序圖 II。





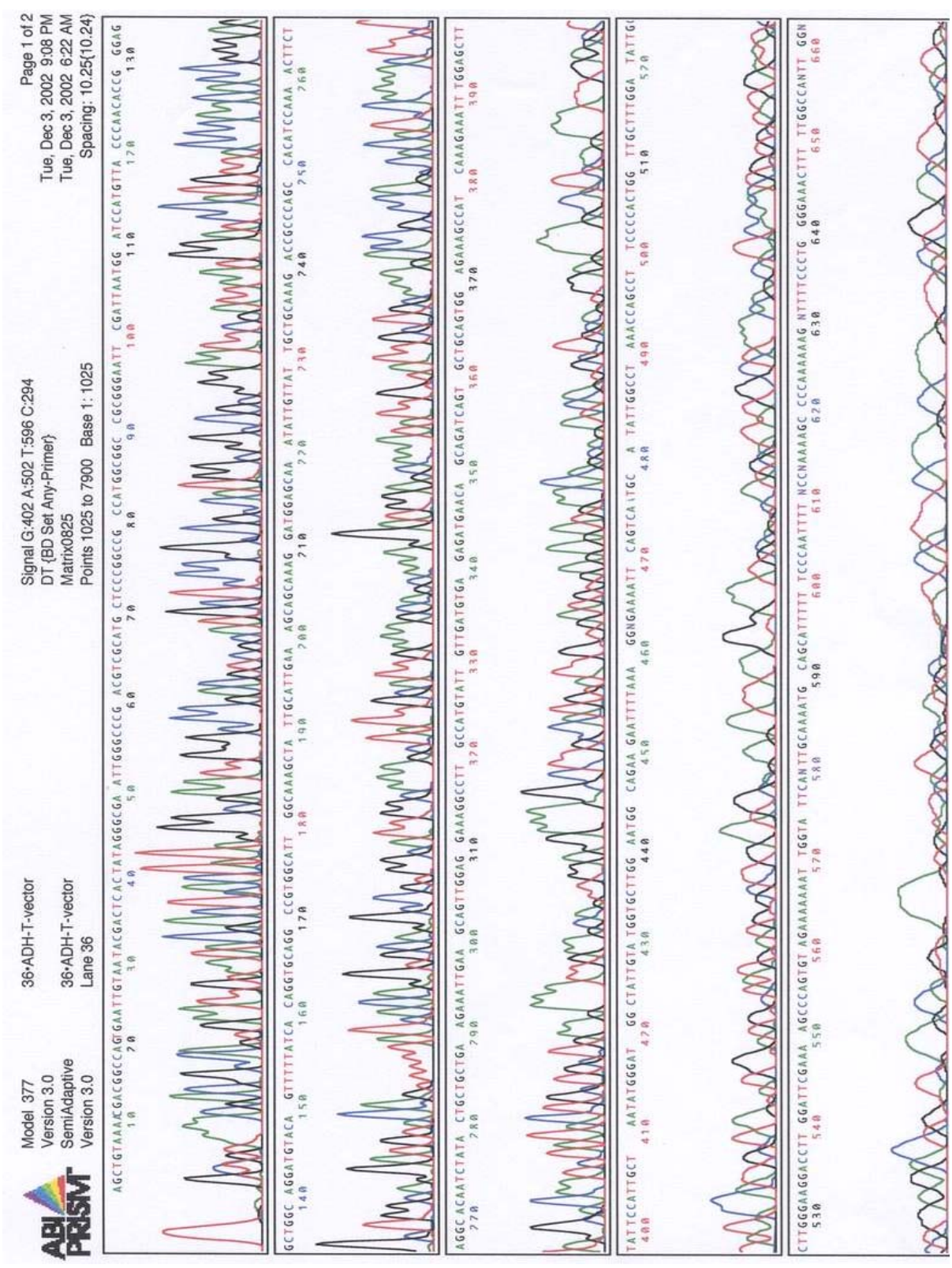
圖十三：Z-HSD 全長 DNA 定序圖 III。

5' UTR-tttaatgtgtcaccaggagccccctctggatcattgcggttcccatatcagacatcatccaggagacacaata 71

ATG	CTG	CAG	AAC	ACA	GGG	AAG	CTG	GCA	GGA	TGC	ACC	ATT	TTC	ATC	116
<b>Met</b>	Leu	Gln	Asn	Thr	Gly	Lys	Leu	Ala	Gly	Cys	Thr	Ile	Phe	Ile	
ACC	GGA	GCA	AGT	CGA	GGT	ATT	GGC	AAG	GCC	ATT	GCT	CTG	AAA	GCT	161
Thr	Gly	Ala	Ser	Arg	Gly	Ile	Gly	Lys	Ala	Ile	Ala	Leu	Lys	Ala	
GCG	CAG	GAT	GGA	GCC	AAT	GTA	GTC	ATC	GCT	GCC	AAA	ACA	GCT	GAT	206
Ala	Gln	Asp	Gly	Ala	Asn	Val	Val	Ile	Ala	Ala	Lys	Thr	Ala	Asp	
CCA	CAT	CCC	AAA	CTC	CCC	GGC	ACC	ATC	TAC	ACA	GCT	GCA	GCA	GAG	251
Pro	His	Pro	Lys	Leu	Pro	Gly	Thr	Ile	Tyr	Thr	Ala	Ala	Ala	Glu	
ATT	GAA	GCA	GCT	GGA	GGG	AAA	GCA	CTG	CCG	TGT	ATT	GTG	GAC	GTC	296
Ile	Glu	Ala	Ala	Gly	Gly	Lys	Ala	Leu	Pro	Cys	Ile	Val	Asp	Val	
CGT	GAT	GAG	AAG	CAG	ATC	AAT	GAT	GCT	GTT	GAA	CAG	GCT	GTG	GAG	341
Arg	Asp	Glu	Lys	Gln	Ile	Asn	Asp	Ala	Val	Glu	Gln	Ala	Val	Glu	
AAG	TTT	GGA	GGG	ATT	GAC	ATA	TTG	GTC	AAC	AAT	GCC	AGT	GCC	ATC	386
Lys	Phe	Gly	Gly	Ile	Asp	Ile	Leu	Val	Asn	Asn	Ala	Ser	Ala	Ile	
AAT	TTA	ACA	GGG	ACT	CTT	CAG	ACT	CCA	ATG	GAA	AAG	GCT	GAC	CTC	431
Asn	Leu	Thr	Gly	Thr	Leu	Gln	Thr	Pro	Met	Glu	Lys	Ala	Asp	Leu	
ATG	CTG	GGC	ATC	AAT	CTC	AGG	GGA	ACA	TAC	CTG	ACG	TCT	AAA	CTG	476
Met	Leu	Gly	Ile	Asn	Leu	Arg	Gly	Thr	Tyr	Leu	Thr	Ser	Lys	Leu	
TGC	ATT	CCA	CAT	CTT	CTG	AAG	AGC	AAA	AAC	CCT	CAC	ATA	CTA	AAC	521
Cys	Ile	Pro	His	Leu	Lys	Lys	Ser	Lys	Asn	Pro	His	Ile	Leu	Asn	
CTC	AGC	CCA	CCT	CTC	AAC	CTT	CAC	CCC	ATT	TGG	TTT	AAA	AAC	CAC	566
Leu	Ser	Pro	Pro	Leu	Asn	Leu	His	Pro	Ile	Trp	Phe	Lys	Asn	His	
ACC	GGC	TTA	CAC	CAT	TGC	AAA	GTA	TGG	CAT	GTC	CAA	TGT	GTG	ACT	611
Thr	Gly	Leu	His	His	Cys	Lys	Val	Trp	His	Val	Gln	Cys	Val	Thr	
TGG	GAA	TGG	CAG	AAG	AGT	TCA	CAG	GAT	CCA	TTG	CCG	TTA	ATG	CCT	656
Trp	Glu	Trp	Gln	Lys	Ser	Ser	Gln	Asp	Pro	Leu	Pro	Leu	Met	Pro	
TAT	GGC	CAA	AGA	CAG	CCA	TTC	AGA	CGG	TTG	CCA	TGG	ACA	TGT	TGG	701
Tyr	Gly	Gln	Arg	Gln	Pro	Phe	Arg	Arg	Leu	Pro	Trp	Thr	Cys	Trp	
GCG	GTT	CTG	AGG	TTG	GCA	AGC	AGT	GCA	GGA	AGG	TGG	AGA	TCA	TGG	746
Ala	Val	Leu	Arg	Leu	Ala	Ser	Ser	Ala	Gly	Arg	Trp	Arg	Ser	Trp	
CTG	ATG	CAG	CAT	ACG	CCA	TCT	TCA	AAC	AAC	CCA	CCA	GCT	TCA	CCG	791
Leu	Met	Gln	His	Thr	Pro	Ser	Ser	Asn	Asn	Pro	Pro	Ala	Ser	Pro	
GAC	AGT	TTG	TTA	TTG	ATG	AGG	ACA	TTC	TCA	AAA	AGG	AGG	GCA	TTA	836
Asp	Ser	Leu	Leu	Leu	Met	Arg	Thr	Phe	Ser	Lys	Arg	Arg	Ala	Leu	
AAG	ATT	TTG	ATG	TTT	ATG	CTG	TTG	AGC	CAG	GTC	ATC	CAT	TGC	TTT	881
Lys	Ile	Leu	Met	Phe	Met	Leu	Leu	Ser	Gln	Val	Ile	His	Cys	Phe	
CCT	GAC	TTT	TTC	TTG	GAC	GGC	CAG	CCT	GAG	GAT	CTA	GTC	AAG	CAT	926
Pro	Asp	Phe	Phe	Leu	Asp	Gly	Gln	Pro	Glu	Asp	Leu	Val	Lys	His	
ATG	GAG	GCA	CAT	GGT	GCC	ACT	CCG	GCG	TTC	ACA	ACT	GCA	AAA	GCA	971
Met	Glu	Ala	His	Gly	Ala	Thr	Pro	Ala	Phe	Thr	Thr	Ala	Lys	Ala	
GAT	CCC	GTT	GCC	GCA	GGA	CCA	GTT	TCT	GAG	ATG	TTC	AAT	ACA	ATC	1016
Asp	Pro	Val	Ala	Ala	Gly	Pro	Val	Ser	Glu	Met	Phe	Asn	Thr	Ile	
AGA	GGA	ATT	ATC	AGT	CCA	GAG	ATG	GTG	AAA	ACC	ACA	CAA	GGA	GTG	1061
Arg	Gly	Ile	Ile	Ser	Pro	Glu	Met	Val	Lys	Thr	Thr	Gln	Gly	Val	
TAC	AAA	TTT	AAC	TTA	GCA	GGC	GAG	CAT	GCT	GGA	GTC	TGG	TAT	CTT	1106
Tyr	Lys	Phe	Asn	Leu	Ala	Gly	Glu	His	Ala	Gly	Val	Trp	Tyr	Leu	
GAC	CTG	AAG	AAC	GAT	GCT	GGA	AGT	GCT	GGA	AAT	GGG	GAA	CCT	CCT	1151
Asp	Leu	Lys	Asn	Asp	Ala	Gly	Ser	Ala	Gly	Asn	Gly	Glu	Pro	Pro	
GTC	AAA	GCT	GAT	GTT	GTC	ATG	TCG	ATG	GAC	AGT	GAG	GAT	TTT	GTC	1196
Val	Lys	Ala	Asp	Val	Val	Met	Ser	Met	Asp	Ser	Glu	Asp	Phe	Val	
AAG	ATG	TTT	GGA	GGG	AAA	TTA	AAG	CCA	ACC	ATG	GCC	TTC	ATG	TCT	1241
Lys	Met	Phe	Gly	Gly	Lys	Leu	Lys	Pro	Thr	Met	Ala	Phe	Met	Ser	
GGA	AAG	CTG	ACT	ATT	AAG	GGT	GAC	ATG	GCC	CTT	GTC	ATC	AAA	CTG	1286
Gly	Lys	Leu	Thr	Ile	Lys	Gly	Asp	Met	Ala	Leu	Val	Ile	Lys	Leu	
GAG	AAG	ATG	ATG	GCC	ATG	ATG	AAG	TCT	AAA	CTG	TGA	actcatc	cacgca	1334	
Glu	Lys	Met	Met	Ala	Met	Met	Lys	Ser	Lys	Leu	***				

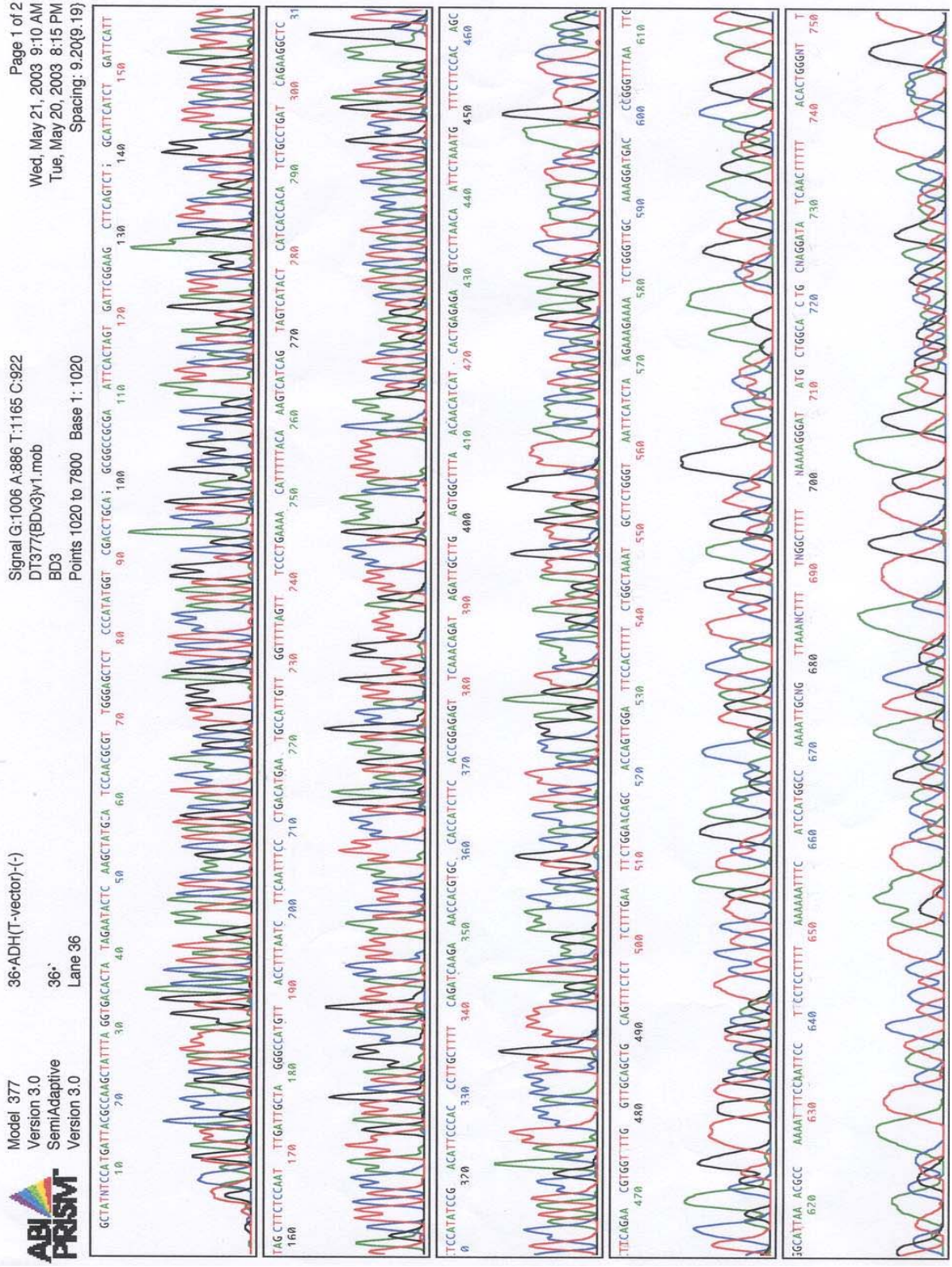
Gtatacctcccttacaacatgtacatcatgcacctcttttctatatcagtggttttttgattaaactaattgttcagatt  
gtactaatttttaataaagaactatcaggcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1469

圖十四：Z-HSD之開放閱讀框架。序列72~1319為蛋白質編碼區，ORF為1469bp，具有416個推定胺基酸。



圖十五：H-HSD 全長 DNA 定序圖 I。





圖十六：H-HSD 全長 DNA 定序圖 II。

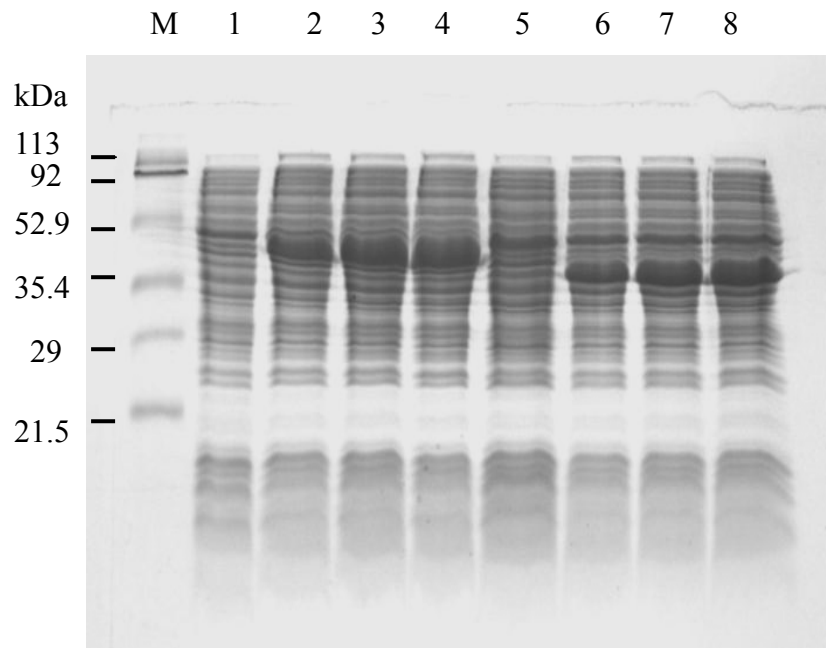
5'-UTR : gataaaatgcgaggagcggtccagctttagctctctgctcgccgcctatttacgctccctgccaggtcgaaa  
tcgagagacgagcggcgccgctgtgcccacctcctctgatctacgaaagtcATG TTA  
Met Leu 135

CCC AAC ACC GGG AGG CTG GCA GGA TGT ACA GTT TTT TTT ATC ACA GGT 180  
Pro Asn Thr Gly Arg Leu Ala Gly Cys Thr Val Phe Ile Thr Gly  
GCA AGC CGT GGC ATT GGC AAA GCT ATT GCA TTG AAA GCA GCA AAG 225  
Ala Ser Arg Gly Ile Gly Lys Ala Ile Ala Leu Lys Ala Ala Lys  
GAT GGA GCA AAT ATT GTT ATT GCT GCA AAG ACC GCC CAG CCA CAT 270  
Asp Gly Ala Asn Ile Val Ile GCT Ala Lys Thr Ala Gln Pro His  
CCA AAA CTT CTA GGC ACA ATC TAT ACT GCT GCT GAA GAA ATT GAA 315  
Pro Lys Leu Leu Gly Thr Ile Tyr Thr Ala Ala Glu Glu Ile Glu  
GCA GTT GGA GGA AAG GCC TTG CCA TGT TTT GAT GAT GTG AGA GAT 360  
Ala Val Gly Gly Lys Ala Val Cys Pro Cys Ile Val Asp Val Arg Asp  
GAA CAG CAG ATC AGT GCT GCA GTG GAG AAA GCC ATC AAG AAA TTT 405  
Glu Gln Gln Ile Ser Ala Ala Val Glu Lys Ala Ile Lys Lys Phe  
GGA GCT TAT ACC ATT GCT AAG TAT GGT TTT ATG TCT ATG TAT GTG CTT 450  
Gly Ala Tyr Thr Ile GCT AAG Tyr Gly Met Ser Met Tyr Val Leu  
GGA ATG GCA GAA GAA TTT AAA GGT GAA ATT GCA GTC AAT GCA TTA 495  
Gly Met Ala Glu Glu Phe Lys Gly Glu Ile Ala Val Asn Ala Leu  
TGG CCT AAA ACA GCC ATA CAC ACT GCT Ala GCT ATG Met GAT ATG CTG GGA 540  
Trp Pro Lys Thr Thr Ile His Thr Ala Thr Met Asp Met Leu Gly  
GGA CCT GGT ATC GAA AGC CAG TGT AGA AAA GTT GAT ATC ATT GCA 585  
Gly Pro Gly Ile Glu Ser Gln Cys Arg Lys Val Asp Ile Ile Ala  
GAT GCA GCA TAT TCC ATT TTC CAA AAG CCA AAA AGT TTT ACT GGC 630  
Asp Ala Ala Tyr Thr Ile Phe Gln Lys Pro Lys Ser Phe Thr Gly  
AAC TTT GTC ATT GAT GAA AAT ATC TTA AAA GAA GAA GGA ATA GAA 675  
Asn Phe Val Ile Asp Glu Asn Ile Leu Lys Glu Glu Gly Ile Glu  
AAT TTT GAC GTT TAT GCA ATT AAA CCA GGT CAT CCT TTG CAA CCA 720  
Asn Phe Asp Val Tyr Ala Ile Lys Pro Gly His Pro Leu Gln Pro  
GAT TTC TTC TTA GAT GAA TAC CCA GAA GCA GTT AGC AAG AAA GTG 765  
Asp Phe Phe Leu Asp Glu Tyr Pro Glu Ala Val Ser Lys Lys Val  
GAA TCA ACT GGT GCT GTT CCA GAA TTC AAA GAA GAG AAA CTG CAG 810  
Glu Ser Thr Gly Ala Val Pro Glu Phe Lys Glu Glu Lys Leu Gln  
CTG CAA CCA AAA CCA CGT TCT GGA GCT GTG GAA GAA ACA TTT AGA 855  
Leu Gln Pro Lys Pro Arg Ser Gly Ala Val Glu Glu Thr Phe Arg  
ATT GTT AAG GAC Ser TCT CTC AGT GAT GAT GTT GTT AAA GCC ACT CAA 900  
Ile Val Lys Asp Ser Leu Ser Asp Val Val Lys Ala Thr Gln  
GCA ATC TAT CTG TTT GAA CTC TCC GGT GAA GAT GGT GGC ACG TGG 945  
Ala Ile Tyr Leu Phe Glu Leu Ser Gly Glu Asp Gly Gly Thr Trp  
THT CTT GAT CTG AAA AGC AAG GGT GGG AAT GTC GGA TAT GGA GAG 990  
Leu Leu Asp Leu Lys Ser Lys Gly Asn Val Gly Tyr Gly Glu  
CCT TCT GAT CAG GCA GAT GTG GTG ATG AGT ATG ACT ACT GAT GAC 1035  
Pro Ser Asp Gln Ala Asp Val Val Met Ser Met Thr Thr Asp Asp  
THT GTA AAA ATG TTT TCA GGG AAA CTA AAA CCA ACA ATG GCA TTC 1080  
Phe Val Lys Met Phe Ser Lys Lys Leu Lys Pro Thr Met Ala Phe  
ATG TCA GGG AAA TTG AAG ATT AAA GGT AAC ATG GCC CTA GCA ATC 1125  
Met Ser Gly Lys Leu Lys Ile Lys Gly Asn Met Ala Leu Ala Ile  
AAA TTG GAG AAG CTA ATG AAT CAG ATG AAT GCC AGA CTG TGA AGG 1170  
Lys Leu Glu Lys Leu Met Asn Gln Met Asn Ala Arg Leu \*\*\*

aaaaataaaaaaaaaagtcgactgctatgctcaaaaagtaaaaaagctcaacagttaaaaatctaagtgtttg  
tctttcctgttatattataaggatgacgctttgttctggaaaaagatagaatttgtctctaaaaagacttga  
tgaattaaaaatggcaagctaatcaaacataaagcttcaatfaagtgggattcctaaagacagctctgtgt  
tcaaggggttaaacctttgagccttacatctcactgtctttctccaagaaaagtaattttggcgggacagc  
agatcaagcagtaaaaattagctctttcaaatcttctgtcatgtaaaaatgaagctagctgtttttaa  
gttttggattgtataactaaagaaaatcctaatgatgttttattttataacttatttfaagaaaatctt  
atagtacattttcaaaaaatfataaaaaatgaattagtactggcgaggactaaatgaaaacaataattttt  
tgaactagctttccaggtggacttagccatagaaaaatattactaatgtaattfaacaaattgctgcatg  
tccattttaaataatgttttaattgtcctaaaaaacaataaattttctccctaggagatgcatttggctac  
tttgaaacagaaaaccttaagaataggtcattggtatgggctgaactgtgtatcccccaattcatttgg  
ctaaactcccatttcttttgaatgactgttcggagatgaggcctttaaagaggtgacttaagttcaaagg  
tgttagtctaatacaacatgggtgctctttggacataaagagataccagcaatggtgacacagaacaag  
gagacacagtgagaaggcagttatctgcaagcaagagagaggcttcagaagaaaacaaaatcaccagca  
atctttgacttctaatctccagaatagtgagaaaataaatttctgttggtaaagccgtccactgtggg  
caggaggattgcttgaggccaggagttcaaggccagcctggacaacatagtaagaccctatctctacc  
taaaatfaattttaaagcccccaatctgtggtatfataatggcagccctagcaagctaaatacagtggt  
gaggtgggagggttgagggaagataaaacttttaaaaaagctcttattcttccattcaatcagttaaa  
gctcagtgtaacaattttgcttccagcttccactcctaattattgttggccattaaagcaatttagctaa  
catttcttagattcataatggttaggagcatttaatctgtattttacaagttaggaagcagaggatcag  
aaaggactagcccaagggccaacatfaacaagcctctaaacaaaaacttacaatcattttagttgaatg  
ccaagatctccactccatccaggaatggagcctcatgtaatacaagtgaaacttaaaaaataggacag  
agtcaggagattcacagcaactgatcaaaagggagtcagctcaacgtgagcaagcgtgattatgatgag  
cctctgttaataccacaaggaacgtaacctgaagtaacctgatgttaaccatctgctgtgtctactatg  
gttctctgttctgctagtgtgctttcaaaaatgcagaccattctatcacctggcagggctctctgttt  
ttgtaggctggatgctaccaggttcctgtaatcgctaaataaaagccaatttagatcttfaaaaaaa  
aaa

3003

圖十七：H-HSD 之開放閱讀框架。序列 130~1164 為蛋白質編碼區，ORF 為 3003bp，具有 345 個推定胺基酸。



圖十八：過量表現新穎之 H-HSD 及 Z-HSD 蛋白質的 SDS-PAGE 分

析。pQE9/H-HSD 與 pQE31/Z-HSD 質體轉形至 E coli. M15 strain 中，37°C 培養，分別在 RT(室溫)、30°C 和 37°C，以 IPTG(200µg/ml)誘發蛋白質表現，取其菌液以 12.5% SDS-PAGE 分析。

Lane M：low range protein marker (Bio-Rad)。

Lane 1、5：Z-HSD 和 H-HSD 沒有誘發表現蛋白質。

Lane 2：斑馬魚 HSD 在室溫下之過量表現蛋白質。

Lane 3：斑馬魚 HSD 在 30°C 下之過量表現蛋白質。

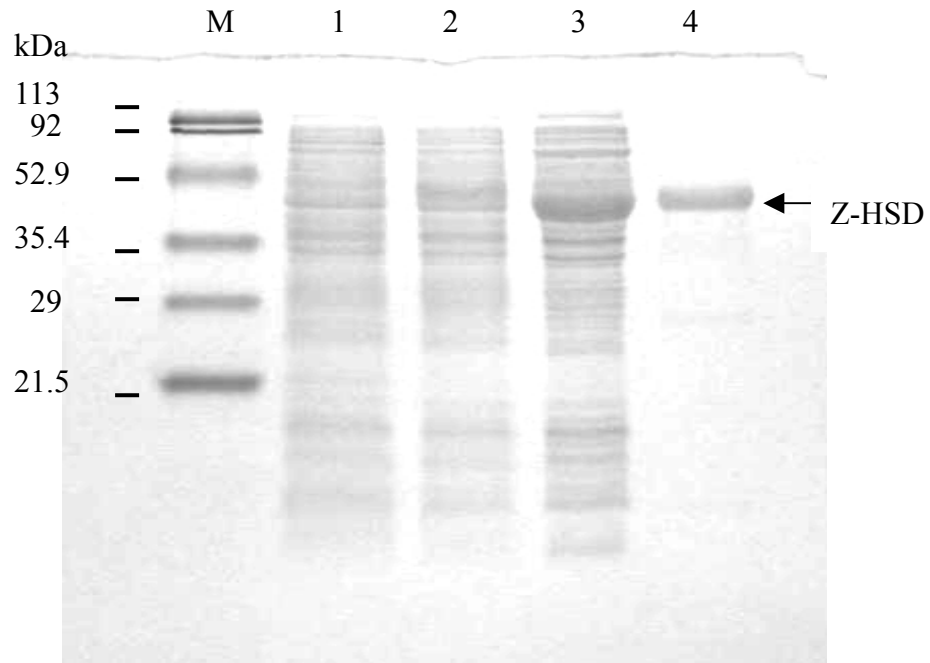
Lane 4：斑馬魚 HSD 在 37°C 下之過量表現蛋白質。

Lane 6：人類 HSD 在室溫下之過量表現蛋白質。

Lane 7：人類 HSD 在 30°C 下之過量表現蛋白質。

Lane 8：人類 HSD 在 37°C 下之過量表現蛋白質。





圖十九：新穎之 Z-HSD 純化蛋白質的 SDS-PAGE 分析。pQE31/Z-HSD

質體轉形至 *E. coli. M15 strain* 中，37°C 隔夜培養，在 30°C 以 IPTG(200µg/ml) 誘發蛋白質表現，經 Ni-NTA 管柱純化，以 12.5% SDS-PAGE 分析。

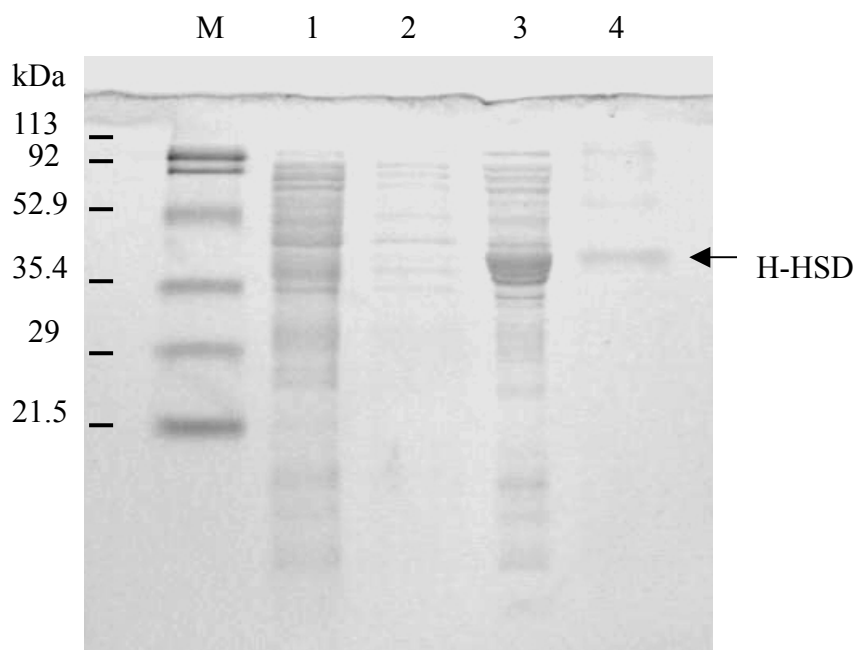
Lane M：low range protein marker (Bio-Rad)。

Lane 1：M15 strain 無轉形 pQE31/Z-HSD 質體。

Lane 2：pQE31/Z-HSD 質體轉形至 *E. coli. M15 strain*，無以 IPTG 誘發。

Lane 3：pQE31/Z-HSD 質體轉形至 *E. coli. M15 strain*，以 IPTG 誘發。

Lane 4：經 Ni-NTA 管柱純化之蛋白質，箭號處為標的蛋白質 Z-HSD。



圖二十：新穎之 H-HSD 純化蛋白質 SDS-PAGE 分析。pQE9/H-HSD 質

體轉形至 *E. coli. M15 strain* 中，37°C 隔夜培養，於 30°C 以 IPTG(200µg/ml)誘發蛋白質表現，經 Ni-NTA 管柱純化，以 12.5% SDS-PAGE 分析。

Lane M：low range protein marker (Bio-Rad)。

Lane 1：M15 strain 無轉形 pQE9/H-HSD 質體。

Lane 2：pQE9/H-HSD 質體轉形至 *E. coli. M15 strain*，無以 IPTG 誘發。

Lane 3：pQE9/H-HSD 質體轉形至 *E. coli. M15 strain*，以 IPTG 誘發表。

Lane 4：經 Ni-NTA 管柱純化之蛋白質，箭號處為標的蛋白質 H-HSD。

表一：Z-HSD 及 H-HSD 放大及序列分析所使用的引子。

Technique	Primer sequences 5' to 3'
<b>3' RACE PCR</b>	oligo dT-anchor primer :
	GACCACGCGTATCGATGTCGACT <sub>16</sub> V
	<i>Mlu</i> I <i>Cla</i> I <i>Sal</i> I V= A, C or G
	First PCR(3-1) : CTCAGGGGAACATACCTGACGTCTAAA
	Next PCR(3-2) : CACACGGCTTACACCATTGCAAGTAT
	<b>Zebrafish</b>
<b>Full length PCR</b>	Forward : AATGTCGACATGCTGCAGAACACAGGGAA
	<i>Sal</i> I
	Reverse : CGCAAGCTTCAGTTTAGACTTCATCATG
	<i>Hind</i> III
	<b>Human</b>
	Forward : AATGGATCCATGTTACCCAACACCGGGAG
	<i>Bam</i> HI
	Reverse : CGGAAGCTTCAGTCTGGCATTTCATCTGATT
	<i>Hind</i> III
<b>Sequencing</b>	T7 : TAATACGACTCACTATAGGG
	SP6 : ATTTAGGTGACACTATAGAA
	PQE S : CGGATAACAATTCACACAG
	PQE AS : GTTCTGAGGTCATTACTGG

註：Z-HSD 延長 3' 端序列(3' RACE PCR)使用 Clontech 之 Marathon kit 提供之 oligo dT primer 及自行設計之二段 3' 端引子(3-1/3-2)來進行 PCR；H-HSD 及 Z-HSD 全長序列放大，分別設計正反二股(含內切酵素位點)引子進行放大作用；核酸序列分析使用 Z-HSD 及 H-HSD，各別所轉殖之載體上具有之 promoter primer(pGEM-T Vector :T7/SP6；pQE 9/31: pQES/pQEAS)。

表二：新穎之 Z-HSD 及 H-HSD 蛋白質純化表 I。

	Fraction	Protein (mg)	Specific Activity (unit/mg)	Recover of activity (%)
Zebrafish HSD (Estriol/NAD <sup>+</sup> )*	supernatant	48.05	0.13	100
	Ni-column purify	1.73	3.16	42.9
Zebrafish HSD (Estriol/NADP <sup>+</sup> )*	supernatant	48.05	0.17	100
	Ni-column purify	1.73	3.71	33.79
Human HSD (Estriol/NAD <sup>+</sup> )	supernatant	28.56	0.19	100
	Ni-column purify	0.115	10.74	22.38
Human HSD (Estriol/NADP <sup>+</sup> )	supernatant	28.56	0.18	100
	Ni-column purify	0.115	10.73	12.59

※ \*：表示 Estriol 所使用之濃度為 1mM；NAD<sup>+</sup>及 NADP<sup>+</sup>所使用之濃度為 1.5mM。

※表中 Ni-column purify 之 Recovery of activity 係以 supernatant 作為 100%計算。

表三：新穎之 Z-HSD 及 H-HSD 蛋白質純化表 II。

	Fraction	Protein (mg)	Specific Activity (unit/mg)	Recover of activity (%)
Zebrafish HSD (5-Androstene /NAD <sup>+</sup> )*	supernatant Ni-column purify	48.05 1.73	0.07 1.75	100 95.58
Zebrafish HSD (5-Androstene /NADP <sup>+</sup> )*	supernatant Ni-column purify	48.05 1.73	0.06 1.41	100 80.53
Human HSD 5-Androstene /NAD <sup>+</sup> )	supernatant Ni-column purify	28.56 0.115	0.30 12.81	100 16.99
Human HSD (5-Androstene /NADP <sup>+</sup> )	supernatant Ni-column purify	28.56 0.115	0.21 13.62	100 26.30

※ \*：表示 5-Androstene(5-Androstene-3 $\beta$ ,16 $\alpha$ ,17 $\beta$ -triol) 所使用之濃度為 0.1mM；NAD<sup>+</sup>及 NADP<sup>+</sup>所使用之濃度為 0.2mM。

※ 表中 Ni-column purify 之 Recovery of activity 係以 supernatant 作為 100%計算。

## 柒、附錄

AI477544

fb58f08.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone  
IMAGE:3716103 5' similar to WP:C45B11.3 CE05427 ALCOHOL  
DEHYDROGENASE ;, MRNA sequence  
gi|4335155|gb|AI477544.1|[4335155]

source : zebrafish

Journal : unpublished (1998)

Sex : mixed

tissue\_type : 26 somite embryos, adult livers, shield stage embryos

### SEQUENCE

```
TTTAATGTGTCAACCAGGAGCCCCTCTGGATCATTGCGTTCATATCAGACATCATCCAG  
GAGACACAATAATGCTGCAGAACACAGGGAAGCTGGCAGGATGCACCATTTTCATAACCG  
GAGCAAGTCGAGGTATTGGCAAGGCCATTGCTCTGAAAGCTGCGCAGGATGGAGCCAATG  
TAGTCATCGCTGCCAAAACAGCTGATCCACATCCCAAACCTCCCCGGCACCATCTACACAG  
CTGCAGCAGAGATTGAAGCAGCTGGAGGGAAAGCACTGCCGTGTATTGTGGACGTCCGTG  
ATGAGAAGCAGATCAATGATGCTGTTGAACAGGCTGTGGAGAAGTTTGGAGGGATTGACA  
TATTGGTCAACAATGCCAGTGCCATCAATTTAACAGGGACTCTTCAGACTCCAATGAAAA  
AGGCTGACCTCATGCTGGGCATCAATCTCAGGGGAACATACCTGACGTCTAAACTGTGCA  
TTCCACATCTTCTGAAGAGCAAAAACCCTCACATACTAAACCTCAGCCCACCTCTCAACC  
TCAACCCCATTTGGTTTAAAAACCACACGGCTTACACCATTTGCAAGTATGGCATGTNCA  
TGTGTGTACTGGGAATGGCAGAAGAGTTCAGA
```

Quality: High quality sequence stops at base: 480

(摘錄自 NCBI 網站)

附錄 1：Z-HSD(編號 AI477544)之資料。NCBI 資料庫發佈之斑馬魚(編號 AI477544)片段序列之參考核酸序列(reference sequences)，缺乏 3'端之聚腺嘌呤(poly A tail)，初步定義其 5'端序列相似於 alcohol dehydrogenase，其 cDNA clone 來源為斑馬魚胚胎之 26 體節(26 somite embryos)、成魚肝臟和胚胎盾狀期(shield stage embryos)。

NP\_115679

Hypothetical protein MGC10940 [Homo sapiens] gi |14150062| ref  
|NP\_115679.1 [14150062]

source organism : Homo sapiens

chromosome : 9

map : 9q32

tissue\_type : Uterus, endometrium adenocarcinoma

#### human nucleotide (ORF)

```
1 atgttaccca acaccgggag gctggcagga tgtacagttt ttatcacagg tgcaagccgt
61 ggcattggca aagctattgc attgaaagca gcaaaggatg gagcaaatat tgttattgct
121 gcaaagaccg cccagccaca tccaaaactt ctaggcacaa tctatactgc tgctgaagaa
181 attgaagcag ttggaggaaa ggcccttgcca tgtattgttg atgtgagaga tgaacagcag
241 atcagtgtcg cagtggagaa agccatcaag aaatttggag cttataccat tgctaagtat
301 ggtatgtcta tgtatgtgct tggaatggca gaagaattta aagtgaaat tgcaagtcaat
361 gcattatggc ctaaaacagc catacacact gctgctatgg atatgctggg aggacctggt
421 atcgaaagcc agtgtagaaa agttgatatc attgcagatg cagcatatc cattttccaa
481 aagccaaaaa gttttactgg caactttgtc attgatgaaa atatcttaaa agaagaagga
541 atagaaaatt ttgacgttta tgcaattaaa ccaggctatc ctttgcaacc agatttcttc
601 ttagatgaat acccagaagc agttagcaag aaagtggaat caactgggtg tgttccagaa
661 ttcaagaag agaaactgca gctgcaacca aaaccacgtt ctggagctgt ggaagaaaca
721 tttagaattg ttaaggactc tctcagtgat gatgtttgta aagccactca agcaatctat
781 ctgtttgaac tctccggtga agatggtggc acgtggtttc ttgatctgaa aagcaagggt
841 gggaatgtcg gatatggaga gccttctgat caggcagatg tggatgatgag tatgactact
901 gatgactttg taaaaatggt ttcagggaaa ctaaaaccaa caatggcatt catgtcaggg
961 aaattgaaga ttaaaggtaa catggcccta gcaatcaaat tggagaagct aatgaatcag
1021 atgaatgcca gactgtga
```

#### human amino acid

```
1 mlpntgrlag ctvfitgasr gigkaialka akdganivia aktaqphpkl lgtiytaee
61 ieavggkalp civdvrdeeq isaavekaik kfgaytiaky gmsmyvlgma eefkgeiavn
121 alwpktaiht aamdmlggpg iesqcrkvdi iadaaysifq kpkstgnfv idenilkeeg
181 ienfdvyaik pghplqpdf ldeypeavsk kvestgavpe fkeeklqlqp kprsgaveet
241 frivkdsisd dvvkatqaiy lfelsgedgg twfldlkskg gnvgygepsd qadvmsmtt
301 ddfvkmfsgk lkptmafmsg klkikgnmal aikleklmnq mnarl
```

(摘錄自 NCBI 網站)

附錄 2：H-HSD(編號 NP\_115679)之資料。NCBI 資料庫公佈之 H-HSD 之參考序列(reference sequences)及對其之初步定義為 hypothetical protein，並說明此人類之基因位於 9 號染色體上 q 臂之 32 位置，組織來源為子宮及子

宮内膜癌。



```

          *           20           *           40           *           6
z-ADH3   : -MDTTGKVIKCKAAVAWEA---GKPLTTEVEVEVAPPKAHEVR-VKIH--ATGVCHTDAY : 52
zebrafish : MLQNTGKLAGCTIFITGASRGI GKATAAKAAQDGANVVIAAKTADPHPKLPGTIYTAAA : 59
          6  TKG6  C   6           GK 6 6  2           4   H   G   T   A

          0           *           80           *           100           *           1
z-ADH3   : TLGSDPEGLFPVILGHEGAGTVEVSGEGVTKFKPGDVIPLYVPPQCGECKFCCKNPKTN : 111
zebrafish : ELEAAGGKALPCIVDVRDEKQINDAVBQAVEKEKGGIDILVNN-ASAINLTGTLQTP-M- : 115
          6           L   66           V 2 V KF   D 66           P

          20           *           140           *           160           *
z-ADH3   : LCQKIRVVTQGGQLMPDNTSRFTCKGKQLRHFEMGTSTFSEYTVVAEISLAKV--DEHAPL : 168
zebrafish : --EKADIMLGINLRGTYLTSKLCIP---HLLKSKNPHILNLSPPNLHPTWFKNHTGL : 168
          2K 6  G  L           3  C           H 6 3           6 6  L 6  H  L

          180           *           200           *           220           *
z-ADH3   : D--KVCLLGGISTGYGAINTAKVEAGSTCAVEFG-----LGAVG-LAVVMGCKSAGAT : 219
zebrafish : HHCKVWHVQCVTWEWQKSSQDPLPLMPYGORQPFRRLPWTCMAVLR LASSAGRRRSWLM : 227
          KV 6  C           1 6           F           AV  LA  G

          240           *           260           *           280           *
z-ADH3   : RIIGIDVNP---DKFELAKKEG---ATEFVNPKDHSKPTQEVLVVELTDG-GVDYS--F : 268
zebrafish : QHTPSSNNEPASPDSLLMRTESKRRALKILMFMLLSQVTHCFPDFFLDQQPEDLVKHM : 286
          NP  D 6 4 F  A 6  S  I           DG  D

          240           *           260           *           280           *
z-ADH3   : RIIGIDVNP---DKFELAKKEG---ATEFVNPKDHSKPTQEVLVVELTDG-GVDYS--F : 268
zebrafish : QHTPSSNNEPASPDSLLMRTESKRRALKILMFMLLSQVTHCFPDFFLDQQPEDLVKHM : 286
          NP  D 6 4 F  A 6  S  I           DG  D

          300           *           320           *           340           *
z-ADH3   : ECIGNVG-IMRAALEACHKCGWTSVIIGVACAGQEISTRPFQLVTGRTWKCTAFCGM-- : 324
zebrafish : EAHGATPAFTTAKADPVAAGPVSEMFNTIRGIISPENVTTQGVYKFNLAGEHAGVMYL : 345
          E  G           A           G 3 6 6  G           4  Q  V           G  G W

          360           *           380           *           400           *
z-ADH3   : ---KSVEV---PKLVNDY---MNKMLVDEE---VHTLPEA--QINEAFDIMHAG- : 367
zebrafish : DLKNDAGSAGNGEPPVKADVMSMDSSEDFVKMEGGKLPKPTMAFMMSGKLTIKGD MALVIK : 404
          S   P 6  D  M1  V  F 6  T6  F 6  D6

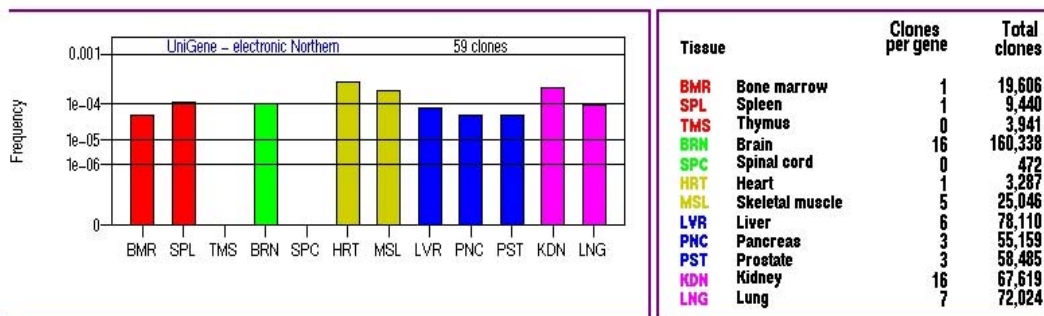
          420
z-ADH3   : -KSIRAVLQF-- : 376
zebrafish : LEKMMAMMKSKL : 416
          6 A66

```

附錄 3：Z-HSD 與 Z-ADH3 之胺基酸比對。Z-HSD 之全長胺基酸序列與 Asok K. Dasmahapatra 等人於 2001 年發表之斑馬魚 ADH3(alcohol dehydrogenase 3)胺基酸序列(accession number AF399909)比對，發現並無有義意之相似性存在。黑色區塊表示胺基酸序列相同。

MGC10940 expression in normal human tissues according to [GeneNote/GeneAnnot](#)

Affymetrix probe-set array sensitivity specificity # of associated genes



Transcripts/Translation Summary

**NM\_032303** Stable ID: ENST00000262542 Exons: 11 Transcript length: 3345 bp Translation length: 418 residues  
[\[View transcript information\]](#) [\[View exon information\]](#) [\[View protein information\]](#)

**Similarity Matches**  
 This Ensembl entry corresponds to the following database identifiers:  
 Affymx Microarray U133: [HG-U133A:209512\\_at](#) [HG-U133A:209513\\_s\\_at](#)  
 Affymx Microarray U95: [HG-U95A:34838\\_at](#) [HG-U95Av2:34838\\_at](#)  
 EMBL: [BC004331](#) [\[align\]](#)  
 LocusLink: [84263](#) [\[align\]](#)  
 Protein ID: [AAH04331](#) [\[align\]](#)  
 RefSeq: [NM\\_032303](#) [Target %id: 100; Query %id: 82] [\[align\]](#)  
 SpTrEMBL: [Q9BT58](#) [Target %id: 100; Query %id: 82] [\[align\]](#) [\[Search GO\]](#)

**GO**  
 The following GO terms have been mapped to this entry via Swissprot/SpTrEMBL:  
[GO:0005498](#) [\[sterol carrier\]](#)  
[GO:0008152](#) [\[metabolism\]](#)  
[GO:0016491](#) [\[oxidoreductase\]](#)

**InterPro**  
[IPR002198](#) Short-chain dehydrogenase/reductase SDR - [\[View other Ensembl genes with this domain\]](#)  
[IPR003033](#) Sterol-binding - [\[View other Ensembl genes with this domain\]](#)

**Protein Family**  
[ENSF0000003053](#) : AMBIGUOUS  
 This cluster contains 1 Ensembl gene member(s)

**Transcript Structure**

**Protein structure**

(摘錄自 GeneCards 網站)

附錄 4：H-HSD 在 GeneCards 網站搜尋資訊。由此網站中的資料顯示此人類 (NP\_115679) 基因之組織分布 (Electronic Northern) 與轉錄/轉譯資訊。

human (1) -----

zebrafish (1) TTTAATGTGTCACCAGGAGCCCTCTGGATCATTGCGTTCCCATATCAGA

human (1) -----ATGTTACCCAACACCGGAGGCTGGCAGG

zebrafish (51) CATCATCCAGGAGACACAATAATGCTGCAGAACACAGGGAAGCTGGCAGG

human (30) ATGTACAGTTTTTATCACAGGTGCAAGCCGTGGCATTGGCAAAGCTATTG

zebrafish (101) ATGCACCATTTTTCATAACGGAGCAAGTCGAGGTATTGGCAAAGCCATTG

human (80) CATTGAAAGCAGCAAAGGATGGAGCAATATTGTTATTGCTGCAAAGACC

zebrafish (151) CTCTGAAAGCTGCGCAGGATGGAGCCAATGTAGTCATCGCTGCCAAACA

human (130) GCCAGCCACATCCAAACTTCTAGGCACAATCTATACTGCTGCTGAAGA

zebrafish (201) GCTGATCCACATCCAAACTCCCGGCACCATCTACACAGCTGCAGCAGA

human (180) AATTGAAGCAGTTGGAGGAAAGGCCTTGCCATGTATTGTTGATGTGAGAG

zebrafish (251) GATTGAAGCAGCTGGAGGAAAGCAC TGCCGTGTATTGTGGAAGTCCGTG

human (230) ATGAACAGCAGATCAGTGCTGCAGTGGAGAAAGCCATCAAGAAATTTGGA

zebrafish (301) ATGAGAAGCAGATCAATGATGCTGTTGACAGGCTGTGGAGAAATTTGGA

human (280) GCTTATACCATTGCTAAGTATGGTATGTCTATGTATGTGCTTGGAAATGGC

zebrafish (351) GGG-----ATTGACA--TATTGGTCAACAATGCCAGTGCC-----

human (330) AGAAGAAATTTAAAGTGAAATTGCAAGTCAATGCAATTATGGCCTAAAACAG

zebrafish (384) --ATCAATTTAAAGGGACTCTTCAGAC--TCCA--ATG-----AAAAG

human (380) CCATACACACTGCTGCT--ATGGATATGCTGGGAGGA--CCTGGTATCGAA

zebrafish (423) GCTGACCTCATGCTGGGCATCAATCTCAGGGGAACATACCTGACGCTCTAA

human (427) AGCCAGTGTAGAAAAGTTGATATCAT--TGCAGATGCAGCATATTCCATTT

zebrafish (473) A--CTGTGCAT-----TCCACATCTTCTGAAGA--GCAAAAACCCTCACAT

human (476) TCCAAAAGCCAAAAGTTTTACTGGCAACTTTGTCAATTGATGAAAATATC

zebrafish (515) ACTAAACCTCA----GCCACCTCTCAACCTCAACCCCATTTGGTTTA--

human (526) TTAAAAAGAAAGGAAAGAAAATTTGACGTTTATGCAATTAACCCAGG

zebrafish (559) --AAAACCACAGGCTTACACCATTTGCAAGT--ATGGCATGTN--CATG

human (576) TCATCCTTTGCAAC--CAGATTTCTTCTTAGATGAATACCCAGAAGCAGT

zebrafish (603) TGTGTACTGGAAATGGCAGAGAGTTCAGA-----

human (624) TAGCAAGAAAGTGAATCAACTGGTGTGTTCCAGAATTCAAAGAAGAGA

human (674) AACTGCAGCTGCAACCAAAACCACGTTCTGGAGCTGTGGAAGAAACATTT

human (724) AGAATTGTTAAGGACTCTCTCAGTGATGATGTTGTTAAAGCCACTCAAGC

human (774) AATCTATCTGTTTGAACCTCCGGTGAAGATGGTGGCACGTGGTTTCTTG

human (824) ATCTGAAAAGCAAGGGTGGGAATGTCGGATATGGAGAGCCTTCTGATCAG

human (874) GCAGATGTGGTGTGAGTATGACTACTGATGACTTTGAAAAATGTTTTC

human (924) AGGGAAACTAAAACCAACAATGGCATTGATGTCAGGGAAATTGAAGATTA

human (974) AAGGTAACATGGCCCTAGCAATCAAATTGGAGAAGCTAATGAATCAGATG

human (1024) AATGCCAGACTGTGA

附錄 5：H-HSD 與 Z-HSD 之核酸序列比對。框格處為 H-HSD 之起始和終止密碼子。

UniGene Cluster Hs.47986 *Homo sapiens*

### MGC10940 Hypothetical protein MGC10940

SEE ALSO: [LocusLink](#) | [HomoloGene](#)

#### SELECTED MODEL ORGANISM PROTEIN SIMILARITIES

organism, protein and percent identity and length of aligned region

- H.sapiens*: [pir:S59136](#) - S59136 estradiol 17beta-dehydrogenase 27 % / 167 aa  
(EC 1.1.1.62) type 4 - human (see [ProtEST](#))
- M.musculus*: [sp:P51660](#) - DHB4\_MOUSE Estradiol 17 25 % / 267 aa  
beta-dehydrogenase 4 (17-beta-HSD 4) (see [ProtEST](#))  
(17-beta-hydroxysteroid dehydrogenase 4)
- R.norvegicus*: [sp:P97852](#) - DHB4\_RAT Estradiol 17 24 % / 267 aa  
beta-dehydrogenase 4 (17-beta-HSD 4) (see [ProtEST](#))  
(17-beta-hydroxysteroid dehydrogenase 4)
- C.elegans*: [pir:T34105](#) - T34105 hypothetical protein C17G10.8 - 45 % / 247 aa  
Caenorhabditis elegans (see [ProtEST](#))

附錄 6-1：H-HSD 在 Uni-Gene 查尋所得之資訊與訊息。H-HSD，分別與人類 (*H.sapiens*)、大鼠 (*R.norvegicus*)、小鼠 (*M.musculus*)和線蟲 (*C.elegans*) 等具有同源相似。

NCBI **LocusLink**

PubMed Entrez BLAST OMIM Map Viewer Taxonomy Structure

Search LocusLink Display Brief Organism: All

Query:  Go Clear

View Hs MGC10940 One of 1 Loci Save All Loci

ABCDEFGHIJKLMN OPQRSTUVWXYZ

Click to Display mRNA-Genomic Alignments (spanning 92467 bps)

ACEVIEW UNIGENE MAP VAR HOMOL e! UCSC

**Homo sapiens Official Gene**

Symbol and Name

None Available

**Interim Gene Symbol and Name:**  
MGC10940: hypothetical protein MGC10940

**LocusID:** 84263

Overview ?

**Locus** gene with protein product, function  
**Type:** unknown  
**Product:** hypothetical protein MGC10940

Function [Submit GeneRIF](#) ?

**Gene Ontology™:**

Term	Evidence	Source	Pub
• <a href="#">sterol carrier activity</a>	IEA	<a href="#">GOA</a>	
• <a href="#">metabolism</a>	IEA	<a href="#">GOA</a>	
• <a href="#">oxidoreductase activity</a>	IEA	<a href="#">GOA</a>	

Map Information ?

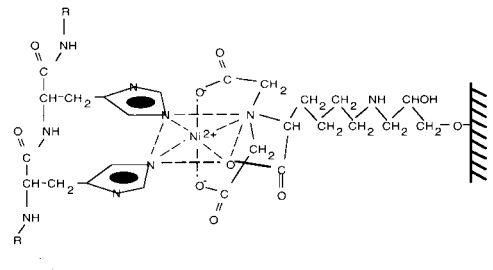
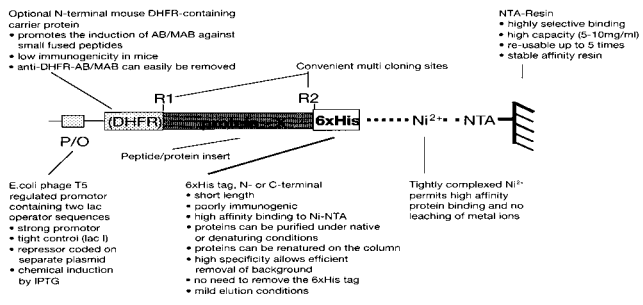
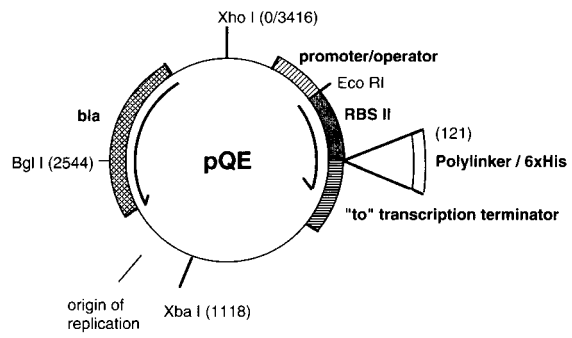
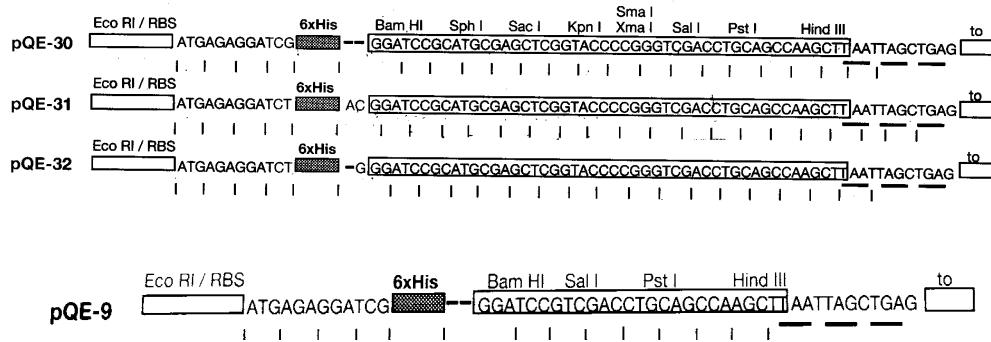
**Chromosome:** 9 mv  
**Cytogenetic:** 9q32 RefSeq  
**Markers:** Chr. 9 [D9S2084](#) D9S2084 mv  
Chr. 9 [A004Y48](#) mv

NCBI Reference Sequences (RefSeq) ?

**Category:** **PREDICTED**

**mRNA:** [NM\\_032303](#)  
**Protein:** [NP\\_115679](#) hypothetical protein BL  
MGC10940

附錄 6-2：H-HSD 在 Uni-Gene 查尋所得之資訊與訊息。



(摘錄自 QIAGEN 公司商品說明書)

附錄 7：pQE 質體表達及純化系統。



NCBI : blastx / nr

[gi|14150062|ref|NP\\_115679.1|](#) (NM\_032303) hypothetical protein  
MGC10940 [Homo sapiens]

[gi|13279254|gb|AAH04331.1|AAH04331](#) (BC004331) Similar to RIKEN  
cDNA 2610207I16 gene [Homo sapiens]

Length = 345

Identities = 69/97 (71%), Positives = 78/97 (80%)

Frame = +3

Query: 72

MLQNTGKLAGCTIFITGASRGIGKAIALKAAQDGANVVIAAKTADPHPKLPGTIYTXXX

X 251

ML NTG+LAGCT+FITGASRGIGKAIALKAA+DGAN+VIAAKTA PHPKL

GTIYT

Sbjct: 1

MLPNTGRLAGCTVFITGASRGIGKAIALKAAKDGANIVIAAKTAQPHPKLLGTIYTAAE

E 60

Query: 252 XXXXXXXXLPCIVDVRDEKQINDAVEQAVEKFGGIDI 362

LPCIVDVRDE+QI+ AVE+A++KFG I

Sbjct: 61 IEAVGGKALPCIVDVRDEQQISAAVEKAIKKFGAYTI 97

Identities = 18/25 (72%), Positives = 19/25 (76%)

Frame = +3

Query: 558 KNHTAYTICKYGMXMCVLGMAEEFR 632

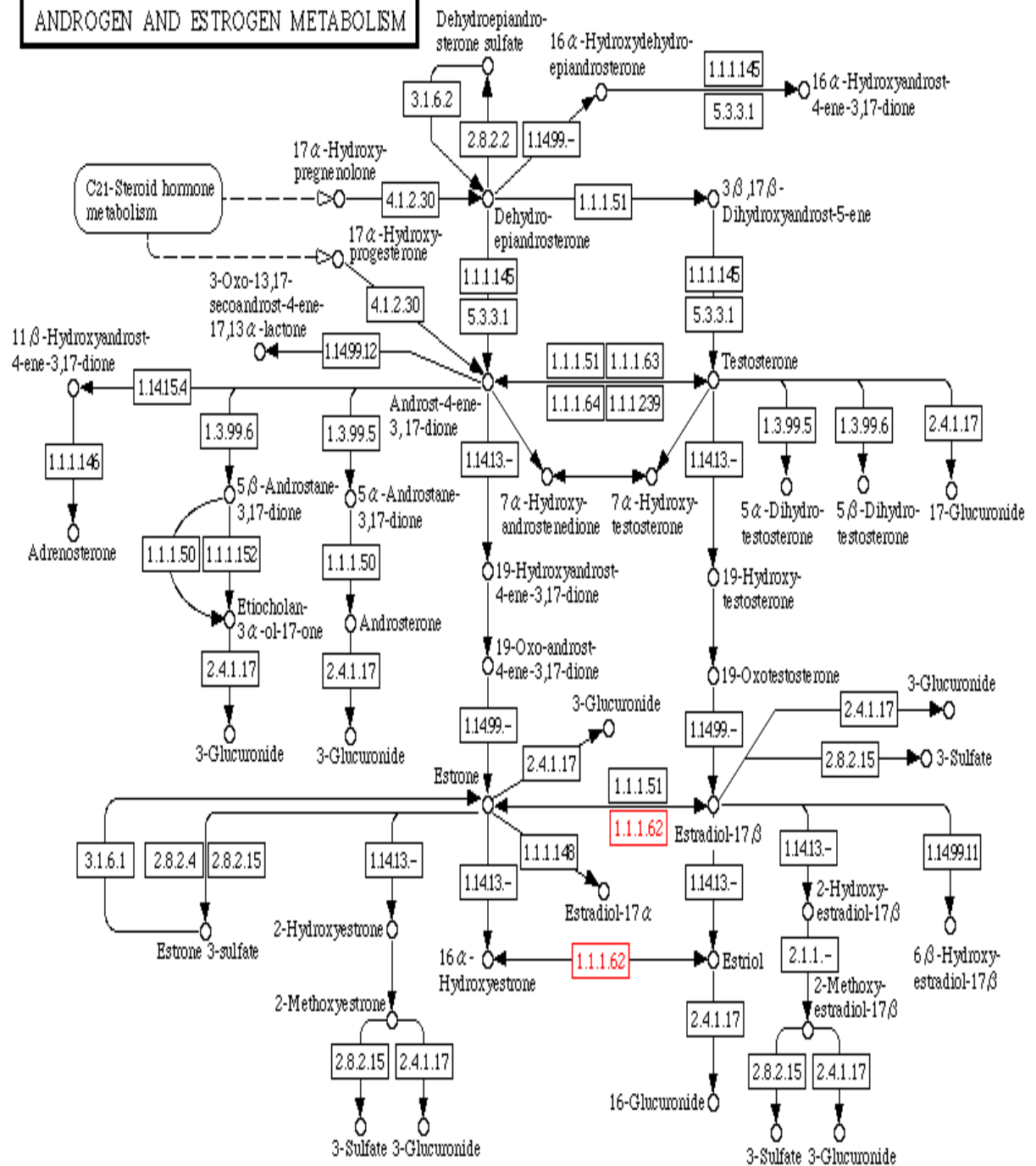
K AYTI KYGM M VLGMAEEF+

Sbjct: 90 KKFGAYTIAKYGMSMYVLGMAEEFK 114

附錄 8：Z-HSD 序列 BLAST/ non-redundant 資料庫之結果。以 Z-HSD 之片段序列(附錄 1 序列)為 Query sequence 詢問 NCBI 下之 BLAST / non-redundant (nr) database。比對到一相互對應之人類蛋白質，彼此間之相似性(identity)為 71%。



**ANDROGEN AND ESTROGEN METABOLISM**



00150 3/22/02

(摘錄自 KEGG 網站)

附錄 9：Androgen and Estrogen Metabolism Pathway.

Entrez 入口網址

點選核酸資料庫

詢問之關鍵字

Entrez is a retrieval system for searching several linked databases. It provides access to:

- [PubMed](#): The biomedical literature (PubMed)
- [Nucleotide](#): sequence database (GenBank)
- [Protein](#): sequence database
- [Structure](#): three-dimensional macromolecular structures
- [Genome](#): complete genome assemblies
- [PopSet](#): population study data sets
- [OMIM](#): Online Mendelian Inheritance in Man

附錄 10-網頁 1：NCBI 之 Entrez 入口網址。

NCBI *translating* **BLAST**

Nucleotide Protein Translations Retrieve results for an RID

Search

輸入詢問之核酸序列

Choose a translation TRANSLATED query - PROTEIN database [blastx]

Set subsequence From: To:

Choose database nr non-redundant database

Genetic codes Standard (1)

Now: **BLAST!** or **Reset query** **Reset all**

附錄 10-網頁 2：NCBI 之 BLAST/non-redundant database 服務輸入 Query 資料之位址。

網址 http://www.ncbi.nlm.nih.gov/BLAST/

NCBI BLAST

PubMed Entrez BLAST OMIM Taxonomy Structure

**What's NEW in BLAST®**

**NEW April 1st 2003:** We will be reorganizing our FTP site in order to allow easier access to the preformatted BLAST databases that users of NCBI BLAST should be using. The FASTA format database files will be moved to <ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/> and the BLAST preformatted database files will be moved up one level to <ftp://ftp.ncbi.nlm.nih.gov/blast/db/>. This will occur on April 8, 2003. For more information please see "New and Noteworthy" link in the sidebar.

**Nucleotide BLAST** ?

- [Standard nucleotide-nucleotide BLAST \[blastn\]](#)
- [MEGABLAST](#)
- [Search for short nearly exact matches](#)

核酸與胺基酸相似序列比對

**Translated BLAST Searches** ?

- [Nucleotide query - Protein db \[blastx\]](#)
- [Protein query - Translated db \[tblastn\]](#)
- [Nucleotide query - Translated db \[tblastx\]](#)

結合結構區(domain)搜尋

**Search for conserved domains** ?

- [Search the Conserved Domain Database using RPS-BLAST](#)
- [Search by domain architecture \[CDART\]](#)

**Pairwise BLAST**

- [BLAST 2 Sequences](#)

比對兩序列之相似度

**Genomic BLAST pages** ?

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Fugu rubripes](#)
- [Zebrafish](#)
- [Flies](#)
- [Nematodes](#)
- [Plants](#)
- [Yeasts](#)
- [Malaria](#)
- [Other eukaryotes](#)
- [Microbial](#)

搜尋 EST 延長或找尋完整序列

**Specialized BLAST pages**

- [VecScreen - BLAST-based detection of vector contamination](#)
- [IgBLAST - Analysis of immunoglobulin sequences in GenBank](#)
- [Trace BLAST - A page optimized for cross-species comparisons](#)

**Retrieve results for an existing Request ID** ?

- [Retrieve results with a Request ID](#)

**JavaScript free BLAST pages** ?

- [Get the BLAST home page with JavaScript free links](#)

Disclaimer Privacy statement

Revised January 29, 2001

http://www.ncbi.nlm.nih.gov/BLAST/Genome/NematodeBlast.html

網際網路

附錄 10-網頁 3：NCBI 之 BLAST 服務入口網址。

[Search](#)

[Choose a translation](#) TRANSLATED query - TRANSLATED database [tblastx]  輸入詢問之核酸序列

[Set subsequence](#) From:  To:

[Choose database](#) est\_others  est others database

[Genetic codes](#) Standard (1)

Now:  or

附錄 10-網頁 4：NCBI 之 BLAST/est\_other database 服務輸入 Query 資料之地址。

**BLAST**  
overview  
FAQs  
news  
manual

## Blast The Zebrafish Genome



**references** Blast your sequence against zebrafish specific sequences.

MegaBLAST  
Danio Traces  
enter a  
nucleotide

Database: EST Traces Program: blastn

use MegaBLAST

Begin Search

EST 或 EST Traces  
database

核酸比對

query  
sequence  
and get a  
match, see  
a tblastx  
alignment  
of your  
two  
sequences

Enter an accession, gi, or a sequence in FASTA format:

輸入詢問  
之核酸序

Optional parameters

**Expect Filter Descriptions Alignments**

0.01 default 100 100

Advanced options: [ ]

Zebrafish Genome  
比對服務網址

網址 http://bioinformatics.weizmann.ac.il/cards/

GeneCards 入口網址

**GeneCards™** an academic web site of the WEIZMANN INSTITUTE OF SCIENCE

[Terms of Use](#) | [GeneCards Homepage](#) | [Search Examples](#) | [Comment F](#)

**Notice** - Please read carefully prior to linking to any third-party site.

**About GeneCards**  
[What's New](#)  
[Data Sources](#)  
[Data Extraction](#)  
[Citing This Resource](#)  
[Publications](#)

**Mirror sites**

**Version: 2.27**  
**Release: May 5, 2003**  
**Entries: 47,794**  
**Approved\*: 16,038**

<b>Genes</b>	<b>21,135</b>
<b>Predicted genes with support</b>	<b>10,588</b>
<b>Other predicted genes</b>	<b>3,372</b>

**SEARCH** **Notice** - Please read carefully prior to linking to any third-party site.

| [Quick Start](#) | [Guided Tour](#) | [More search examples](#) |

• Search/Display GeneCards™ by


輸入欲詢問之 NCBI accession


*For example, you can display the GeneCard for the (sensitive) **symbol** [FMR1](#), for a **random** gene, for a **random** [HUGO approved gene](#), or see GeneCards™ for the **keyword(s)**:*


- [p53](#)
- [apolipoprot\\*](#)**AND** ([hyper\\*](#) **OR** [Alzheim\\*](#))
- [GC21M032\\*](#) *region on chromosome/strand*
- [U85267](#) *GenBank accession No.*
- [Hs.1288](#) *UniGene cluster*

附錄 10-網頁 6：GeneCards 入口網址。



網址  http://bioinfo.weizmann.ac.il/cards-bin/cardsearch.pl?search=NP\_115679



an academic web site of the WEIZMANN INSTITUTE OF SCIENCE 

[Terms of Use](#) | [GeneCards Homepage](#) | [Search Examples](#)

[Notice - Please read carefully prior to linking to any third-party site](#)

**SEARCH** GeneCards

NP\_115679

[\[Open a small search window\]](#)

[\[Quick start help, Search help\]](#)

**RESULT:**

**1 GeneCard** matches **your precise query** for "NP\_115679";  
 It is represented by a minicard.  
 Click "Display" on the left to get the full GeneCard.

<p><a href="#">Display</a> the complete GeneCard for this gene (MGC10940)</p> <p><a href="#">More like this</a></p>	<p>Gene: MGC10940 [Locus: --]</p> <p>The following lines in the GeneCard text contribute to matching your query:</p> <p>- LOC: MGC10940   OMIM:   chromosome:9   location:9q32   genename:hyp        unclassified   DICTED   Start: 106874530   End: 106966854   Contig:        - S: NP_115679   Validation: +[Expression Data: LocationType=mrna-utr; mRNA        frequency ContigLocationTo:5423845 ContigLocationFrom:5423845]        Data: LocationType=mrna-utr; mRNAAcc=NM_032303; [ExperimentalValidation: by-        frequency ContigOrientation: forward ContigLocationTo:5424062 ContigLocationFrom:5424062]        Data: ProtAcc=NP_115679; LocationType=intronic; mRNAAcc=NM_032303; [Experimen        tValidation: by-frequency ContigOrientation: reverse ContigLocationTo:5411589 ContigLocationFrom:5411589]        Data: ProtAcc=NP_115679; LocationType=intronic; mRNAAcc=NM_032303; [Experimen        tValidation: by-frequency ContigOrientation: forward ContigLocationTo:5411495 ContigLocationFrom:5411495]        Data: ProtAcc=NP_115679; LocationType=intronic; mRNAAcc=NM_032303; [Experimen        tValidation: by-frequency ContigOrientation: forward ContigLocationTo:5392132 ContigLocationFrom:5392132]        Data: ProtAcc=NP_115679; LocationType=intronic; mRNAAcc=NM_032303; [Experimen        tValidation: by-frequency ContigOrientation: forward ContigLocationTo:5358855 ContigLocationFrom:5358855]</p>
---	--

點選 Display 查  
看詳細資料

附錄 10-網頁 7 : GeneCards search 基因資料之入口。



網址(D) <http://www.genome.ad.jp/kegg/>

# KEGG: Kyoto Encyclopedia of Genes and Genomes

KEGG 網站之入口網址

Release 26.0, April 2003 (plus daily updates)

---

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a bioinformatics resource for understanding higher order functional meanings and utilities of the cell or the organism from its genome information. The KEGG project is undertaken in the Bioinformatics Center, Institute for Chemical Research, Kyoto University with supports from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Japan Society for the Promotion of Science (JSPS), and the Japan Science and Technology Corporation (JST).

- **Introduction**
- **Open KEGG (Table of Contents)**
- **Utilize KEGG (XML and API)**
- **Links to Related Databases**
- **KEGG Distribution**
- **Go to GenomeNet Home**

附錄 10-網頁 8：KEGG 網站之入口網址。



# KEGG - Table of Contents

[ [PATHWAY](#) | [GENES](#) | [SSDB](#) | [LIGAND](#) | [EXPRESSION](#) | [BRITE](#) | [BLAST/FASTA](#) | [DBGET](#) | [API](#) ]

Search  All organisms  for

## 1. Pathway Information

### 1-1. Pathway Maps and Ortholog Tables -- Update Status

Category	Pathway Map Ortholog Table	Search & Compute	DBGET Search
Pathway	<b>Metabolic pathways</b>	XML Search objects in pathway maps Color objects in pathway maps Search or color genes in ortholog tables	PATHWAY
	<b>Regulatory pathways</b>	Search similar sequences in pathway maps Search similar sequences in ortholog tables Generate possible reaction pathways	

### 1-2. Hierarchical Classifications

Category	Classification	DBGET Search	Chemscape Search	Link
Disease	ICD disease classification OMIM gene map OMIM morbid map			
Ortholog	<b>KEGG Orthology (KO)</b>	KO		SSDB
Enzyme	EC number classification	ENZYME		LIGAND
Reaction		REACTION	LIGAND	
Compound	Compound classification	COMPOUND		COMPOUND
Element	Periodic table			

點選進入以查尋酵素代謝路徑

附錄 10-網頁 9：KEGG 查尋酵素代謝反應路徑之位址。

## Search ENZYME database using DBGET

Database: ENZYME

Ligand Chemical Database for Enzyme Nomenclature  
Release 26.0+/06-18, Jun 03  
Institute for Chemical Research, Kyoto University  
4,171 entries

Input your search keyword(s) into the search box

example) kinase

Submit

Reset

bfind mode    Max number of hits:    
 bget mode

輸入欲查尋之酵素

EC (Enzyme Commission) numbers assigned by [IUPAC-IUBMB](#)

[ [1st Level](#) | [2nd Level](#) | [3rd Level](#) | [4th Level](#) | [Text Search](#) ]

[1. Oxidoreductases](#)

[2. Transferases](#)

[3. Hydrolases](#)

[4. Lyases](#)

[5. Isomerases](#)

[6. Ligases](#)

點選欲查尋之酵素種類

附錄 10-網頁 10：KEGG 查尋酵素反應及參與受質之位址。

## H-HSD SNPs

GATAAATGCGGAGGGACGGTCCAGCTTTAGCTCTCTGCTCGCCGCCGCCGCTGTCGCCGC  
CACCTCCTCTGATCTACGAAAGTCATGTTACCCAACACCCGGGAGGCTGGCAGGATGTACA

.....-R-L-A-G-C-T-  
GTTTTATCACAGGTGCAAGCCGTGGCATTGGCAAAGCTATTGCATTGAAAGCAGCAAAG  
-V-F-I-T-G-A-S-R-G-I-G-K-A-I-A-L-K-A-A-K-  
GATGGAGCAAATATTGTTATTGCTGCAAAGACCGCCAGCCACATCCAAAACCTCTAGGC  
-D-G-A-N-I-V-I-A-A-K-T-A-Q-P-H-P-K-L-L-G-  
ACAATCTATACTGCTGCTGAAGAAAATTGAAGCAGTTGGAGGAAAGGCCTTGCCATGTATT  
-T-I-Y-T-A-A-E-E-I-E-A-V-G-G-K-A-L-P-C-I-  
GTTGATGTGAGAGATGAACAGCAGATCAGTGCTGCAGTGGAGAAAAGCCATCAAGAAATTT  
-V-D-V-R-D-E-Q-Q-I-S-A-A-V-E-K-A-I-K-K-F-  
GGAGCTTATACCATTGCTAAGTATGGTATGTCTATGTATGTGCTTGGAAATGGCAGAAGAA  
-G-A-Y-T-I-A-K-Y-G-M-S-M-Y-V-L-G-M-A-E-E-  
TTTAAAGGTGAAATTGCAGTCAATGCATTATGGCCTAAAACAGCCATACACACTGCTGCT  
-F-K-G-E-I-A-V-N-A-L-W-P-K-T-A-I-H-T-A-A-  
ATGGATATGCTGGGAGGACCTGGTATCGAAAAGCCAGTGTAGAAAAGTTGATATCATTGCA  
-M-D-M-L-G-G-P-G-I-E-S-Q-C-R-K-V-D-I-I-A-  
GATGCAGCATATTCCATTTTCCAAAAGCCAAAAGTTTACTGGCAACTTTGTCATTGAT  
-D-A-A-Y-S-I-F-Q-K-P-K-S-F-T-G-N-F-V-I-D-  
GAAAATATCTTAAAAGAAGAAGGAATAGAAAATTTTGACGTTTATGCAATTAAACCAGGT  
-E-N-I-L-K-E-E-G-I-E-N-F-D-V-Y-A-I-K-P-G-  
CATCCTTTGCAACCAGATTTCTTCTTAGATGAATACCCAGAAGCAGTTAGCAAGAAAAGTG  
-H-P-L-Q-P-D-F-F-L-D-E-Y-P-E-A-V-S-K-K-V-

R

GAATCAACTGGTGCTGTTCCAGAATTCAAAGAAGAGAAAAGCTGCAGCTGCAACCAAAAACCA  
-E-S-T-G-A-V-P-E-F-K-E-E-K-L-Q-L-Q-P-K-P-  
CGTTCTGGAGCTGTGGAAGAAACATTTAGAATTGTTAAGGACTCTCTCAGTGATGATGTT  
-R-S-G-A-V-E-E-T-F-R-I-V-K-D-S-L-S-D-D-V-  
GTTAAAGCCACTCAAGCAATCTATCTGTTTGAACCTCTCCGGTGAAGATGGTGGCACGTGG  
-V-K-A-T-Q-A-I-Y-L-F-E-L-S-G-E-D-G-G-T-W-  
TTTCTTGATCTGAAAAGCAAGGGTGGGAATGTCGGATATGGAGAGCCTTCTGATCAGGCA  
-F-L-D-L-K-S-K-G-G-N-V-G-Y-G-E-P-S-D-Q-A-  
GATGTGGTGATGAGTATGACTACTGATGACTTTGTAAAAATGTTTTCAGGGAAACTAAAA  
-D-V-V-M-S-M-T-T-D-D-F-V-K-M-F-S-G-K-L-K-  
CCAACAATGGCATTTCATGTCAGGGAAAATTGAAGATTAAGGTAACATGGCCCTAGCAATC  
-P-T-M-A-F-M-S-G-K-L-K-I-K-G-N-M-A-L-A-I-



AAATTGGAGAAGCTAATGAATCAGATGAATGCCAGACTGTGAAGGAAAATATAAAAAAAAA  
-K--L--E--K--L--M--N--Q--M--N--A--R--L--\*-. . . . .  
AGTCGACTGCTATGCTCAAAAAAGTAAAAAAGCTCAACAGTTAAAATCTAATGTTTGT  
TCTTTCTGTTATATTATAAGGATATGCACGTTTGTCTGGAAAAGATAGAATTTGTCTC  
TAAAAGACTTGAAATTGTAATTA AAAATGGCAAGCTAATCAAACATAAGCTTCATTAAGTG  
GGATTCTAAGACAGTCTGTGTTTTTATATTTCAAGGGTTAACCCTTTGAGCCTTACATC  
TCATTCACTGTCTTTCTCCAAGAAAAGTATTTTGGGCGGACAGTCAGATCAAGCAGTAAA  
ATTAGCTCTTTCAAATCTTCTGTGCATGTA AAAATGAAGCTAGTCTGTTTTAAAATTTTTA  
GTTTTGGATTGTATACTAATGAAAATCTTAATGATGTTTTTGATTTTTATATACTTATTT  
TAAAGAAAATCTTATATAGTACATTTTACAAAAATTATAAAAAATGAATTAGTACTGGCG  
AGGACTAAATGAAACAATAATTTTTCATTTTGATAACTAGCTTTCCAGGTGGACTTAGCC  
ATAGGAAAATATTACTAATGTAATTTAACAAATGCTGCATGTATCCATTTAAAATAT  
GTTTAAATTTGTCTTAAAACAAAAATAATTTTCTCCCTAGGAGTATGCATTTGGCTACAGTG  
TTTTGAAACAGAAACCTTAGAATAGGTCATTGGTATGGGCTGAACTGTGTATCCCCAAT  
TCATTTGTTGAGGTCCTAACTCCCATTCTTTTGAATGTGACTGTTCCGGAGATGAGGCCT  
TTAAAGAGGTGACTTAAGTTCAAAGGAGGCTGTTAGTCTAATCCAACATGGTGTCTTTG  
GACATAAGAGATACCAGCAATGTGTGCACAGAACAAGACCAGGAGAGGACACAGTGAGA  
AGGCAGTTATCTGCAAGCAAAGAGAGAGGCTTCAGAAGAAAACAAAATCACCAGCACCTTG  
ATCTTTGACTTCTAATCTCCAGAATAGTGAGAAAATAAATTTCTGTTGTTAAGCCGTCCAC  
TGTGGGAGGCCGACGCAGGAGGATTGCTTGAGGCCAGGAGTTCAAGGCCAGCCTGGACAA  
CATAGTAAGACCTATCTCTACCCCCTAATAAATTAATTTAAAAAGCCCCCAATCTGT  
GGTATTTTATTATGGCAGCCCTAGCAAGCTAATACAGTGGTTTGAGAGGCTGGGAGGGTT  
GAGGGGAAGATAAACTTTTAAAAAGCTCTTATCTTTCAATTCAGTTAAAAATACTT  
GCTCAGTGTAACAATTTTGCTTCTCAGCTTCCACTCTAATATTGTTGTGCCATTAAGCAA  
TTTAGCTAATCCTGACATTTCTTAGATTCATAATGTTAGGAGCATTTAATCTGTATTTTA  
CAAGTTAGGAAGCAGAGGATCAGAGATGGGAAAAGGACTAGCCCAAGGCCAACATTAACAA  
GCCCTCTAACAAAACTTTACAATACATTTATGTTGAATGGAACTCCAAGATCTCACCTC

S

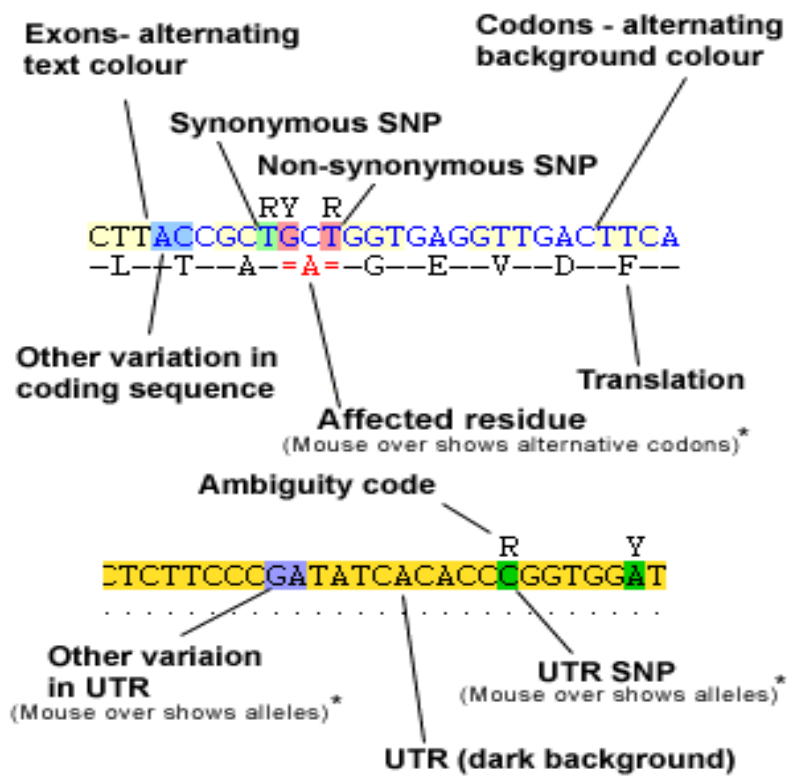
TCCATCCAGGAATGGAGTCCATGTAATCAAAGTGAACCTAAAAATAGGACAGTTTCAACA  
AGTCAGGAGATTCACAGCAACTGATCAAAGGGAGTCCAGTCAACGTGAGCAAGCGTGATT  
ATGATGAGGAAGCCCCCTGCTTTAATCCACACAAGGAACGTAACCTGAAGTAACCTGA  
TGTTAACCAATCTGCTGTGTCTACTATGCTGTTTCTTGTTCCTGCTAGTGCTGCTTTAC

R

AAATGCAGACCATTCTATCATACCTGGCGGGGCTTCTGTTTTATTTTGTAGGCTGGATGC  
TACCCAGTTCATGAATCGCTAATAAAAAGCCAATTAGATCTTTACATAA

(摘錄自 Gene Cards 網站)

附錄 11-1 : H-HSD 之 single nucleotide polymorphism (SNPs)。



\* - for HTML4 compliant browsers (NS6+, IE5+)

(摘錄自 Gene Cards 網站)

附錄 11-2：single nucleotide polymorphism (SNPs)之標示說明。

TTTAATGTGT CACCAGGAGC CCCTCTGGAT CATTGCGTTC CCATATCAGA  
 CATCATCCAG GAGACACAAT AATGCTGCAG AACACAGGGA AGCTGGCAGG  
 ATGCACCATT TTCATAACCG GAGCAAGTCG AGGTATTGGC AAGGCCATTG  
 CTCTGAAAGC TGCGCAGGAT GGAGCCAATG TAGTCATCGC TGCCAAAACA  
 GCTGATCCAC ATCCCAAACCT CCCCGGCACC ATCTACACAG CTGCAGCAGA  
 GATTGAAGCA GCTGGAGGGA AAGCACTGCC GTGTATTGTG GACGTCCGTG  
 ATGAGAAGCA GATCAATGAT GCTGTTGAAC AGGCTGTGGA GAAGTTTGGGA  
 GGGATTGACA TATTGGTCAA CAATGCCAGT GCCATCAATT TAACAGGGAC  
 TCTTCAGACT CCAATGAAAA AGGCTGACCT CATGCTGGGC ATCAATCTCA  
 GGGGAACATA CCTGACGTCT AAACTGTGCA TTCCACATCT TCTGAAGAGC  
 AAAAACCCCTC ACATACTAAA CCTCAGCCCA CCTCTCAACC TCAACCCCAT  
 TTGGTTTAAA AACCACACGG CTTACACCAT TTGCAAGTAT GGCATGTNCA  
 TGTGTGTACT GGGAAATGGCA GAAGAGTTCA GAGGATCCAT TGCCGTTAAT  
 GCCTTATGGC CAAAGACAGC CATTGAGACG GCTGCCATGG ACATGTTGGG  
 CGGTTCTGAG GTTGGCAAGC AGTGCAGGAA GGTGGAGATC ATGGCTGATG  
 CAGCATACGC CATCTTCAA CAACCCACCA GCTTCACCGG ACAGTTTGTT  
 ATTGATGAGG ACATTCTCAA AAAGGAGGGC ATTAAAGATT TTGATGTTTA  
 TGCTGTTGAG CCAGGTCATC CATTGCTTCT GACTTTTTCT TGGACGGCCA  
 GCCTGAGGAT CTAGTCAAGC ATATGGAGGC ACATGGTGCC ACTCCGGCGT  
 TCACAACCTGC AAAAGCAGAT CCCGTTGCCG CAGGACCAGT TTCTGAGATG  
 TCAATACAAT CAGAGGAATT ATCAGTCCAG AGATGGTGAA AACCACACAA  
 GGAGTGTC AAATTAACCTGG CAGGCGAACA TGCTGGAGCT GGNATCTTGC  
 CTGAAGAACG ATCTGGAATG CTGGAAATGG GGAACCTTCT GTNAACTGAT  
 GTTGCATGTC GATGGCCATG AGGATTTGNC AANATGNTTG GANGGAAATT  
 AAANCCANCC ATGNCCTTCA TGTCTGGAAA GCTGACTATT AAGGGTGACA  
 TGGCCCTTGT CATCAAACCTG GAGAAGATGA TGGCCATGAT GAAGTCTAAA  
 CTGTGAACNC ATCACGCAGT NTCCTCCCTT ACAANATGTA CATNATGCC  
 TCTTTTCTAN ATCAGTGTTT TTTGATTAAA CTAATTGTTC AGATTGTACT  
 AATTTTAATA AAGAACTATC AGGCAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAA

附錄 12: Z-HSD 全長核酸序列組裝。564~632 為 Z-HSD 原序列(黑色)與 3' RACE  
 PCR 延長序列(紫色)之重疊部分(黃色部分), 第 72~74 及 1304~1306 號之核酸(紅  
 色)為推定之起始密碼子和終止密碼子。

H-HSD (1) MLPNTGRLAGCTVFITGASRGIGKAIALKAAKDGANIVIAAKTAQPHPKL  
Z-HSD (1) MLQNTGKLAGCTIFITGASRGIGKAIALKAAQDGANVIVIAAKTADPHPKL

H-HSD (51) LGTIYTAAEEIEAVGGKALPCIVDVRDEQQISAAVEKAIKKFGAYTIKAKY  
Z-HSD (51) PGTIYTAAAEIEAAGGKALPCIVDVRDEKQINDAVEQAVEKFGGIDILVN

H-HSD (101) GMSMYVLG-----MAEE  
Z-HSD (101) NASAINLITGTLQTPMEKADMLGINLRGTYLTSKLCIPHLLKSKNPHILN

H-HSD (113) FKGEIAVNALWPK--TAIH TAAMDMLGGPGIESQCRKVDIADAAYSIFQ  
Z-HSD (151) LSPPLNLHPIWFKNHTGLHCKVWHVQCVTWEWQKSSQDPLPLMPYGQRQ

H-HSD (161) -----KPKSFTGNFVIDENIL  
Z-HSD (201) PFRRLPWCWAVLRLASSAGRWRWSWLMQHTPSSNPPASPDSLLLMRTFS

H-HSD (177) KEEGIENFDVYAIKPGHPLQPDFFLDEYPEAVSKKVESTGAVPEFKEEKL  
Z-HSD (251) KRRALKILMFMLLSQVIHCFPDFFLDGQPEDLVKHMEAHGATPAFTTAKA

H-HSD (227) QLQPKPRS GAVEETFRIVKDSLSDDVVKATQAIYLFELSGEDGGTWFLLDL  
Z-HSD (301) DP---VAA GPVSEMFNTIRGIISPPEMVKTTQGVYKFNLAGEHAGVWYLLDL

H-HSD (277) KSKGNGVGYGEP SDQADVVMSTTD DFKMFS GKLKPTMAFMSGKLIK  
Z-HSD (348) KNDAGSAGNGEP PVKADVVMST DSEDFVKMFGKLKPTMAFMSGKLIK

H-HSD (327) NMAIAIKLEKLMNQMNARL  
Z-HSD (398) DMALVIKLEKMMAMMKS KL

附錄 13 : conserved motifs of SDR(short chain dehydrogenase/reductase) family。

NAD(P)<sup>+</sup>-binding rossmann fold : G17~G21~G23(H-HSD/Z-HSD) 、  
Y130~K134(Z-HSD) 、 G137~S143(H-HSD) 、 Y156~K161(H-HSD) 、  
P191G192(H-HSD)。