

00381

2
e.



**UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO**

**FACULTAD DE CIENCIAS
División de Estudios de Posgrado**

**Génes de Hemoglobina en plantas superiores:
Busqueda en Monocotiledoneas mediante
el uso de una Sonda Consenso.**

**TESIS CON
FALLA DE ORIGEN**

T E S I S

**Que para obtener el Grado Académico de
DOCTOR EN CIENCIAS
(BIOLOGIA)
presenta**

RAUL ARREDONDO PETER

México, D. F.

1992



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO.

RESUMEN.....	i
ABSTRACT.....	ii
ABREVIATURAS.....	iii
I. INTRODUCCION.....	1
II. OBJETIVOS.....	18
III. MATERIALES Y METODOS.....	19
IV. RESULTADOS Y DISCUSION.....	31
V. CONCLUSIONES.....	50
VI. BIBLIOGRAFIA.....	55
APENDICE 1.....	60
APENDICE 2.....	80

RESUMEN. Los genes de las hemoglobinas y las proteínas que codifican se han encontrado en diversas plantas Dicotiledóneas. La detección de estos genes se ha facilitado debido al uso de los procedimientos de la biología molecular; en cambio, las proteínas se han detectado mediante técnicas de espectroscopía de alta sensibilidad. Sin embargo, el uso de las sondas moleculares, que corresponden a los genes de hemoglobina de determinadas especies de Dicotiledóneas, son de uso limitado debido a fenómenos de divergencia. Por lo tanto, en este trabajo se obtuvo una secuencia consenso para las hemoglobinas de las plantas y sus genes, con el objeto de detectar y de seleccionar las regiones más conservadas, las cuales pudieran ser de utilidad en la búsqueda de genes de hemoglobina en plantas no-Dicotiledóneas, particularmente en las Monocotiledóneas. Las regiones más conservadas correspondieron a aquellas que codifican para las α -hélices A, C y H; de entre ellas, se seleccionó a la región que codifica para la α -hélice H, a partir de la cual se sintetizó una sonda molecular. Esta sonda se marcó con derivados biotinilados y se hibridizó mediante procedimientos del tipo "Dot" y "Slot-Blot" con el DNA total de algunas Leguminosas y plantas no-Dicotiledóneas. Se obtuvieron señales de hibridización fuertes con los DNAs de Leguminosas y, de manera muy interesante, también con los DNAs de las Gramíneas, aunque éstas dieron señales con menor intensidad. Estos resultados sugieren fuertemente que las Monocotiledóneas contienen genes de hemoglobina. Además, al utilizar las regiones conservadas, se logró amplificar por PCR fragmentos discretos a partir del DNA del Maíz, del Sorgo y del Trigo, los cuales se deben clonar y secuenciar para confirmar la presencia de los genes de hemoglobina en las plantas no-Dicotiledóneas.

ABSTRACT. Hemoglobin (Hb) genes and transcripts have been found in a variety of dicot plants. In addition to spectral analysis, genetic approaches using specific gene probes have facilitated the detection of *hb* genes; however, the use of these probes is limited due to divergence events. The detection of *hb* genes in non-dicot plants is of evolutive, as well as physiological interest. In order to test other strategies in the search of *hb* genes in the monocots and other plants, we constructed a consensus sequence. The most conserved regions were those that code for the α -helices A, C and H. The α -helix H coding segment was selected for the synthesis of an oligo-probe. This probe was biotin-labelled and hybridized with total DNA from legume and non-dicot plants by dot- and slot-blot procedures. High hybridization signals were observed with legume DNAs; interestingly, monocot DNAs hybridized with the oligo-probe as well, thus suggesting the presence of *hb*-like sequences in these non-dicot plants. The confirmation of these monocot DNA hybridizing fragments as authentic *hb* genes must come from cloning and sequencing procedures. In addition, PCR-derived *hb* fragments have been amplified for corn, wheat and sorghum species, which will be further sequenced.

SELECCION DE ABREVIATURAS

Amp	Ampicilina
BCIP	5-bromo-4-cloro-3-indolilfosfato
BSA	Albúmina sérica de bovino
CAT	Cloranfenicol acetil transferasa
cdNA	DNA copia, que resulta de la transcripción inversa del RNAm en DNA al utilizar la enzima reversotranscriptasa
CTAB	Bromuro de hexadeciltrimetilamonio
2,4-D	Acido 2,4-diclorofenoxiacético
dATP	Desoxiadenosina de trifosfato
dNTP	Desoxinucleótido de trifosfato
EDTA	Acido etilendiaminotetracético
GUS	β -Glucuronidasa
Hb	Hemoglobina
hb	Gen de hemoglobina
IFTG	Isopropiltiogalactósido
KDa	KiloDalton
Kpb	Kilopares de bases
Lb	Leghemoglobina
Mb	Mioglobina
NADH	Nicotinamida de adenina dinucleótido reducido
NBT	Azul de nitrotetrazolium
NC	Papel de nitrocelulosa
nif	Grupo de genes bacterianos que codifican para la síntesis de las enzimas que participan en el proceso de la fijación del nitrógeno
nod	Grupo de genes bacterianos que codifican para la síntesis de proteínas que permiten el establecimiento de la simbiosis entre la planta Leguminosa y <i>Rhizobium</i> , lo que resulta en la formación de un nódulo
NPTII	Neomicina fosfotransferasa
O₂⁻	Anión superóxido
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa ("Polymerase Chain Reaction")
pHBCS/H	Plásmido que se deriva del pUC19 que contiene al inserto SCHBP/H
PM	Peso molecular
ppO₂	Presión parcial de oxígeno
PVP	Polivinilpirrolidona
SCHBP/H	Sonda molecular que corresponde al fragmento de la secuencia consenso del gen <i>hb</i> de las plantas que codifica para la α -hélice H
SCTHBP	Secuencia consenso total de las hemoglobinas de las plantas
SDS	Dodecilsulfato de sodio
Tris	Tris(hidroximetil)aminometano
x-Gal	5-bromo-4-cloro-3-indolil- β -D-tiogalactósido

I. INTRODUCCION.

Se calcula que la vida surgió en la Tierra hace unos 3.5×10^9 años según evidencias geoquímicas y por microfósiles que se han encontrado en sedimentos que datan de esta antigüedad (1, 2). Las condiciones que prevalecían en esta Tierra primitiva eran distintas de lo que son en la actualidad: se cree que fue hace 3.8×10^9 años cuando el H₂O líquida se empezó a acumular en la litósfera, y que el carácter de la atmósfera en aquel entonces era predominantemente reductor (3). Así, se piensa que los primeros organismos vivientes proliferaron adaptándose a tales condiciones; en ellos se desarrolló un metabolismo anaerobio mediante el cual pudieron satisfacer sus requerimientos energéticos, y de quienes tenemos algunos representantes en la actualidad, como son todos los microorganismos anaerobios, por ejemplo los clostridia. Estas condiciones cambiaron con el advenimiento de un tipo nuevo de metabolismo: la fotosíntesis oxigénica. Esta innovación metabólica tuvo efectos drásticos en la naturaleza de la atmósfera primitiva, y en general en el curso que habría de tomar la vida en la Tierra, ya que el oxígeno molecular, O₂, el subproducto de este proceso, se habría de acumular en la atmósfera tornándola a una atmósfera oxidante, tal como la conocemos hoy en día.

Este O₂, cuya concentración aumentó en relación con la proliferación de los organismos fotosintéticos, fue un factor de selección muy importante hacia los organismos anaerobios ya que es sumamente tóxico debido a que puede alcanzar estados excitados con la generación de radicales libres, como el anión superóxido o el hidroxilo, los cuales son áltamente reactivos con las biomoléculas en presencia de diversos

metales (4). Así, en muchos organismos surgieron diferentes mecanismos que les permitieron adaptarse a la presencia de este elemento tóxico, pero también muchos otros se seleccionaron negativamente por los efectos letales del O₂.

Probablemente uno de los primeros, y más simples, mecanismos de adaptación fue el evitar el estar en contacto con el O₂ al ocupar nichos con ppO₂ muy baja, o el desarrollar estructuras subcelulares que les protegieran de este elemento, tal como las endosporas de muchas bacterias anaerobias actuales. Otro mecanismo de adaptación, y tal vez el más complejo, fue el desarrollo de un aparato metabólico capaz de utilizar al O₂ como el aceptor final de electrones para formar agua, y que durante este proceso se obtuviera energía en forma de ATP; esto es lo que en la actualidad conocemos como la respiración aerobia, la cual ha tenido un gran éxito evolutivo ya que aporta una cantidad mayor de energía a las células, unas 686 Kcal durante la oxidación de la glucosa hasta el CO₂ y el H₂O, en comparación con las 56 Kcal que se producen durante la respiración anaerobia, en donde se oxida la glucosa hasta el ácido láctico (4).

Existen además otros mecanismos de adaptación que son comunes entre los organismos aerobios y que resultan de la presencia de metalo y de hemo-enzimas, las cuales pueden eliminar a los radicales libres. De ellas, las más importantes son la superóxido dismutasa, que cataliza la siguiente reacción: $O_2^- + O_2^- + 2H^+ \rightarrow H_2O_2 + O_2$; así como la catalasa y la peroxidasa que eliminan, a su vez, al H₂O₂ mediante las siguientes reacciones, respectivamente: $H_2O_2 + H_2O_2 \rightarrow O_2 + 2H_2O$ y $H_2O_2 + RH_2 \rightarrow R + 2H_2O$ (4). Existen también otras hemo-proteínas, sin actividad catalítica, que interactúan con el oxígeno molecular mediante

su grupo prostético, el grupo hemo (Fe-protohemo IX), aunque en este caso no existe modificación alguna del O₂. En general, estas hemo-proteínas, que constituyen a la familia de las globinas (5), captan, almacenan y transportan el oxígeno atmosférico dentro de los tejidos en los que se encuentran. Mediante estudios de difracción de rayos-X, Perutz (6) descubrió cómo captan y liberan el O₂ las Hbs tetraméricas mediante mecanismos halostéricos en respuesta a algunos metabolitos celulares que actúan como ligandos modulares, tal es el caso del ácido 3-fosfoglicérico. La función de estas Hbs como reservorio interno del O₂ resulta de gran importancia para el metabolismo aerobio celular.

Tradicionalmente, se han detectado las Hbs en diversos grupos de animales, tanto invertebrados como vertebrados, en donde se han estudiado con mucho detalle. Ingram (7) fue el primero en inferir relaciones de descendencia entre las globinas monoméricas (Mb) y las tetraméricas (Hb) de animales vertebrados, en donde él propuso a la subunidad α de la Hb como el componente ancestral del que se derivaron las demás subunidades, que incluye a la Mb. Además de los animales, las Hbs se encuentran en organismos más simples, como son los hongos (8) y los protoctistas (9); sin embargo, estas Hbs han sido poco estudiadas hasta el momento.

Para el grupo de las plantas, en 1947 Keilin y Smith (10) propusieron, después de un análisis químico y espectral, que el pigmento rojo que da la coloración a los nódulos radicales de las Leguminosas fijadoras del nitrógeno presenta características similares a las Hbs de los animales. En efecto, en el reino Plantae se sintetiza una Hb monomérica que se conoce como leghemoglobina (Lb) debido a que

por mucho tiempo se le encontró asociada solamente con las plantas Leguminosas. Las Hbs de origen vegetal y animal se encuentran muy conservadas, particularmente en relación con la estructura terciaria: la sobreposición de las estructuras tridimensionales de la Lb de la Soya y la Mb del esperma de ballena es casi perfecta, aún cuando existen diferencias en las estructuras primarias (11).

Las Lbs se han estudiado con mucho detalle desde su descubrimiento, y debido a la gran cantidad de información que se ha acumulado en relación con la genética, la estructura y la función de estas proteínas, se recomienda consultar las revisiones de Appleby (12, 13), y de Arredondo (14, 15).

Durante mucho tiempo las Lbs solamente se pudieron detectar en los nódulos radicales y caulinares de las Leguminosas que fijan el nitrógeno cuando establecen simbiosis con bacterias de los géneros *Rhizobium*, *Bradyrhizobium* y *Azorhizobium*. Actualmente se han encontrado Hbs en los nódulos de las plantas actinorrizas, que son infectadas por *Frankia*, y en la no-Leguminosa *Parasponia* (Ulmaceae) que es nodulada por algunas cepas de *Rhizobium*. En todos estos casos la Lb se relaciona intrínsecamente con la eficiencia en la fijación del nitrógeno, ya que funciona aportando las cantidades necesarias del O₂ para apoyar las demandas respiratorias del microsimbionte (12). La concentración del O₂ dentro del nódulo es muy baja, de 10 nM, y para poder llevar a cabo la fijación la bacteria requiere de un aporte importante de energía, unos 18 ATPs/N₂, la cual se obtiene mediante la respiración aerobia del microsimbionte.

Al utilizar como herramienta a los genes *hb* de *Parasponia*, Bogusz et al. (16) pudieron detectar tanto a los genes como a los transcritos de

Hb en *Trema*, y posiblemente también en *Celtis*, dos géneros que están emparentados con *Parasponia* y que pertenecen a la misma familia Ulmaceae. El descubrimiento de que estos genes se expresan predominantemente en la porción meristemática del tejido radical de estas plantas, es decir, en los últimos 0.5 mm del ápice de la raíz, condujo a Appleby et al. (17) a sugerir que esta Hb puede funcionar como un "sensor" de los niveles internos del O₂ para determinar el tipo de metabolismo que presenta este tejido: oxidativo o fermentativo. Sin embargo, hasta el momento no se han obtenido evidencias experimentales que apoyen esta proposición.

En el nivel genético, los genes *hb* de los animales y de las plantas se encuentran áltamente conservados: ambos están interrumpidos por dos intrones que se ubican exactamente en la misma posición. Sin embargo, los genes *hb* en las plantas contienen un intrón adicional que interrumpe al exón central. Resulta de mucho interés el hacer notar que este intrón adicional que está presente en los genes *hb* de las plantas fue predicho por Gó (18) al estudiar los mapas de las distancias atómicas de la subunidad β de la desoxiHb del caballo. Ella encontró que el mapa se divide naturalmente en cuatro unidades estructurales, y no en tres de acuerdo con el número de exones que presentan los genes *hb* en los animales, en donde el exón central podía ser interrumpido potencialmente por un intrón. Hasta este momento, todos los genes *hb* que se han estudiado en las plantas contienen un intrón adicional que interrumpe al exón central.

Como se mencionó, los genes *hb* se expresan predominantemente en los nódulos radicales y/o caulinares de las plantas fijadoras del nitrógeno. Estos órganos especializados son el resultado de los

productos de la expresión genética del micro y del macro-simbionte. Por parte de la planta se han detectado algunos flavonoides, como la luteolina y la naringenina (19), que tienen efecto activador sobre los genes *nod* de la bacteria, quienes son los responsables de que la nodulación se lleve a cabo. Recientemente, Dénarié *et al.* (en 20) purificaron el factor NodRm-1 de *Rhizobium melloti*, el cual se sintetiza cuando los genes *nod* son funcionales en la bacteria. El análisis químico ha mostrado que NodRm-1 es un $\beta(1-4)$ tetrasacárido sulfatado de D-glucosamina en el que tres grupos amino están acetilados, y uno está acilado con un ácido graso insaturado (C16) (20). La interacción de este tipo de señales químicas conduce a la formación de un nódulo funcional.

Durante los procesos ontogenéticos que se presentan en la formación del nódulo, se expresan de manera más o menos específica una serie de genes de la planta, a los cuales se les conoce como "nodulinas" debido a la especificidad en cuanto al tejido en el que se sintetizan: el nódulo. En el Apéndice I se encuentra una revisión actualizada sobre estas proteínas. Aunque es cierto que la expresión de las nodulinas depende en gran medida de activadores de origen bacteriano, recientemente se ha encontrado que la nodulina temprana SrEnod2 de *Sesbania rostrata* puede ser inducida por citocininas (21), lo que sugiere que las hormonas de la planta, así como sus niveles de síntesis, también regulan la expresión de los genes que codifican para las nodulinas. Además, es factible proponer que las señales de la bacteria pueden regular indirectamente la síntesis de las nodulinas a través del sistema hormonal de la planta, es decir, pueden funcionar como activadores o inhibidores de algunos componentes de la ruta

biosintética de las hormonas vegetales.

En los últimos años se han descrito una gran cantidad de nodulinas (22) , pero la "nodulina por excelencia" puede seguir siendo la Lb cuando se sintetiza en los nódulos, aunque esto se pone en duda con frecuencia debido a que también se le detecta en tejidos no simbióticos. La expresión espacio-temporal de la Lb, así como la de otras nodulinas, es áltamente específica: su síntesis se localiza en las células de las plantas que están infectadas por el micro-simbionte, y forma parte de las nodulinas tardías porque antecede a la expresión de los genes *nif* y por lo tanto a la actividad del complejo enzimático nitrogenasa (revisado en 14).

Por lo tanto, la expresión específica de los genes *hb* debe depender de las regiones reguladoras. Stougaard *et al.* (23) detectaron elementos *cis* dentro de la región promotora 5' del gen *lbc3* de la Lb de la Soya. Mediante el análisis en plantas transgénicas de *Lotus corniculatus*, así como por deleciones en esta región 5' y mediante la construcción de genes "quiméricos" (híbridos), que resultan de la fusión con un gen "reportador", en este caso CAT, estos autores encontraron un elemento regulador positivo fuerte (SPE, del inglés "Strong Positive Element") entre -1100 y -950, y uno débil (WPE, de "Weak Positive Element") entre -230 y -270; además, encontraron un elemento regulador *cis* entre -139 y -102 que se requiere para la expresión específica de este gen en el nódulo (OSE, de "Organ Specific Element") .

Estos mismos autores (23) no detectaron elementos de control en la región 3' no codificante. Sin embargo, al obtener la secuencia de varias clonas del cDNA de Lb del Haba, Kuhse y Pühler (24) encontraron una señal probable de poliadenilación, con la secuencia ATAAAT, que

está conservada entre las diferentes clonas de cDNA de la Lb del Haba, en donde se le encuentra en posiciones similares. Además, la comparación entre estas secuencias reveló la existencia de otra región altamente conservada con unas 16 pb entre +380 y +396, con la secuencia 5' ACTAAACTTGTTTAA 3'; la mayor parte de esta secuencia, 5' AA/GAACTTGT 3', también se encuentra conservada en otras Leguminosas. Aún cuando Stougaard *et al.* (23) no detectaron disminución alguna en la actividad de CAT, cuando se le fusionó con la región 3' flanqueante del gen *lbc3* de la Soya, lo que les condujo a proponer que esta región no tiene funciones reguladoras, Kuhse y Pühler (24) han sugerido que la porción conservada de 16 pb que se encuentra en el extremo 3' puede participar en la formación de una estructura secundaria con el extremo 5' del RNAm, ya que parte de esta región de 16 pb también se localiza en la porción codificante del gen, en +73 a +83.

Al estudiar extractos de diversos tejidos de la Soya, y mediante experimentos de retardamiento en geles, Jensen *et al.* (25) pudieron aislar un factor *trans* específico a partir de los nódulos, el cual se une a las regiones -223 a -246, el elemento 1, y -161 a -176, el elemento 2. Mediante experimentos de competencia, estos autores concluyeron que este factor se une a ambos elementos pero con afinidades diferentes. Recientemente, Bergman *et al.* (26) dedujeron la secuencia de amino ácidos a partir de los nucleótidos 1 a 276 del cDNA de la ureato oxidasa (*uox*); este polipéptido, el drp90, presenta dominios característicos de las proteínas que se unen al DNA. Mediante experimentos por retardamiento en geles, los autores encontraron que drp90 interactúa con las regiones 5' y 3' de *lbc3* y N23, pero con mayor afinidad con el oligonucleótido sintético de 24 pb que corresponde al

elemento 1 de *lbc3*. Este péptido, *drp90*, que se codifica por el gen *uox*, puede corresponder a un factor regulador que está implicado en la expresión de estos genes.

Sin embargo, a diferencia de los ejemplos anteriores, en algunas nodulinas la regulación es por factores extrínsecos a la planta, como en el caso de la uricasa, o N35, y la sacarosa sintetasa, o N100, para quienes los niveles del RNAm responden a fluctuaciones en la concentración del O₂: cuando disminuye la cantidad del O₂ disuelto, como ocurre en condiciones naturales cuando se desarrolla el nódulo, los RNAm de estas nodulinas aumentan entre 4 y 5 veces los niveles basales que se observan con la concentración atmosférica del O₂ (27).

La regulación de la síntesis de la Lb también puede ser en el nivel post-transcripcional. Los trabajos de Jensen *et al.* (28) muestran que al utilizar fusiones de las regiones 5' y 3' flanqueantes del gen *lbc3* con CAT y NPTII en la mutante TMI de la levadura (*Saccharomyces cerevisiae*), el grupo hemo regula específicamente al estimular la actividad de estos genes reporteros en el nivel post-transcripcional, probablemente al interactuar con la región 5' no codificante de *lbc3*.

Al comparar los elementos reguladores *cis* entre los genes *lb* de la Soya y de *Sesbania rostrata*, Metz *et al.* (29) encontraron una similitud alta entre ambos elementos, los que aparentemente interactúan con el mismo factor *trans* (¿universal?) que se encuentra en extractos de los nódulos. Con base en estas observaciones, estos autores sugieren que el mecanismo de regulación de los genes *lb* puede estar muy conservado en las Leguminosas.

Los genes *lb* en la Soya comparten elementos reguladores con otras nodulinas. Cuando se llevan a cabo deleciones en el elemento OSE (-139,

-102), que contiene a las secuencias consenso para nodulinas (5' AAAGAT y 5' CTCTT), la actividad del gen reportero CAT disminuye significativamente ; este mismo efecto se observa también en los estudios con promotores híbridos que contienen a las regiones reguladoras de la nodulina N23 de la Soya (30). Esto sugiere que los elementos reguladores pueden estar muy conservados entre los genes de las nodulinas, aunque es necesario estudiar de manera comparativa la secuencia de estas regiones flanqueantes en 5' para poder conocer el panorama general de los mecanismos de regulación de estos genes.

Adicionalmente a la presencia de los elementos *cis* y a los factores *trans* que se describieron con anterioridad para los genes *hb* de las plantas, se han localizado una serie de regiones potenciadoras ("enhancers") en la familia de genes *hb* de los vertebrados las cuales estimulan la expresión de estos genes (31). Según el gen que se trate, estas regiones potenciadoras se pueden localizar en los extremos 5' ó 3' no codificantes, e inclusive en la región codificante. La presencia de estas secuencias potenciadoras en los genes *lb* es una posibilidad que no se ha explorado hasta este momento; resultaría de mucho interés el localizar estas secuencias, así como el estudiar su efecto en la expresión de los genes *hb* durante el desarrollo del nódulo.

Como se mencionó, los genes *hb* de las plantas se encuentran altamente conservados entre diferentes especies de Leguminosas (32), pero a juzgar por los resultados que se describen en los párrafos anteriores, también lo están los elementos *cis* y los factores *trans* que regulan la expresión de estos genes. Adicionalmente, esto se ha podido comprobar ya que Stougaard *et al.* (33) lograron transferir el gen *lbc3* de la Soya completo a *Lotus corniculatus* al usar como vector a

Agrobacterium rhizogenes; como resultado se obtuvo la expresión específica de *lbc3* en los nódulos, en donde los niveles de síntesis alcanzaron los niveles endógenos de la Lb original de *Lotus*. Por lo tanto, dado el grado de conservación que se observa tanto en la región codificante como en las regiones reguladoras, se cree que los genes *hb* de las plantas descienden del mismo estado ancestral, por lo menos dentro de la familia Leguminosae; además, las relaciones inmunológicas entre las Lbs reflejan las relaciones de parentesco que existen entre algunas de las tribus de la subfamilia Papilionoideae de las Leguminosae (34).

Brown *et al.* (35) han propuesto que la tasa de variación entre los residuos de las Hbs de los vegetales es mayor que en las de los animales. En una revisión reciente, Ohta (36) propone que las copias que resultan de genes que se han duplicado recientemente presentan una tasa de sustitución alta, mucho mayor que la del gen ancestral; esto se debe a que la presión de selección es menor en las copias, sin embargo, esta presión aumenta con el transcurso del tiempo y así la variación tiende a disminuir. Si este caso general se puede aplicar a los genes *hb* de las plantas, entonces la tasa de variación que observan Brown *et al.* (35) puede ser el resultado de fenómenos de duplicaciones genéticas recientes, tal vez durante los últimos 60 a 70 x 10⁶ años; Lee y Verma (37) han propuesto que el gen *hb* ancestral de la tribu Phaseoleae se duplicó antes de la divergencia entre la Soya y el Frijol para generar una familia de 4 genes *hb*.

Debido a que durante mucho tiempo las Lbs sólo se conocieron dentro del grupo de las Leguminosae, se pensó que estas Hbs de vegetales tuvieron su origen en las Hbs de los insectos, las cuales se

transfirieron horizontalmente mediante una partícula viral, como vector, hacia una Leguminosa ancestral (38). Sin embargo, esta proposición se descartó cuando se determinó la secuencia de los genes *hb* del insecto *Chironomus thumii* (39), la cual carece de intrones; debido a que los genes *hb* de las plantas presentan tres intrones, se esperaba un número parecido en los insectos en el caso de que estos últimos hubieran sido los donadores de los genes *hb* de las plantas.

En 1960, Davenport (40) reportó que en los nódulos de algunas plantas actinorrizas se detecta un contenido igual o mayor de hematina que en las Leguminosas, lo que sugiere que estas plantas también pueden sintetizar Hbs. Estos resultados no se confirmaron durante mucho tiempo debido, sobretudo, a que en los extractos de estos nódulos existe una actividad alta de fenol oxidasas, así como concentraciones considerables de la peroxidasa (41): las primeras sintetizan quinonas que tienen una afinidad alta por las biomoléculas en general, y dan lugar a pigmentos pardos; y las segundas son contaminantes importantes porque comparten muchas de las características espectrales de las Hbs. Por estas razones se requieren condiciones extremas de anaerobiosis y métodos espectroscópicos de alta sensibilidad para poder detectar Hbs en estas plantas (C.A. Appleby, comunicación personal).

Mediante estudios de espectroscopia de preparaciones de nódulos de plantas actinorrizas, Tjepkema (42) logró obtener máximos de absorción a 560 (en una atmósfera de N₂), 540 y 580 nm (en una atmósfera de O₂), además de 540 y 570 nm bajo CO, los cuales son característicos de las Hbs. Posteriormente, Hattory y Johnson (43) y Roberts *et al.* (44) lograron obtener señales después de hibridizar una sonda de *hb* de la Soya, que se obtuvo a partir del cDNA, con los DNAs digeridos de las

plantas actinorrizas, y así confirmar el descubrimiento de Tjepkema. Estos resultados permitieron ampliar el espectro de distribución de los genes *hb* a las plantas actinorrizas.

Por otro lado, *Parasponia*, una planta que pertenece a la familia Ulmaceae, es de interés ya que es nodulada por cepas de *Rhizobium*, que también nodulan a las Leguminosas. En 1983, Appleby *et al.* (45) lograron purificar la Hb de los nódulos radicales de esta planta por medio de procedimientos electroforéticos, como el enfoque isoeléctrico. Esta proteína es una forma dimérica de Hb soluble que presenta un pI de 6.28 en la forma oxi, y de 6.67 en la forma férrica, en comparación con 4.59 y 4.79 para las formas oxi y férrica de la Lb del Chicharo, respectivamente. La purificación de esta Hb permitió, mediante procedimientos genéticos, la obtención de la sonda pL29, que se deriva del cDNA de *Parasponia*, la cual fue de utilidad para la detección de los genes *hb* en otra planta ulmacea: *Trema*. Con anterioridad se había intentado la detección de genes *hb* en esta última planta mediante el uso de sondas específicas de *hb* que se obtuvieron de Leguminosas, pero nunca fue posible detectar señales de hibridización debido a los fenómenos de divergencia (C.A. Appleby, comunicación personal).

El descubrimiento de los genes y transcritos de Hb en *Trema* (46, 47) resultó de mucho interés debido a que este fue el primer caso demostrado de la presencia de estos genes en una planta que no nodula, y por lo tanto que no fija el nitrógeno, de tal manera que esto condujo a pensar que estos genes son comunes a todas las plantas con flores, y que comparten el mismo ancestro con las Hbs de los animales, tal vez desde hace más de 600×10^6 años (48).

Los genes *hb* se encuentran conservados en las plantas (32); esto ha

permitido construir árboles evolutivos para estas Hbs (49), los cuales concuerdan bien con las relaciones filogenéticas que se presentan entre las plantas que las sintetizan. Además, y de manera muy interesante, los elementos *cis* que regulan la expresión de los genes *hb* en las no-Leguminosas *Parasponia* y *Trema* también se encuentran conservados, ya que cuando se introducen las regiones 5' y 3' no codificantes fusionadas con GUS en *Lotus corniculatus*, se detecta la actividad de β -glucuronidasa en los nódulos después de infectar a *Lotus* con su simbiote *Rhizobium loti*; en cambio, cuando se transforma a *Nicotiana tabacum*, la expresión de la actividad se presenta sólo en el meristemo de la raíz (50). Esto demuestra nuevamente que los mecanismos de regulación de estos genes, sean elementos *cis* o factores *trans*, están muy conservados en las plantas, así como también la característica de la expresión específica en los nódulos o el meristemo radical, según el caso, ya que no se observa actividad de β -glucuronidasa en otros tejidos de la planta.

Aún cuando existe la creencia generalizada de que los genes *hb* son comunes a todas las plantas, y que las Hbs animales se han estudiado con mucho detalle, los árboles que reflejan la filogenia de estos genes todavía no muestran la raíz, o el estado ancestral, común a las Hbs de plantas y animales (51). En este sentido, resulta de mucho interés la presencia de una Hb dimérica en la bacteria ferruginosa *Vitreoscilla*, que pertenece a la familia *Beggiatoa*, cuyo gen ya se ha clonado y secuenciado (52).

Aunque divergente, la secuencia de esta Hb bacteriana presenta una similitud mayor con las Hbs vegetales que con las animales (32), lo que implica un nivel de parentesco mayor entre las Hbs de las plantas y de

las bacterias, que las de éstas últimas con las de los animales. Resulta de interés el notar que la región 3' no codificante de este gen *hb* bacteriano tiene similitud con la región correspondiente en el gen *hb* de las plantas (53); esto descarta la posibilidad de convergencia entre ambos genes. Por lo tanto, dadas estas similitudes, el estado ancestral común a los genes *hb* de las plantas y de los animales pudo ser el de algún procarionte, del tipo *Vitreoscilla*, y así el origen de estos genes puede ser mayor a los 600×10^6 años.

Riggs (54), en una excelente revisión acerca del origen y la evolución de las Hbs, propone que las Hbs de las plantas y de los animales divergieron hace unos 1×10^9 años, de donde el linaje de las Hbs animales perdió el intrón central en alguna época temprana de su evolución, lo que, sin embargo, requiere de confirmación al encontrar, por ejemplo, la combinación 4 exones/3 intrones en los genes *hb* de los nemátodos. Además, el mismo autor sugiere que las Hbs pudieron haber surgido en alguna época temprana poco después del origen de la fotosíntesis: *Chromatium* es un anaerobio obligado que sintetiza una hemo-proteína de 11 KDa de PM que capta al O_2 , ya que contiene un grupo hemo del tipo *c* (no un protohemo IX como las globinas), la cual puede funcionar como detoxificadora del O_2 . Se ha sugerido, además, que estas posibles Hbs primitivas pudieron funcionar en sus inicios como oxidasas terminales ya que se ha encontrado que *Alcaligenes eutrophus*, una bacteria que metaboliza al hidrógeno, contiene una Hb de 43 KDa de PM con dominios tanto para la flavina como para el grupo hemo: se combina reversiblemente con el O_2 y se reduce directamente con el NADH; esta flavo-hemoglobina puede reducir al citocromo *c* y utilizar al NADH como donador de electrones. Así, la estructura de estas proteínas pudo

resultar de la fusión de genes distintos o de exones diferentes que codifican para dominios específicos ("exon shuffling" [55]), que da como resultado proteínas quiméricas. En este sentido, Runnegar (56) propuso que las globinas se derivan parcialmente de una proteína semejante al citocromo *b5* que está presente en bacterias que son estrictamente anaerobias, en donde el segundo exón de la globina se deriva de la región del citocromo *b5* que codifica para la histidina distal.

Appleby *et al.* (57), también en una excelente revisión, concuerdan con Riggs en el sentido de que las Hbs pudieron haber surgido hace unos 1.5×10^9 años, ya que se calcula que la transición de una atmósfera microaerobia a una aerobia ocurrió hace unos 1.5 a 2.0×10^9 años, y que durante ese lapso aparecieron los linajes principales de los eucariontes, en uno de los cuales ya existía el proto-organismo que contenía Hb y que en el futuro llegaría a ser el ancestro común de las plantas y de los animales.

Se conocen fósiles de estromatolitos con 3.4×10^9 años de antigüedad, y como lo establece Walker *et al.* (58) "...resulta difícil el imaginar cómo pudo surgir la fotosíntesis en microorganismos que no están pre-adaptados para tolerar al O_2 , que es un subproducto de dicho proceso metabólico..."; con base en esta suposición, las Hbs pudieran tener una antigüedad mayor a los 1.5×10^9 años en el caso de haber funcionado como un mecanismo de adaptación al O_2 . Walker *et al.* (58) calcularon que en el océano del Arqueano la concentración del O_2 que se producía por efecto fotoquímico era alrededor de 1 Pa, o unos 14 nM de O_2 disuelto; con base en este cálculo, Appleby *et al.* (57) sugieren que la afinidad tan alta que presentan las Hbs de las plantas

hacia este elemento es un reflejo de lo que fue su estado ancestral, ya que ellas funcionan con ppO_2 similares en los tejidos vegetales, de 10 nM; luego, la pérdida de la afinidad fue un carácter evolutivo posterior que se presentó en el linaje de las Hbs de los animales.

En resumen, los genes *hb* de las plantas y de los animales probablemente evolucionaron por líneas verticales a partir del mismo ancestro: un procarionte que surgió y proliferó en el momento en el que la naturaleza de la atmósfera del Arqueano cambió de microaerofílica a aerobia. Esta Hb ancestral pudo funcionar como un mecanismo de adaptación al O_2 . Sin embargo, dadas las evidencias experimentales con las que contamos actualmente en el reino Plantae, la mayor antigüedad de los genes *hb* corresponde al origen de las plantas con flores, en particular al origen de las Dicotiledóneas que es en donde se han detectado estos genes, lo cual ocurrió hace unos 100×10^6 años, durante el Cretácico Inferior (59).

II. OBJETIVOS.

Actualmente existe consenso en aceptar que las Gimnospermas y las Angiospermas son grupos monofiléticos (60), así como lo son las Mono y las Dicotiledóneas dentro de las Angiospermas (61, 62); por lo tanto es necesario el estudiar los genes *hb* en estos grupos de plantas para visualizar con mayor claridad la distribución y evolución de estos genes en el reino Plantae. En el pasado se han hecho intentos para detectar los genes *hb* en algunas plantas Monocotiledóneas, tal como el Maíz, mediante el uso de sondas de *hb* específicas, sin embargo, estos intentos no han tenido éxito debido, probablemente, a fenómenos de divergencia (C.A. Appleby, comunicación personal). Por esta razón, en el presente trabajo se propuso una estrategia distinta a la del uso de las sondas específicas, la cual consistió en la construcción de una secuencia consenso (¿universal?) para los genes *hb* en las plantas, de la que se seleccionaron las regiones más conservadas que se utilizaron como sondas moleculares para experimentos de hibridación con los DNAs de las plantas Monocotiledóneas y de otras cuya situación filogenética es de interés.

III. MATERIALES Y METODOS.

ELABORACION DE LA SECUENCIA CONSENSO PARA GENES *hb* EN PLANTAS. En el Apéndice 2 se presenta el detalle de los procedimientos que se realizaron para construir una secuencia consenso para las Hbs de las plantas, además se muestran las secuencias originales, el alineamiento y los programas que se utilizaron para el análisis por computadora de dicha secuencia. A continuación se presentan brevemente estos procedimientos.

Se compilaron 16 secuencias polipeptídicas y 6 secuencias nucleotídicas de Hbs de diferentes especies vegetales (Figuras 1 y 2 del Apéndice 2, respectivamente). Estas secuencias se alinearon mediante el uso de la rutina MATCH del programa NEWAT'85, que fue donado generosamente por el Dr. Russell Doolittle (Universidad de California en San Diego, EUA).

De la comparación entre las diferentes secuencias se obtuvo una secuencia consenso para el polipéptido y otra para el gen, de las cuales se construyó la secuencia consenso total en la que se detectaron las regiones conservadas, variables e hipervariables (ver la Figura 3 del Apéndice 2). Como resultado de este análisis se seleccionaron las regiones más conservadas con el objeto de sintetizar una sonda que sería de utilidad en los experimentos de hibridización posteriores.

Además, este mismo programa, NEWAT'85, se utilizó para caracterizar con mayor detalle la secuencia consenso total de las Hbs de plantas (SCTHBP): se obtuvo la composición total de amino ácidos con la rutina COMP, el perfil de hidropatía con GREASE, y la similitud con otras secuencias que se encuentran en el banco de datos del NEWAT'85 con SCANA.

OBTENCION DEL DNA TOTAL DE PLANTAS SUPERIORES. Se utilizaron las siguientes especies de plantas como fuente del DNA. Leguminosas: *Medicago sativa* (Alfalfa), *Phaseolus vulgaris* (Frijol), *Lens* sp. (Lenteja), *Glycine max* (Soya) y *Erythrina* sp. (Colorín); Casuarinales: *Casuarina cunninghamiana* (Casuarina); Gramíneas: *Oryza* sp. (Arróz), *Zea mays* ssp. *mays* (Maiz), *Zea mays* ssp. *mexicana* (Teosinte), *Sorghum* sp. (Sorgo), y *Triticum sativum* (Trigo); otras: *Magnolia* sp. (Magnolia) y *Zamia furfuracea* (Cicada).

Con excepción de la Casuarina y de la Magnolia, la fuente del DNA para las demás especies fue el tejido embrionario, el cual se obtuvo después de germinar las semillas en cajas de Petri que contenían papel Whatman No. 3 que se humedeció con agua destilada; estas semillas se incubaron a 30°C hasta que el hipocótilo alcanzó unos 2 a 3 cm de longitud, y entonces se almacenaron a -70°C hasta su uso. En el caso de la Casuarina y de la Magnolia, la fuente de DNA fue el tejido foliar el cual se utilizó directamente para los procedimientos de extracción.

La obtención del DNA total se realizó de acuerdo con el procedimiento descrito por Doyle y Doyle (63). Se homogenizaron de 0.5 a 1.0 g de tejido vegetal en un mortero con un pistilo, bajo nitrógeno líquido, hasta obtener un polvo fino, el cual se disolvió en 5 a 7.5 ml del amortiguador de aislamiento (CTAB al 2% [p/v], NaCl 1.4 M, β -mercaptoetanol al 0.2% [v/v], EDTA 20 mM, Tris-HCl 100 mM, pH=8.0) precalentado a 60°C; se incubó la mezcla en esta temperatura durante 30 min con agitación ocasional.

Se extrajo la mezcla con un volúmen de cloroformo:alcohol isoamílico (24:1) y se centrifugó a 1600 x g a temperatura ambiente (TA). Se recuperó la fase acuosa, se le adicionaron 2/3 de volúmen de

isopropanol frío, se agitó y se centrifugó a 5000 x g por 2 minutos. Se lavó el botón durante 20 min con etanol al 76% (v/v)-acetato de amonio 10 mM. Se centrifugó la solución a 1600 x g durante 10 min a TA; se desechó el sobrenadante y se secó el botón.

Se resuspendió el botón en 1 ml de TEB (Tris-HCl 10 mM, EDTA 1 mM, pH=8.0) y se le adicionó RNasa de páncreas de bovino (Calbiochem) a una concentración final de 10 µg/ml; la solución se incubó a 37°C durante 30 minutos. Al cabo de ese tiempo se adicionaron 2 volúmenes de TEB, acetato de amonio 7.5 M a una concentración final de 2.5 M, y 2.5 volúmenes de etanol frío. Se mezcló por inversión y se recuperó el DNA con una varilla de cristal.

Se disolvieron las hebras del DNA que se adhirieron a la varilla de cristal en 1 ml de TEB, y se dializó la solución contra este mismo amortiguador durante 16 h a 4°C.

Se caracterizó el DNA que se obtuvo por este procedimiento de la siguiente manera: a) mediante el perfil de absorción en el intervalo de 200 a 300 nm; b) la determinación de los cocientes R260/280 y R260/230; c) cuantificación del DNA por la absorción a 260 nm, considerando el coeficiente de extinción molar (ϵ) del DNA como 13200 M⁻¹ cm⁻¹ (64); y d) mediante el corrimiento electroforético en un gel de agarosa al 1% (p/v) en TBES (Tris-HCl 89 mM, ácido bórico 89 mM, EDTA 2.5 mM, pH=8.0), el cual se tizó con bromuro de etidio (5 µg/ml) durante 3 min, se destizó con agua y se visualizó en un transiluminador.

Cuando el DNA de la Casuarina se obtuvo a partir del tejido foliar, mediante el mismo procedimiento, la preparación siempre acarreo un pigmento pardo el cual podía interferir con los ensayos de hibridización posteriores, por lo que se intentó separar del DNA el

pigmento mediante los siguientes procedimientos: a) adición de PVP a una concentración final del 1% (p/v); b) adición de EDTA a una concentración final de 50 mM; c) aumento en cinco veces, 5x, de la cantidad original del β -mercaptoetanol en el amortiguador de extracción; d) adición de BSA a una concentración final del 0.1% (p/v); e) diálisis contra 500 ml de TES durante 24 h a 4°C; f) extracción con un volumen de cloroformo: alcohol isoamílico (24:1); y g) filtración a través de una columna de Sephadex G-75.

DIGESTION DEL DNA TOTAL CON ENZIMAS DE RESTRICCIÓN. Se digirieron los DNAs totales con las enzimas *EcoRI*, *HindIII* o *PstI* (Sigma) en la siguiente mezcla de reacción: 2 μ l del DNA (1 μ g/ μ l), 2 μ l del amortiguador de la enzima (ver abajo), 2 μ l de espermidina 10 mM, 1 μ l de β -mercaptoetanol al 0.2% (v/v), 10 U de la enzima, y agua destilada estéril a un volumen final de 20 μ l. La digestión se incubó a 37°C durante toda la noche, y luego se paró la reacción al incubar a 65°C durante 3 minutos. Para verificar la digestión, se corrieron 5 μ l de la solución en un gel de agarosa al 1% (p/v), o de poliacrilamida al 5% (p/v), los que se tificaron posteriormente con bromuro de etidio. El digerido restante se mantuvo a -70°C hasta su uso.

Los amortiguadores que se utilizaron para cada enzima fueron los siguientes. *EcoRI*: Tris-HCl 100 mM, MgCl₂ 10 mM, NaCl 50 mM, BSA (100 μ g/ml), pH=7.5; *HindIII*: Tris-HCl 50 mM, MgCl₂ 10 mM, NaCl 50 mM, BSA (100 μ g/ml), pH=8.0; y *PstI*: Tris-HCl 10 mM, MgCl₂ 10 mM, NaCl 100 mM, BSA (100 μ g/ml), pH=7.5.

SINTESIS DEL OLIGONUCLEOTIDO (SONDA). Se seleccionó al fragmento que codifica para una porción de la α -hélice H de la Hb de las plantas, ya que es la más conservada (residuos de amino ácidos 143 a 153 [ver la

Figura 3 del Apéndice 2)). El fragmento corresponde a un oligonucleótido de 33 pb, el cual fue amablemente sintetizado por los Dres. Franz J. deBruijn (Universidad de Michigan, EUA) y Peter Welters (Instituto Max-Planck, Alemania) con la siguiente secuencia: 5' GTA GCC TAT GAT GAA TTG GCA GCT GCA ATT AAG 3' (sonda 1913). La cadena complementaria (sonda 1914) se sintetizó con sitios para *SphI* y *PstI* (se indican con minúsculas en la secuencia), respectivamente: 5' catg CTT AAT TGC AGC TGC CAA TTC ATC ATA GGC TAC tgca 3'. A esta sonda se le denominó SCHBP/H (Secuencia Consenso de las Hbs de Plantas/región H).

CLONACION DE LA SONDA SCHBP/H. Los procedimientos que se realizaron fueron los descritos en Sambrook et al. (65).

a) Alineamiento de las sondas 1913 y 1914. Se obtuvo el valor de T_m de la sonda SCHBP/H (67.1°C) el cual se utilizó para el apareamiento de las dos cadenas. Se disolvieron las dos sondas, 1913 y 1914, en 50 μl de agua destilada estéril, se calentaron a 67°C durante 5 min, y se dejaron enfriar a TA lentamente; luego se mantuvo la solución así durante 1 hora. Se separaron alícuotas que se congelaron a -70°C hasta su uso. La concentración final fue de 16 μg de SCHBP/H/ μl de agua destilada.

b) Linearización del vector. El vector, pUC19, se linearizó con *PstI* y *SphI* en la siguiente mezcla de digestión: 2 μl del amortiguador para *SphI* 10x (Tris-HCl 10 mM, NaCl 100 mM, MgCl_2 7 mM, 2 μl de β -mercaptoetanol 7 mM, pH=7.5), 2 μl de espermidina 10 mM, 1 μl de β -mercaptoetanol al 2% (v/v), 20 U de *PstI*, 10 U de *SphI*, y agua destilada estéril a un volumen final de 20 μl . Se incubó la digestión a 37°C durante toda la noche, y al cabo de ese tiempo se corrió 1 μl en un gel de agarosa al 1% (p/v). Después de verificar la digestión, se

paró la reacción al incubar a 65°C durante 3 minutos.

c) Ligamiento pasajero-vector. Se utilizó una proporción 5:1 en la relación pasajero (SCHBP/H)-vector (pUC19) para el ligamiento. La mezcla de reacción fue la siguiente: 7 µl del DNA pasajero (de la dilución 1:100 que resultó en una concentración final de 1.6 µg de SCHBP/H/µl), 1 µl del DNA vector, 2 µl del amortiguador 10x para la ligasa de T4 (Tris-HCl 50 mM, MgCl₂ 10 mM, ATP 1 mM, BSA [100 µg/ml], pH=7.8), 2 µl de espermidina 10 mM, 2.5 U de la ligasa de T4 (Sigma), y agua destilada a un volumen final de 20 µl. Se incubó la mezcla a 14°C durante toda la noche y se paró la reacción al incubar a 65°C durante 3 minutos.

d) Transformación de *E. coli* CMK. Se inocularon 100 ml del medio LB (por litro: 10 g de bacto-triptona, 5 g de extracto de levadura, 10 g de NaCl, pH=7.5 [para medio sólido se adicionan 15 g de agar]) con 1 ml de un medio que se creció previamente durante toda la noche, se incubó a 37°C durante 3 h o hasta alcanzar una DO de 0.4 a 0.5, a 650 nm, y se colocó el cultivo en hielo durante 10 minutos. Al cabo de ese tiempo se centrifugó el medio a 2000 x g durante 5 min a 4°C.

Las células se hicieron competentes al resuspenderlas en 50 ml de CaCl₂ 50 mM fríos, y se mantuvieron en hielo durante 15 minutos. Posteriormente, se centrifugó la solución a 2000 x g durante 5 min a 4°C; se resuspendió el botón en 3 a 5 ml de CaCl₂ 50 mM fríos y se tomaron alícuotas de 0.8 ml.

Se mezclaron de 5 a 10 µl de la solución que contenía al DNA pasajero-vector (pUC19::SCHBP/H, al que ahora se le denomina pHBGS/H) con 150 µl de las células competentes, y se mantuvo la solución en hielo durante 30 minutos; al cabo de ese tiempo, se dió un choque

térmico al incubar a 37°C durante 5 minutos. Se añadió 1 ml del medio LB y se creció a 37°C durante 1 hora.

Se sembraron 150 µl de las células transformadas en medio LB selectivo (medio LB, Amp [50 µg/ml], IPTG [25 µg/ml], x-Gal [25 µg/ml]). Además, se creció un control negativo con células no transformadas, y un control positivo con células transformadas con el pUC19. Se incubaron las placas a 37°C durante 24 horas. Se recuperaron las colonias blancas (Amp^r, x-Gal⁻) por punción y se resembraron por estría en el medio selectivo. Los palillos con los que se picaron las colonias blancas se colocaron en 1 ml del medio selectivo y se incubaron a 37°C durante 24 h con agitación vigorosa para obtener posteriormente los plásmidos.

e) Obtención de plásmidos por minipreparaciones (66). Se tomaron 1.5 ml del medio que se creció por toda la noche y se colocaron en un tubo Eppendorf para centrifugar posteriormente en una microfuga durante 1 minuto. Se descartó el sobrenadante y se resuspendió el botón en 100 µl de una solución fría de glucosa 50 mM, EDTA 10 mM, Tris-HCl 25 mM (pH=8.0), lisozima (4 mg/ml); se mantuvo a TA durante 5 minutos. Al cabo de ese tiempo se agregaron 200 µl de una solución fría y recién preparada de NaOH 0.2 N - SDS al 0.1% (p/v); se mezcló por inversión, 2 ó 3 veces, y se mantuvo la solución en hielo durante 5 minutos. Se agregaron 150 µl de una solución fría de acetato de potasio que se preparó de la siguiente manera: se mezclaron 60 ml de acetato de potasio 5 M con 11.5 ml de ácido acético y 28.5 ml de agua destilada; se ajustó el pH a 4.8. Se agitó suavemente por 10 s y se colocó el tubo en hielo durante 5 minutos. Se centrifugó la solución en una microfuga a 4°C durante 5 min, se recuperó el sobrenadante y se desproteinizó

adicionando 1 volúmen de cloroformo-alcohol isoamílico (24:1); se centrifugó en una microfuga durante 2 min y se recuperó la fase acuosa a la que se le agregaron dos volúmenes de etanol absoluto para precipitar el DNA plasmídico al centrifugar por otros 5 min adicionales. Se secó el botón y se resuspendió en 50 μ l de TES que contenía RNasa (20 μ g/ml), se agitó y se incubó a 37°C durante 1 hora.

f) Liberación de la sonda SCHBP/H del plásmido pHBCS/H. Se linearizaron los plásmidos pHBCS/H que se obtuvieron por minipreparaciones utilizando las enzimas *Sph*I y *Pst*I de acuerdo con los procedimientos que se describieron anteriormente. Se corrieron 10 μ l de la solución de digestión en un gel de poliacrilamida (gel concentrador al 3% [p/v]; gel resolvidor al 7% [p/v]) de acuerdo con el sistema discontinuo de Davis (67). Se tñó el gel con bromuro de etidio durante 30 min, y se destñó con agitación en agua destilada durante el mismo tiempo.

MARCAJE NO RADIOACTIVO DE LA SONDA SCHBP/H. Se utilizó el sistema de "Nick Translation" que incorpora nucleótidos biotinilados, dATP-biotina, mediante el uso del kit "BioNick Labeling System" (GIBCO BRL). Los procedimientos fueron de acuerdo con las intrucciones del fabricante.

Se marcaron 1.6 μ g de la sonda SCHBP/H, de doble cadena, en la siguiente mezcla de reacción: 5 μ l de la solución 10x de dNTPs (que contiene dATP-biotina), 1 μ l de la sonda (1.6 μ g del DNA), 5 μ l de la solución de enzimas (con la DNasa y el fragmento Klenow de la DNA polimerasa I) y agua destilada para un volúmen final de 45 μ l. Se mezcló centrifugando en una microfuga durante 5 s, y se incubó a 16°C durante 1 hora. Al cabo de ese tiempo, se adicionó 1 μ l de EDTA 300 mM

para detener la reacción, y se separaron los nucleótidos que no se incorporaron al precipitar la sonda de la siguiente manera: se añadió 1/10 de volúmen de acetato de sodio 3 M y dos volúmenes de etanol absoluto, se mezcló por inversión y se mantuvo a -70°C durante 20 minutos; al cabo de ese tiempo, se centrifugó en una microfuga a 4°C durante 15 min, y se descartó el sobrenadante para resuspender el botón, que contuvo a la sonda marcada, en 50 μl de TE8. La sonda se almacenó a -70°C hasta su uso.

HIBRIDIZACION DEL TIPO "DOT" Y "SLOT-BLOT". Los procedimientos fueron siguiendo el protocolo del fabricante del kit de detección "BluGene" (GIBCO BRL). Se aplicaron 20 μg del DNA total de cada especie a tiras de nitrocelulosa (NC) (MicroFiltration Systems, 0.45 μm). Las Leguminosas fueron los controles positivos ya que se sabe que contienen genes *hb*; se utilizaron los DNAs de testículos del salmón (Sigma) y de timo de ternera (Merck) como controles negativos. La aplicación de las muestras a la NC fue mediante el uso del aparato para "Dot" y "Slot-Blot" (Milliblot-D, Millipore) utilizando vacío.

Se incubó el filtro en un horno a 80°C durante 2 h para fijar al DNA a la NC, y al cabo de ese tiempo se rehidrató el filtro en una solución de SSC 2x (SSC 20x es: NaCl 3 M, citrato de sodio 0.3 M, pH=7.0) y se prehibridizó a 42°C durante 4 h con la solución de prehibridización (SSC 2x, Denhardt 5x [Denhardt 50x es: ficoll al 1% (p/v), PVP al 1% (p/v), BSA al 1% (p/v)], SDS al 0.1% [p/v], DNA acarreador [500 $\mu\text{g}/\text{ml}$]). Posteriormente se hibridizó el filtro a 42°C durante toda la noche con la solución de hibridización que fue la misma que para la prehibridización pero a la cual se le añadió la sonda desnaturalizada en una concentración de 1.6 $\mu\text{g}/5$ ml.

Se lavó el filtro de acuerdo con la siguiente secuencia: SSC 2x/SDS al 0.1% (p/v) durante 3 min a TA, dos veces; SSC 0.2x/SDS al 0.1% (p/v) a TA, dos veces; SSC 0.16x/SDS al 0.1% (p/v) durante 15 min a 65°C, dos veces; y unos segundos en SSC 2x a temperatura ambiente. Se lavó el filtro con el amortiguador 1 (Tris-HCl 100 mM, NaCl 150 mM, pH=7.5) durante 1 min a TA, y luego se bloqueó con el amortiguador 2 (BSA al 3% [p/v] en el amortiguador 1) a 65°C durante 1 hora.

Finalmente, se incubó el filtro con el conjugado de estreptavidina - fosfatasa alcalina (1 µl de la solución stock [1 mg/ml]/ml del amortiguador 1) durante 10 min a TA, y entonces con el amortiguador 3 (Tris-HCl 100 mM, NaCl 100 mM, MgCl₂ 50 mM, pH=9.5) durante 10 min a temperatura ambiente. Se reveló incubando el filtro en la solución reveladora (12.5 µl de una solución stock de 50 mg de BCIP /ml, 16.5 µl de una solución stock de 75 mg de NBT/ml, 3.75 ml del amortiguador 3) de 30 min a 3 h a temperatura ambiente. Al cabo de ese tiempo se lavó el filtro con Tris-HCl 20 mM (pH=7.5)/EDTA 0.5 mM, y se secó a 80°C durante 2 minutos.

HIBRIDIZACION DIRECTA EN GELES DE AGAROSA. Se digirieron 10 µg de los DNAs totales con *Hind*III, y los fragmentos resultantes se separaron en un gel de agarosa al 1.5% (p/v) para hibridizar posteriormente con la sonda marcada. El procedimiento que se siguió fue una modificación al método de Gontijo *et al.* (68).

Después del corrimiento electroforético de los DNAs digeridos, se tñó el gel con bromuro de etidio y se fotografió. Se desnaturalizó y se despurinó el DNA al tratar el gel con HCl 0.25 M durante 10 min, dos veces, y con NaOH 0.5 M/NaCl 1.5 M durante 15 min, dos veces, respectivamente. Se neutralizó el gel al lavarlo con una solución de

Tris-HCl 1 M (pH=8.0)/NaCl 1.5 M durante 15 min, dos veces. Posteriormente, se secó el gel durante toda la noche mediante un proceso de transferencia doble ("Di-Blot") (69).

El gel seco se hibridizó con la solución de hibridación (SSC 6x, Denhardt 5x, SDS al 0.5% [p/v], Tris-HCl 20 mM [pH=8.0], y la sonda marcada [0.2 µg/ml]) a 42°C durante toda la noche. Posteriormente, se lavó el gel con SSC 6x/SDS al 0.1% (p/v) durante 10 min a TA, tres veces; una vez con esta misma solución pero a 42°C durante 2 min; una vez con SSC 6x/Tritón X-100 al 1% (v/v) durante 10 min a TA; y finalmente una vez con SSC 6x a temperatura ambiente.

Se incubó el gel con la solución del conjugado durante 48 h por lo menos, se lavó con el amortiguador 1 durante toda la noche y con agitación cuidadosa, y finalmente con el amortiguador 2 durante 1 hora. Se reveló el gel mediante el procedimiento que se describió para "Dot" y "Slot-Blot".

AMPLIFICACION POR PCR. Se sintetizaron dos pares de sondas ("primers") para amplificar dos fragmentos que comprenden a la región que se encuentra entre las α -hélices C y H, y la región entre las α -hélices A y H. Las sondas fueron: 510.3 (5' GG ctgcag TTG TTC TCG TTT CTG AAG 3' [las letras minúsculas indican la secuencia para *Pst*I]) y 511.3 (5' CTC gcatgc AGC CAA CTC ATC ATA AGC 3' [las letras minúsculas indican la secuencia para *Sph*I]), complementarias para las regiones C y H, respectivamente; y 399 (5' GG ctgcag CAA AAG CAA GAA GCT CTA GTG 3' [con una secuencia para *Pst*I]) y 398 (5' CTC gcatgc AGC CAA CTC ATC ATA AGC 3' [con una secuencia para *Sph*I]), complementarias para las regiones A y H, respectivamente. Las sondas 398 y 399 fueron sintetizadas por el Dr. Angel Zarain (Universidad de Manitoba, Canada),

y las sondas 510.3 y 511.3 se sintetizaron en el Instituto de Biotecnología de la UNAM por el Dr. Paul Gaytán.

La amplificación se llevó a cabo con los siguientes componentes: 1 μg del DNA genómico, sondas 480 μM (cada una), dNTPs (Boehringer) 400 μM , 1.25 U de AmpliTaq fragmento Stoffel (Cetus, Perkin Elmer), 18 μl del amortiguador de amplificación (Tris-HCl 10 mM, KCl 50 mM, MgCl_2 2 mM, pH=8.3), y agua destilada estéril para un volumen final de 25 μl . Se colocaron unos 30 μl de aceite mineral ligero (Sigma) sobre la solución, y se amplificó automáticamente en un "ThermoCycler" (Stratagene) de acuerdo con el siguiente perfil: 1 ciclo de desnaturalización a 95°C por 3 min; 45 ciclos a 37°C por 1.5 min para el alineamiento, 72°C por 2 min para la extensión, y 95°C por 1 min para la desnaturalización; finalmente se dió 1 ciclo a 72°C por 5 minutos; los cambios de temperatura se realizaron con una velocidad, RAMP, de 3 (según las indicaciones del fabricante).

Posteriormente se extrajo el aceite mineral con 50 μl de cloroformo, y se recuperaron 20 μl de la solución de amplificación, los cuales se corrieron en un gel de agarosa PE (Cetus, Perkin Elmer) al 3% (p/v) en TBES. El gel se tñó con bromuro de etidio y se fotografió en el transiluminador. Las bandas que se amplificaron se recuperaron, después de cortar el fragmento del gel, mediante electroelución (65). Los fragmentos que se recuperaron se precipitaron utilizando el procedimiento de acetato de sodio-etanol y se resuspendieron en 30 a 50 μl de TES.

IV. RESULTADOS Y DISCUSION.

ALINEAMIENTO Y ANALISIS DE LAS SECUENCIAS DE HBs. Durante los últimos años se han purificado y secuenciado las Hbs de diversas plantas; además se han clonado y secuenciado los genes *hb* gracias al desarrollo de las técnicas del DNA recombinante. Después de compilar la mayor parte de estas secuencias, que se han publicado previamente, éstas se alinearon con el objeto de construir una secuencia consenso en la que se pudieran detectar las regiones más conservadas (ver el Apéndice 2).

El alineamiento de las secuencias se realizó con la ayuda de la rutina MATCH del programa NEWAT'85 (68), sin embargo la mayor parte del procedimiento se llevó a cabo de manera manual ya que la mayoría de los programas actuales alinean por pares de secuencias debido a que los alineamientos múltiples ocupan demasiado espacio en la memoria de la computadora, consumen mucho tiempo, e incorporan una gran cantidad de indels (INserción/DELección) en el alineamiento (70). En el caso de las secuencias nucleotídicas y polipeptídicas de las Hbs de las plantas, no resultó complicado el alineamiento ya que estas secuencias están muy conservadas en lo general, y sólomente fue necesario incorporar unos cuantos indels (Figuras 1 y 2 del Apéndice 2).

Como resultado de los procedimientos del alineamiento se construyó una secuencia consenso para las Hbs de las plantas, la cual está formada por 483 nucleótidos y 161 residuos de amino ácidos de longitud (ver la Figura 3 del Apéndice 2). El residuo de amino ácido más abundante en la secuencia fue la Alanina, que correspondió al 13.04% del total de los residuos; el PM estimado de la secuencia fue de 17.481 KDa, lo cual concuerda con el PM de las Hbs monoméricas de las plantas.

La secuencia total presenta un 46.5% de residuos "degenerados" en el nivel de la proteína, y un 44% en el nivel genético; es interesante hacer notar que ambos valores son muy parecidos entre sí, lo que indica que el nivel de variabilidad es similar entre estos genes y sus productos. Generalmente se ha encontrado que la variabilidad es mayor en el nivel genético que en el proteico debido a la redundancia del código genético (71), sin embargo, esto parece no ser el caso para las Hbs vegetales, lo que resulta de interés ya que al estudiar la variabilidad de las Hbs en las poblaciones naturales de las plantas, por ejemplo por procedimientos como la detección de isoformas, es posible sugerir que la variabilidad que se observa en el nivel proteico es un reflejo de la variabilidad que esta presente en el nivel genético.

Se detectaron tres regiones que están muy conservadas en la SCTHBP: la primera, SCTHBP/A, que corresponde al fragmento que se encuentra entre los nucleótidos 25 a 57 que codifica para los residuos de amino ácidos que forman a la α -hélice A; la segunda, SCTHBP/C, que corresponde a los nucleótidos 127 a 159 que codifican para los residuos de la α -hélice C; y la tercera, SCTHBP/H, que corresponde a los nucleótidos 427 a 459 y que codifican para los residuos de la α -hélice H. La cantidad de nucleótidos degenerados para cada una de estas regiones fue la siguiente: 39% para SCTHBP/A, 36.3% para SCTHBP/C, y 27.2% para SCTHBP/H. Como resultado de estos cálculos se determinó que la región más conservada es la SCTHBP/H, la cual forma parte del empaque al grupo hemo en la estructura terciaria de la proteína (11). Por lo tanto, con base en los criterios de la cantidad de nucleótidos degenerados y de la localización en la estructura de la proteína, se supuso que la región SCTHBP/H es la que se encuentra sujeta a la

presión de selección mayor, y por lo tanto debe estar muy conservada en otros grupos de plantas diferentes a las Leguminosas. Por esta razón se seleccionó a esta región como la sonda que se utilizaría en los experimentos posteriores de hibridización, a la cual se le designó SCHBP/H.

Sin embargo, debido a que esta región juega un papel muy importante en el empaque del grupo hemo por la Hb, y por lo tanto en su función, fué necesario descartar la posibilidad de que esta región también se encuentre conservada en otras hemoproteínas. El resultado de la búsqueda de similitud entre la SCHBP/H con las secuencias proteicas que contiene el banco de datos NEWAT'85, demostró que SCHBP/H es exclusiva de las Hbs de las plantas, aunque existe cierta similitud con las Hbs de las bacterias, y en menor grado con las de animales (ver la Tabla II del Apéndice 2). Esto resultó ser significativo porque al parecer la sonda que se deriva de esta SCHBP/H puede ser específica en la búsqueda de los genes *hb* en las plantas; además, la similitud que se encontró entre la sonda y la Hb de *Vitreoscilla* (Tabla II del Apéndice 2) puede reflejar el estado ancestral del cual divergieron las Hbs vegetales. Es interesante el notar que este posible estado ancestral también se refleja fuertemente en el nivel de la composición total de amino ácidos ya que, como se indicó con anterioridad, la Alanina es el residuo más abundante en la SCTHBP, en donde alcanza el 13% del total; este residuo también es el más abundante en las proteínas de *E. coli*, así como en las proteínas de los procariontes en general, en donde alcanza un 9.5% del total de los residuos (70). Por lo tanto, resultaría de interés el probar esta sonda para la búsqueda de genes *hb* en otras bacterias, ya que la región de mayor similitud corresponde al fragmento SCHBP/H

cuando se alinean la SCTHBP con la Hb de *Vitreoscilla*.

OBTENCION DEL DNA TOTAL DE PLANTAS. Se intentaron diversos procedimientos de obtención del DNA total a partir de plantas superiores; de entre ellos, el que utiliza al detergente CTAB (63) fue el que dió los mejores resultados ya que el DNA prácticamente no se encuentra degradado (Figura 1A) en comparación, por ejemplo, con el método de las minipreparaciones de Dellaporta *et al.* (72) (Figura 1B), en donde se observa que una buena parte del DNA está degradado. Por esta razón el procedimiento de obtención del DNA que se siguió a lo largo del trabajo fue el del CTAB de Doyle y Doyle (63).

El DNA que se obtuvo se caracterizó por medios espectroscópicos. En la figura 2 se muestra el perfil de absorción del DNA del Frijol, el cual presenta máximos alrededor de 260 y 220 nm que son característicos de los ácidos nucleicos; el perfil de absorción del DNA de las demás plantas fue parecido al del Frijol, aunque en la mayor parte de ellos se observó mucho ruido en el espectro conforme el barrido se acercaba a los 220 nm, lo que seguramente se debe a la presencia de polisacáridos contaminantes en la preparación. En la tabla I se presentan las características espectrales más importantes y el rendimiento de todos los DNAs que se estudiaron en este trabajo; la mayor parte de los DNAs mostraron un valor del $R_{260/280}$ cercano a los 1.8, lo que indica que se logró eliminar una buena cantidad de las proteínas mediante los procedimientos de extracción, y aún cuando la pureza del DNA no fue total, las proteínas que se mantuvieron en la preparación no influyeron en los procedimientos posteriores, como fue la digestión con las enzimas de restricción.

En términos generales, el rendimiento del DNA de cada especie fue

Figura 1. Electroforésis en un gel de agarosa al 1% del DNA total que se extrajo de algunas plantas superiores por el método (A) del CTAB (63), y (B) de las minipreparaciones (72). El barrido intenso que se observa en (B) corresponde a los productos de degradación del DNA.

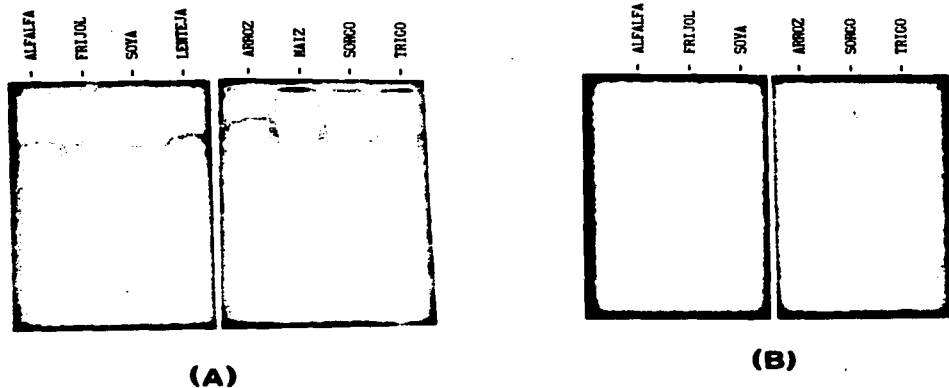


Figura 2. Espectro de absorción del DNA total del Frijol (70 ng/ml de TES) que se obtuvo por el método del CTAB (63).

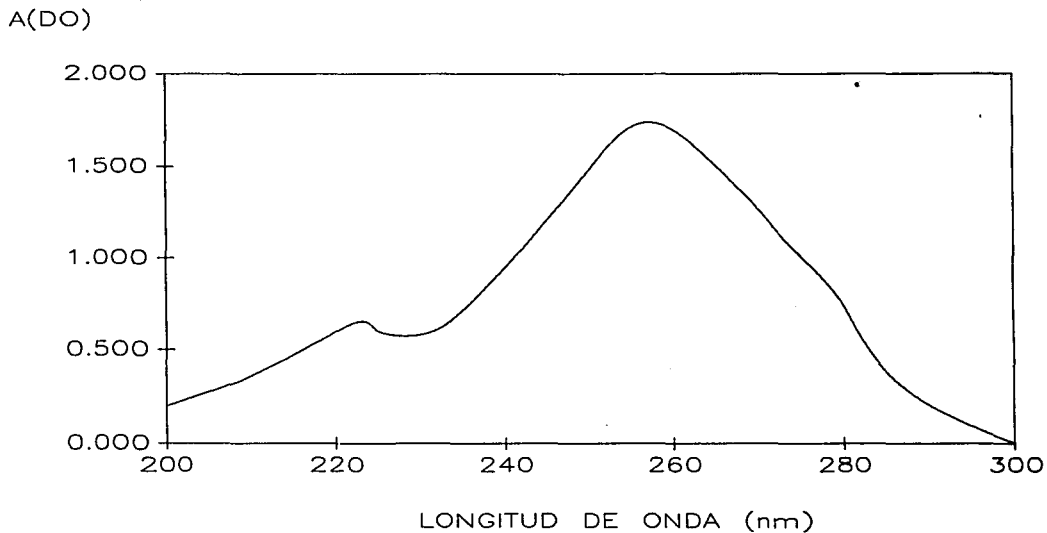


Tabla I. Características espectrales y rendimiento del DNA total que se obtuvo a partir de diferentes especies vegetales por el método del CTAB (63).

ESPECIE ¹	R260/280	RENDIMIENTO ²
LEGUMINOSAS:		
<i>Medicago sativa</i>	1.21	0.346
<i>Phaseolus vulgaris</i>	1.38	0.687
<i>Glycine max</i>	1.21	1.300
<i>Lens</i> sp.	1.09	7.600
<i>Erythrina</i> sp.	1.18	0.870
CASUARINALES:		
<i>Casuarina cunninghamiana</i>	1.23	0.086
GRAMINEAS:		
<i>Oryza</i> sp.	1.44	0.690
<i>Zea mays</i> ssp. <i>mays</i>	1.39	0.560
<i>Zea mays</i> ssp. <i>mexicana</i>	1.63	0.535
<i>Sorghum</i> sp.	1.66	1.300
<i>Triticum sativum</i>	1.32	2.000
OTRAS:		
<i>Magnolia</i> sp.	1.31	1.06
<i>Zamia furfuracea</i>	1.40	0.13

¹El tejido radical fue la fuente de extracción para todas las especies con la excepción de *C. cunninghamiana* y *Magnolia* sp. en donde el material fue el tejido foliar.

²El rendimiento está dado como mg de DNA/g de tejido fresco tomando $c(\text{DNA})_{260} = 13200$ (64).

bueno: las cantidades se encuentran en el intervalo de 0.346 a 7.6 mg de DNA/g de tejido fresco, para la Alfalfa y para la Lenteja, respectivamente. Además, no se observaron diferencias importantes en cuanto al tipo de tejido de donde se extrajo el DNA ya que, aunque el tejido embrionario es con el que se obtienen los mejores resultados, la extracción a partir de las hojas de la Magnolia dio buenos resultados, en donde el rendimiento fue alto y la calidad del DNA fué buena (Tabla I); en cambio, cuando se extrajo el DNA a partir de las hojas de la Casuarina, se presentaron diversos problemas en cuanto al rendimiento y a la presencia de algunos pigmentos contaminantes. El cálculo de la cantidad del DNA que se obtuvo para cada especie coincidió bien cuando se realizó mediante procedimientos espectroscópicos y cuando se comparó con el corrimiento electroforético patrón que contenía cantidades conocidas del DNA del timo de la ternera.

Cuando el DNA se extrajo a partir del tejido foliar de la Casuarina, el rendimiento fue mucho menor, 0.086 mg de DNA/g de tejido fresco, del que se obtuvo para las demás especies (Tabla I); esto concuerda con algunos reportes en donde se establece que el tejido foliar es una fuente rica para la obtención del RNA, pero que el rendimiento en la cantidad del DNA es muy bajo (73); sin embargo, este no es un caso general ya que, como se indicó con anterioridad, la extracción es muy buena cuando ésta es a partir de las hojas de la Magnolia. Otro problema que se presentó cuando se obtuvo el DNA de la Casuarina a partir del tejido foliar, fue la presencia de un pigmento pardo, el cual aparentemente no interfiere con la actividad de las enzimas de restricción, pero que puede dar lugar a un "fondo" en el filtro de NC debido a su color cuando se realizan hibridizaciones del tipo "Dot" o

"Slot-Blot", de tal manera que la presencia de este pigmento pudiera interferir con la detección de la señal que se deriva de la actividad de la fosfatasa alcalina, la que resulta en un color azul. Así, se intentaron diversos procedimientos para eliminar el pigmento de la preparación (Tabla II), lo que a su vez permitió obtener su caracterización parcial.

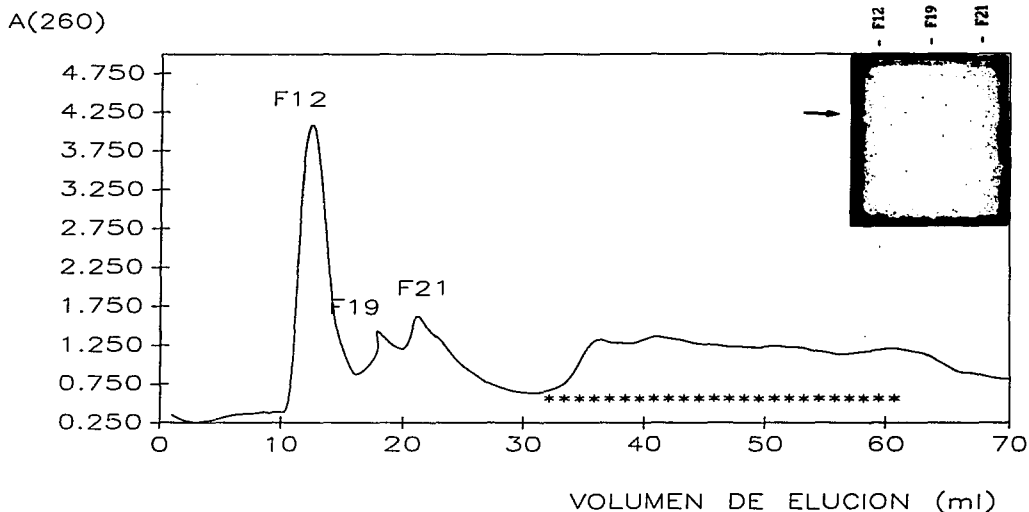
Se cree que este pigmento puede corresponder a compuestos de PM bajo, tales como los polifenoles o las quinonas, los cuales se derivan de la actividad de las fenol oxidasas (74). Estos compuestos deben estar presentes en el tejido foliar *in vivo*, ya que cuando se homogeniza el tejido en condiciones reductoras, en presencia del β -mercaptoetanol, no se inhibe su síntesis y se les encuentra siempre en la preparación; probablemente, están compartimentalizados dentro de la célula en estructuras como los peroxisomas, y simplemente se liberan durante la homogenización.

La mayor parte de estos compuestos se unen al DNA (¿y a otras biomoléculas?) con una afinidad muy alta, ya que no se les encontró en forma libre en el extracto acuoso; este pigmento tampoco pudo ser quelado por el EDTA o fijado por la BSA, además de no ser dializable (Tabla II). No se le pudo extraer con solventes orgánicos, tal como el cloroformo, y siempre se le encontró en la fase acuosa, ya sea por su carácter hidrosoluble o por su alta afinidad hacia el DNA. Finalmente se le pudo separar del DNA de PM alto mediante cromatografía en Sephadex G-75 (Pharmacia-LKB) (Tabla II). En la figura 3 se puede observar que el pigmento eluyó después de los componentes de PM alto, que se encuentran en las fracciones 12, 19 y 21 (ver F12 en el recuadro de la Figura 3), aunque este pigmento continuó asociado con el DNA

Tabla II. Características del pigmento pardo que se obtiene al extraer el DNA total a partir del tejido foliar de *C. cunninghamiana* por el método del CTAB (63).

PARAMETRO	CARACTERISTICA
SOLUBILIDAD	Hidrosoluble, en TES
DIALISIS	No es dializable
QUELANTES:	
a. EDTA	No es quelado
b. BSA	No es fijado
REDUCTORES:	
a. β -MERCAPTOETANOL	No inhibe su síntesis
FILTRACION:	
a. SEPHADEX G-75	Eluye después del DNA de tamaño molecular alto

Figura 3. Perfil de elución en Sephadex G-75 del DNA que se obtuvo a partir de las hojas de la Casuarina. Columna: 25 x 1.5 cm; volúmen vacío: 15 ml; eluyente: TEB. *, fracciones en donde eluyó el pigmento pardo; recuadro, electroforésis en un gel de agarosa de las muestras que eluyeron en las fracciones F12, F19 y F21; →, DNA de tamaño molecular alto que está presente en F12.



degradado, ya que las fracciones en donde eluyó presentan una absorción alta a 260 nm; el pigmento no absorbió en la región del ultravioleta, solamente en el visible en donde tiene máximos en 592 y 657 nm.

Por lo tanto, y con base en las evidencias anteriores, se sugiere que estos compuestos se liberan de las células del tejido foliar en el momento de la homogenización y se pegan al DNA con una afinidad muy alta en donde inducen rompimientos inespecíficos, probablemente en la presencia de metales (Cu^{++} , Fe^{+++}) y/o de radicales libres (HO^- , O_2^-); sin embargo, mediante la filtración en el gel fue posible recuperar el DNA de PM alto que no ha reaccionado con este compuesto, y además fue posible separarlo de los productos de la degradación en los que se ha fijado el pigmento. Probablemente la presencia y el efecto que tiene este compuesto en las preparaciones del DNA sean la causa del rendimiento tan bajo que se obtiene cuando se extrae el DNA a partir del tejido foliar. Será interesante probar en el futuro algunos inhibidores específicos de las enzimas que participan en la biosíntesis de estos metabolitos con el objeto de aumentar el rendimiento en la obtención del DNA a partir de las hojas, que es el tejido más accesible en las plantas.

Se utilizó el DNA total de diversas plantas para los experimentos de hibridación. Sin embargo, fue necesario descartar el DNA de la Casuarina debido a que la cantidad que se obtuvo fue muy baja, aún cuando se le purificara mediante filtración en gel; esto mismo ocurrió cuando se intentó extraer el DNA a partir del tejido foliar de otras plantas, tal como las frondas de los helechos.

CLONACION Y MARCAJE DE LA Sonda SCHBP/H. Se obtuvieron 11 colonias transformantes (Amp^r , x-Gal^-) como resultado de la clonación, para las

cuales se obtuvo el plásmido pHBCS/H, el que se digirió con *Pst*I y *Sph*I, y se corrió electroforéticamente en un gel de poliacrilamida. En la figura 4 se muestra que la sonda SCHBP/H se clonó bien, ya que se recupera como un fragmento discreto de unos 20 KDa de PM después de digerir el plásmido.

Se marcó la sonda SCHBP/H por "Nick Translation" con derivados biotinilados, dATP-biotina, mediante el uso del kit "BioNick Labeling System" según las instrucciones del fabricante (GIBCO BRL). Sin embargo, el tamaño mínimo de la sonda que se recomienda para el uso de este kit es de 50 pb; debido al tamaño de la sonda SCHBP/H, 33 pb, fue necesario el comprobar que se había marcado eficientemente por este sistema. En la figura 5 se aprecia que 300 ng de la sonda marcada dan una señal intensa cuando se fijan al filtro de NC; esta sonda se separó previamente de los nucleótidos que no se incorporaron mediante precipitación con acetato de sodio-etanol. Esto demostró que es posible marcar oligonucleótidos con un tamaño menor a los 50 pb mediante el procedimiento de "Nick Translation"; probablemente el límite inferior en la longitud del oligonucleótido a marcar sea de unos 20 a 25 pb ya que el sitio activo del fragmento Klenow de la DNA polimerasa I acepta unos 19 pb (ver el esquema que se presenta en la página 31 del catálogo "Molecular Biologicals" [1986] de Pharmacia Fine Chemicals). Por lo tanto, este fue el sistema de marcaje de la sonda que se adoptó durante el desarrollo del trabajo.

HIBRIDIZACION POR "DOT" Y "SLOT-BLOT". Los primeros experimentos de hibridación que se realizaron utilizando la sonda SCHBP/H fueron mediante ensayos del tipo "Dot-Blot", en donde se intentó probar la funcionalidad de la sonda al hibridizar con los DNAs de aquellas

Figura 4. Electroforésis nativa en un gel de poliacrilamida al 7% (67) del plásmido pHBGS/H que se linearizó previamente con *Pst*I y *Sph*I. Las colonias corresponden a algunas de las transformantes que contuvieron al plásmido pHBGS/H (ver el texto).

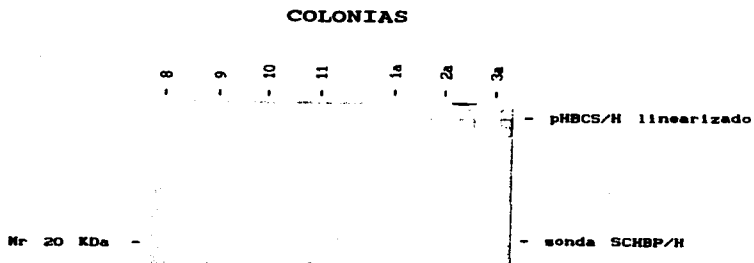


Figura 5. "Dot-Blot" en un filtro de NC de 300 ng de la sonda SCHBP/H que se marcó por "Nick Translation" al utilizar el kit "BioNick Labeling System" (GIBCO BRL). El revelado fue por el sistema NBT/BCIP para la fosfatasa alcalina.



plantas que se sabe que contienen genes *hb*. En la figura 6 se aprecia que la sonda SCHBP/H hibridizó bien con los DNAs sin digerir de las plantas que forman al grupo de los controles positivos, las Leguminosas, en donde la mayor intensidad de la hibridización se obtuvo con la Lenteja, 1.00, y la menor con la Alfalfa, 0.75. Para descartar la existencia de falsos positivos se incluyó en las hibridizaciones a los DNAs de las fuentes animales: de testículos del salmón y de timo de ternera. La hibridización de la sonda SCHBP/H con estos DNAs fue significativamente menor, 0.15 y 0.135 respectivamente, que la que se obtuvo con los DNAs de las Leguminosas. Aún cuando existen los genes *hb* en estos DNAs de animales, la similitud con la sonda debe ser muy baja debido a los fenómenos de divergencia, lo que resulta en señales de hibridización débiles; en la tabla II del Apéndice 2 se puede apreciar que el nivel de similitud es muy bajo cuando se compara la sonda SCHBP/H con las Hbs de diversas fuentes animales. De esta manera, y con base en las evidencias anteriores, es posible descartar la existencia de falsos positivos que resultan de hibridizaciones inespecíficas.

Como se mencionó con anterioridad, la sonda SCHBP/H hibridiza con todos los DNAs de las Leguminosas, pero la intensidad de la señal fue diferente para cada especie (Figura 6). Esto puede explicarse con base en la similitud que existe entre la secuencia de la sonda y el DNA blanco, en donde a mayor similitud se fijará una cantidad mayor de la sonda, y la señal de hibridización será más intensa, y a la inversa. Aún cuando los genes *hb* están reiterados en las Leguminosas (74), y si se considera que esto pudiera determinar la intensidad de la señal en la hibridización, debido a que se fijará una cantidad mayor de la sonda

Figura 6. Hibridización por "Dot-Blot" de 20 μg del DNA total sin digerir de plantas Leguminosas y Gramíneas con la sonda SCHBP/H. Los controles corresponden a: 1) 20 μg del DNA de testículos del salmón (Sigma); y 2) 20 μg del DNA de timo de ternera (Merck). El revelado fue para la fosfatasa alcalina. En los paréntesis se indican los valores de intensidad relativa de la hibridización, los que se obtuvieron de la siguiente manera: se fotografiaron los filtros con los "Dot-Blots" y se realizó un barrido densitométrico al negativo para determinar la especie vegetal que presenta el pico mayor de transmitancia (la Lenteja), en el que se obtuvo la altura máxima, y que sirvió para obtener la proporción relativa de acuerdo con la altura máxima del pico que presenta cada especie.

LEGUMINOSAS

Control 1
(0.13)

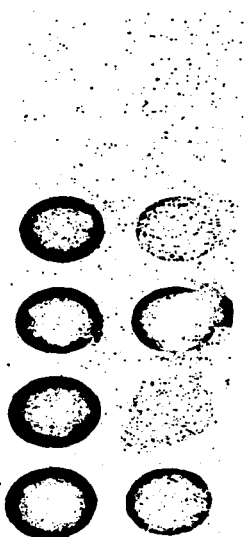
Control 2
(0.13)

Alfalfa
(0.75)

Frijol
(0.90)

Lenteja
(1.00)

Soya
(0.88)



Control 1
(0.17)

Control 2
(0.14)

Arroz
(0.45)

Maiz
(0.84)

Sorgo
(0.45)

Trigo
(0.70)

GRAMINEAS

en aquella planta que contenga el número mayor de copias de los genes *hb*, las diferencias en el número de copias que se han encontrado entre las diversas especies no son lo suficientemente grandes como para explicar las distintas intensidades de hibridización que se observaron cuando se hibridizó la sonda SCHBP/H con los DNAs de las Leguminosas. Si esto es así, entonces llama la atención el hecho de que la señal de hibridización más intensa resultó con el DNA que se obtuvo de la Lenteja, 1.00 (Figura 6), lo que sugiere que la secuencia de la sonda SCHBP/H tiene una similitud mayor con esta planta que con las demás Leguminosas. Esto es relevante desde el punto de vista de que, aunque ya se han purificado las Lbs de la Lenteja, todavía no se han clonado y secuenciado los genes que las codifican, y con base en los criterios anteriores se espera que la región que hibridiza sea altamente similar a la secuencia de la sonda SCHBP/H.









Además de hibridizar con los DNAs de las Leguminosas, la sonda hibridizó con los DNAs de las Gramíneas (Figura 6). Esto resultó de mucho interés ya que éstos fueron los primeros resultados que indicaron la posible existencia de genes *hb*, o secuencias muy relacionadas con ellos, en las plantas Gramíneas. Al igual que con las Leguminosas, la sonda SCHBP/H hibridizó con intensidades diferentes entre las Gramíneas, en donde la señal más intensa se detectó con el DNA del Maíz, 0.84, y la de menor intensidad con los DNAs del Arróz y del Sorgo, 0.45, pero que sin embargo fueron significativamente mayores a las señales que se obtuvieron con los controles negativos. Si se considera nuevamente la posibilidad de que la intensidad de la señal de hibridización se relacione directamente con el nivel de similitud que existe entre el DNA blanco y la sonda, independientemente del número de

copias de los genes *hb*, entonces es factible el proponer que estos posibles genes *hb* tienen un nivel de similitud semejante en la Soya y en el Maíz, ya que ambas muestran señales de hibridación parecidas, 0.88 y 0.84 respectivamente, y en la Alfalfa y el Trigo, con señales de 0.75 y 0.70 respectivamente (Figura 6).

En cuanto a la especificidad de la sonda SCHBP/H en la hibridación con los DNAs de las Leguminosas y de las Gramíneas, la ausencia de señales en la hibridación con los DNAs de las fuentes animales fue un criterio para descartar la existencia de falsos positivos. Sin embargo, la astringencia con la que se lavaron los filtros fue otro parámetro que permitió asegurar con mayor confianza el que las regiones que hibridizan con la sonda corresponden a genes *hb*; la astringencia alcanzó hasta 0.16x de SSC en una solución que contuvo SDS al 0.1%, y los lavados se realizaron a 65°C (ver Materiales y Métodos). En esta astringencia se espera que las sondas que hibridizan de manera inespecífica sean lavadas.

Los resultados que se obtuvieron mediante experimentos de hibridación del tipo "Dot-Blot" (Figura 6) fueron reproducibles bajo las mismas condiciones. En la figura 7 se muestra la hibridación de la sonda SCHBP/H con los DNAs de varias plantas superiores, así como con los controles negativos. En este ensayo se incluyeron los DNAs de algunas plantas que son de interés desde el punto de vista evolutivo, como es el Teosinte, el cual se ha sugerido como el ancestro del Maíz cultivado (76), la Magnolia, cuya ubicación filogenética se encuentra en la dicotomía que dió lugar a los linajes de las mono y dicotiledóneas (77), y una Cicada, que representa a uno de los estados ancestrales con mayor antigüedad entre las plantas vasculares (78).

Figura 7. Hibridación del tipo "Slot-Blot" de 20 µg del DNA total sin digerir de plantas superiores con la sonda SCHBP/H. El revelado fue para la fosfatasa alcalina.

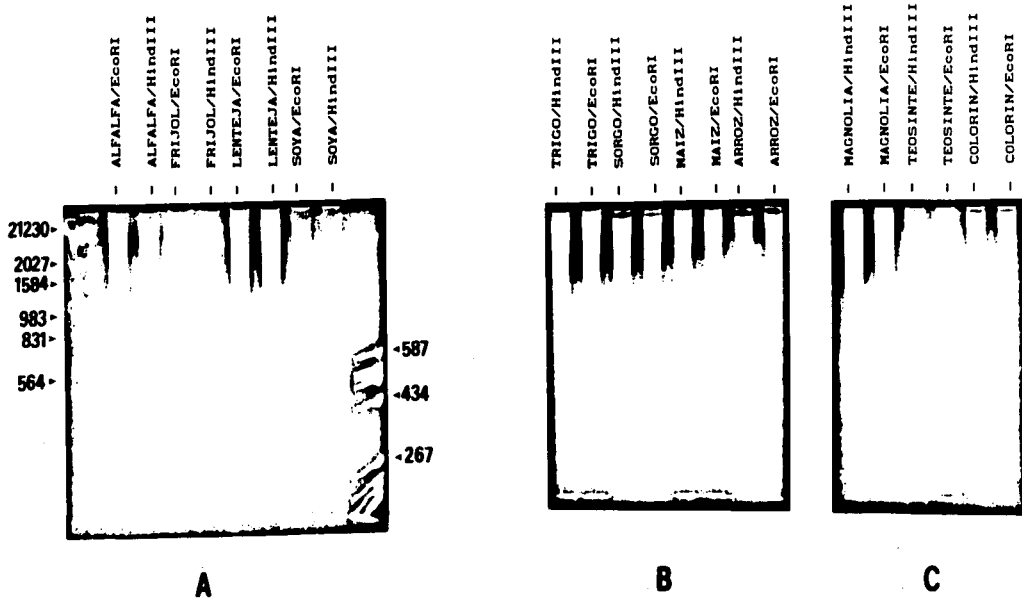
CONTROLES NEGATIVOS	CONTROLES POSITIVOS LEGUMINOSAS	GRAMINEAS	OTRAS
TIPO DE TERNERA		ALFALFA	COLORIN
TESTICULO DE		FRIJOL	TEOSINTE
SALMON		LENTEJA	MAGNOLIA
		SOYA	CICADA
			ARROZ
			MAIZ
			SORGO
			TRIGO

Los resultados de la hibridización fueron de mucho interés ya que la sonda SCHBP/H no dió señal alguna con los DNAs de fuentes animales, pero hibridizó bien con los DNAs de las Leguminosas, y con menor intensidad con los DNAs de las Gramíneas y de las demás plantas no Dicotiledóneas (Figura 7). En términos generales, se puede apreciar cómo disminuye la intensidad de la señal de la hibridización en relación con el aumento en la distancia filogenética, a partir de las Leguminosas y hasta llegar a la Cícada; la excepción a este patrón corresponde al Sorgo, en donde la intensidad de la señal es muy débil.

Los resultados que se muestran en la figura 7 son significativos desde el punto de vista de que, además de estar presentes en las Gramíneas, los genes *hb* también pudieran encontrarse en plantas que representan linajes ancestrales, como lo son la Magnolia y la Cícada, lo que sugiere que estos genes descienden, por lo menos, a partir del ancestro de las plantas vasculares, lo que apoya la hipótesis de la evolución vertical de los genes *hb* en el reino Plantae (48, 54, 57). Además, esta sonda SCHBP/H puede ser útil para aislar genes *hb* en diversas plantas, inclusive en aquellas que son filogenéticamente distantes de las Leguminosas; estos genes se pueden utilizar posteriormente como sondas moleculares.

HIBRIDIZACION DIRECTA EN GELES DE AGAROSA. Además de las hibridizaciones del tipo "Dot" y "Slot-Blot", se intentaron hibridizaciones del tipo "Southern-Blot" (79) (en filtros de NC o de nylon [N⁺, Amersham]) con los DNAs que se digirieron previamente con *EcoRI* y *HindIII*, y que se separaron en geles de poliacrilamida o de agarosa. Aunque los DNAs se digirieron bien (Figura 8), y se aplicaron cantidades altas de los digeridos a los geles, hasta 25 µg, nunca fue

Figura 8. Perfiles electroforéticos en un gel de poliacrilamida al 5% de la digestión con *EcoRI* y *HindIII* de 10 μ g del DNA total de plantas superiores. Los marcadores del tamaño molecular corresponden al digerido de λ con *EcoRI/HindIII*, y al digerido de pBR322 con *HaeIII*, y se expresan en pb. A, Leguminosas; B, Gramíneas; C, varias.

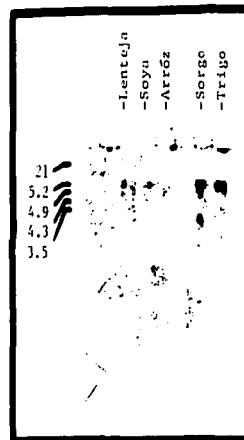
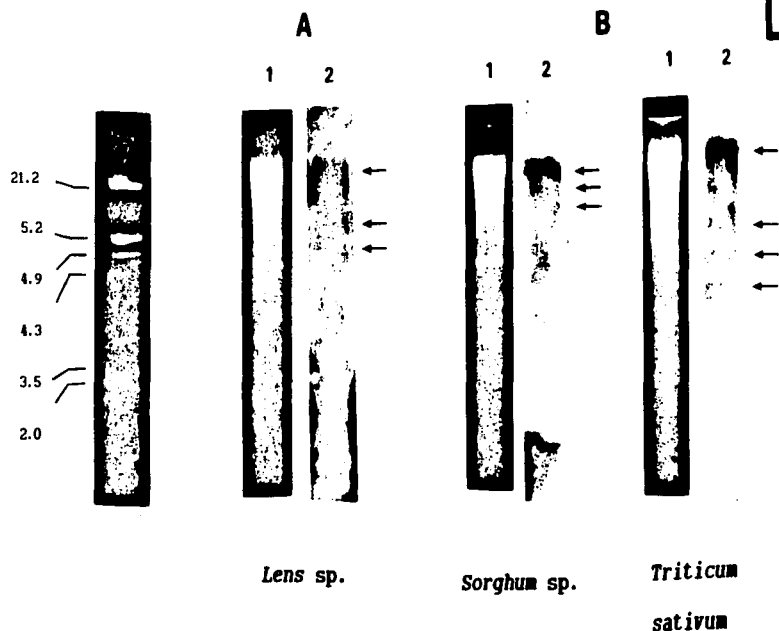


posible obtener resultados positivos debido a que la sonda SCHBP/H es de un tamaño molecular muy pequeño, de 33 pb, y a que está degenerada porque proviene de una secuencia consenso. Estas fueron las causas por las que no se encontraron las condiciones óptimas para la hibridización y/o para el lavado: no se detectaron señales cuando los filtros se lavaron con una astringencia alta, en cambio se desarrolló demasiado "fondo" en el filtro cuando la astringencia del lavado fue baja, con la consecuencia de que no se pudieron detectar las señales.

Por lo tanto, se optó por seguir las recomendaciones de Meinkoth y Wahl (80) en el sentido de realizar las hibridizaciones directamente en el gel cuando se trata del uso de una sonda de tamaño molecular pequeño. En el recuadro de la Figura 9 se observa que fué posible obtener señales de hibridización para algunas Leguminosas, la Lenteja y la Soya, y para algunas Gramíneas, el Sorgo y el Trigo, mediante la aplicación del procedimiento de Gontijo *et al.* (68) (ver Materiales y Métodos); el gel presentó mucho "fondo" debido a que los lavados se realizaron con una astringencia baja de 6x de SSC, sin embargo las señales que se obtuvieron corresponden a bandas que muestran tamaños moleculares que se encuentran entre los 4 y 20 Kpb, adicionalmente a la señal del DNA genómico, cuyo tamaño molecular es mayor a los 21.2 Kpb.

No fue posible descartar la presencia de hibridizaciones inespecíficas en el caso de la hibridización directa en el gel debido a que los lavados fueron con una astringencia muy baja, ya que las señales desaparecieron al aumentar dicha astringencia. Por lo tanto, se debe considerar con cautela a estos resultados. Sin embargo, en el caso de que estas señales de hibridización correspondan a genes *hb* auténticos, entonces éstos pudieran ser polimórficos en la Lenteja, el

Figura 9. Hibridación directa en un gel de agarosa al 1.5% (68) de 5 µg del DNA de plantas superiores (digerido con *Hind*III) con la sonda SCHBP/H. A, Dicotiledóneas; B, Monocotiledóneas. 1, gel teñido con bromuro de etidio; 2, hibridación y revelado para fosfatasa alcalina. →, bandas que hibridizan con la sonda. Recuadro: panorama general del gel después de la hibridación. Los marcadores del tamaño molecular corresponden al digerido de λ con *Eco*RI/*Hind*III, y se expresan en Kpb.



Sorgo y el Trigo, ya que se detectan de 2 a 3 bandas, con tamaños moleculares diferentes, además de la banda que corresponde al genómico. Para confirmar la identidad de estos fragmentos que hibridizan con la sonda, como presuntos genes *hb*, se deberá realizar la clonación y la secuenciación de estos fragmentos; sin embargo, los procedimientos experimentales que condujeran hacia la obtención de la secuencia consumirían demasiado tiempo debido a que estos fragmentos son de un tamaño molecular muy grande, entre 4 y 20 Kpb. Por esta razón se optó por intentar la amplificación específica de la región del DNA que hibridiza con la sonda SCHBP/H mediante procedimientos del tipo PCR (81).

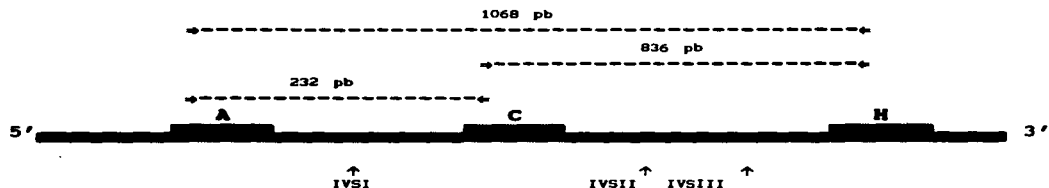
AMPLIFICACION GENETICA POR PROCEDIMIENTOS DEL TIPO PCR. Con base en la secuencia consenso total para las Hbs de plantas (ver la Figura 3 del Apéndice 2) se diseñaron 4 sondas (Tabla III) con el objeto de amplificar específicamente las secuencias que se encuentran entre las regiones más conservadas (Figura 10): el fragmento C-H que tiene el tamaño de unas 836 pb, el cual se amplifica al utilizar el par de sondas 511.3 y 510.3, y el fragmento A-H de unos 1068 pb que se amplifica al utilizar las sondas 399 y 398. Se ha sugerido con anterioridad el "degenerar" a las sondas mediante la adición de secuencias de restricción para facilitar la clonación posterior de los fragmentos que se amplifiquen (82). En nuestro caso, las sondas son degeneradas ya que se les añadieron secuencias para *Pst*I y para *Sph*I (Tabla III) con el objeto de poder clonar los fragmentos resultantes en el vector pUC19. Por lo tanto, se les denominó a estas sondas como "superdegeneradas" debido a que contienen los sitios de restricción y a que provienen de una secuencia consenso; además, por esta última razón,

Tabla III. Características principales de las sondas que se utilizaron para amplificar fragmentos de los genes *hb* de las plantas por PCR.

SONDA	PORCION CODIFICANTE	LONGITUD (PB)	C+G (%)	A+T (%)	T _m ¹ (°C)	SITIO DE RESTRICC.
510.3	α-hélice C	26	50	50	64.8	<i>Pst</i> I
511.3	α-hélice H	27	52	48	66.6	<i>Sph</i> I
399	α-hélice A	29	52	48	68.2	<i>Pst</i> I
398	α-hélice H	27	52	48	66.5	<i>Sph</i> I

¹El T_m (°C) se calculó de acuerdo con la fórmula (65): T_m=69.3+0.41(G+C)%-650/L, en donde L=longitud del oligonucleótido.

Figura 10. Esquema general que muestra la longitud total de los genes *hb* de las plantas Dicotiledóneas. → y ←, posición de las sondas; [■], regiones más conservadas; ↑, posición aproximada de los intrones (IVS). En la parte inferior se indica el tamaño de los intrones y de las regiones codificantes que se han calculado en diversas plantas Dicotiledóneas.



ESPECIE	LONGITUD DEL INTRON (pb)	REGION CODIFICANTE ¹	LONGITUD (pb)	REF.
<i>Trema</i>	IVSI= 127	A a C	105	32
	IVSII= 348	A a H	434	32
	IVSIII= 159	C a H	329	32
<i>Parasponia</i>	IVSI= 122			32
	IVSII= 334			
	IVSIII= 159			
<i>Sesbania</i>	IVSI= 105			32
	IVSII= 307			
	IVSIII= 240			

¹La longitud de estas regiones codificantes corresponde a un valor promedio entre los genes *hb* de diferentes especies.

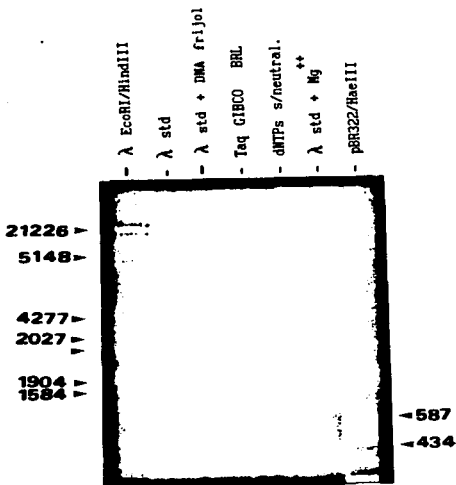
las sondas no son complementarias en su totalidad con el DNA de las plantas que se estudiaron en este trabajo. Esto tuvo implicaciones importantes en cuanto a la selección de la astringencia de la temperatura en la que se llevó a cabo el alineamiento, la cual no debería ser alta.

Se estandarizaron las condiciones de la amplificación al evaluar diferentes parámetros. En la figura 11 se muestra el efecto que tienen algunos de estos factores al amplificar un fragmento de 500 pb de longitud del DNA del fago λ ; los componentes de este sistema de amplificación corresponden al kit "GeneAmp" (Cetus, Perkin Elmer). Se obtuvo un fragmento de unos 500 pb cuando se amplificó bajo las condiciones estándar que recomienda el fabricante, y utilizando la enzima *Ampli Taq* del mismo kit. Esta amplificación no se modificó drásticamente cuando se añadió 1 μg del DNA del Frijol a la preparación, de tal manera que la presencia de un DNA de tamaño molecular grande no influye en la amplificación.

Sin embargo, en la figura 11 se muestra que fueron dos los factores que influyeron de manera importante en la amplificación: la fuente de la enzima *Taq* DNA polimerasa y el pH de la solución de los dNTPs. Prácticamente no se observó amplificación alguna cuando se utilizó la enzima *Taq* DNA polimerasa de GIBCO BRL, en comparación con la *Ampli Taq* de Cetus; por lo tanto, parecen existir diferencias importantes entre los lotes y la fuente de las enzimas; sin embargo en algunos trabajos se reporta la amplificación eficiente cuando se utiliza la enzima de GIBCO BRL, o de otras fuentes (ver 83).

En fechas recientes se ha utilizado con más frecuencia el fragmento Stoffel de la enzima *Ampli Taq* (Cetus), el cual consiste de una *Taq* DNA

Figura 11. Efecto de diversos parámetros en la amplificación por PCR de un fragmento de 500 pb del fago λ . La separación electroforética fue en un gel de agarosa al 1%. Los componentes de la amplificación estándar (λ std) fueron de acuerdo con las instrucciones del fabricante del kit "GeneAmp" (Cetus). Los ciclos de amplificación fueron los siguientes: 1 ciclo: 95°C/2 min; 37 ciclos: 46°C/1.5 min, 72°C/1.5 min, 95°C/1.0 min; 1 ciclo: 72°C/ 5 min.



polimerasa que se ha modificado al eliminar unos 30 residuos de amino ácidos del extremo carboxilo terminal; esto le confiere una termo-estabilidad mayor a la enzima, además de un intervalo más amplio en cuanto a su actividad en diferentes concentraciones de Mg^{++} (84). Esta enzima fue la que se seleccionó en este trabajo para los experimentos de amplificación.

El segundo factor que influyó de manera importante en la amplificación del fragmento de 500 pb del fago λ fue el pH de la solución de los dNTPs. En la figura 11 se observa que no existe amplificación alguna cuando la solución de dNTPs no se neutraliza. Esto se debe seguramente a que cuando la solución de amplificación contiene a los dNTPs sin neutralizar, el pH disminuye de 8.0 a 6.5, lo que debe afectar la actividad de la enzima cuyo máximo se encuentra alrededor del pH 8.0; los valores de pH de una solución 20 mM de los dNTPs se encuentran entre 3.5 y 5.0, según el desoxinucleótido de que se trate. Además, se ha reportado que la fidelidad de la amplificación disminuye notoriamente cuando no se neutraliza la solución de los dNTPs (85).

Por lo tanto, con base en experimentos como el anterior se establecieron las condiciones de amplificación por PCR de los genes *hb* en las plantas de estudio; sin embargo, las condiciones generales variaron de especie a especie debido, principalmente, a la complementariedad que existe entre las sondas y el DNA blanco.

Así, se logró amplificar un fragmento con alrededor de 1.2 Kpb de longitud para el DNA del Frijol al utilizar las sondas 511.3 y 510.3, y una temperatura para el alineamiento de $46^{\circ}C$ (Figura 12). El fragmento que resultó de la amplificación tuvo un tamaño mayor al que se esperaba, de unos 0.8 Kpb. Sin embargo, aunque estas condiciones

Figura 12. Amplificación por PCR de los genes *hb* de (A) la Alfalfa y (B) del Frijol al utilizar las sondas S10.3 y S11.3. Los marcadores del tamaño molecular corresponden al digerido de λ con *EcoRI/HindIII*, y se expresan en pb. →, fragmento del Frijol que se amplificó con una longitud de 1.0 a 1.2 Kpb.

A B

1904-
1584-
1330-
980-

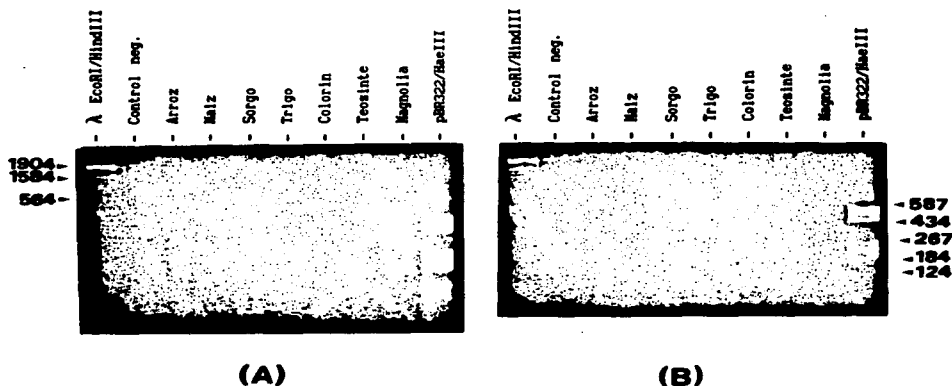


funcionaron para el Frijol, esto no ocurrió para la Alfalfa en donde se obtuvo un barrido.

Los resultados anteriores fueron poco reproducibles cuando se utilizó el par de sondas 511.3 y 510.3 ya que con mucha frecuencia se obtuvo un barrido con un tamaño molecular grande. Bell y De Marini (86) reportan que cuando el número de ciclos para la amplificación es excesivo, mayor a 32, se puede presentar un barrido con la disminución consecuente del producto específico de la amplificación. Por lo tanto, al reproducir las condiciones de amplificación de estos autores, pero para el sistema de los genes *hb*, se encontró que el rendimiento en el producto aumenta substancialmente cuando se incrementa el número de ciclos de la amplificación a un número mayor a 40, sobretodo al disminuir la astringencia de la temperatura del alineamiento hasta los 37°C .

De esta manera, fue posible obtener fragmentos discretos de la amplificación para las Gramíneas, en particular para el Maíz, el Sorgo y el Trigo (Figura 13A), al utilizar las condiciones anteriores y el par de sondas 399 y 398, que en teoría pueden amplificar un fragmento de 1 Kpb, que incluye tres intrones, al considerar la estructura de los genes *hb* en las Dicotiledóneas (Figura 10). Los tamaños moleculares de los fragmentos que se amplificaron se encuentran entre los 100 y 400 pb, en donde la banda más evidente fue la de 400 pb del Sorgo. En cambio, no se obtuvo producto alguno cuando se utilizó el par de sondas 511.3 y 510.3, bajo las mismas condiciones que para la amplificación anterior (Figura 13B). Por lo tanto, el par de sondas 399 y 398, que amplifican a la región A-H, son adecuadas para amplificar genes *hb* en las Gramíneas.

Figura 13. Amplificación por PCR de los genes *hb* de plantas no-Dicotiledóneas al utilizar (A) el par de sondas 399 y 398, y (B) el par de sondas 510.3 y 511.3. El control negativo contuvo todos los componentes de la amplificación pero sin el DNA genómico. La separación electroforética fue en un gel de agarosa PE (Cetus) al 3% en TBES. Los marcadores del tamaño molecular corresponden a los digeridos de λ con *EcoRI/HindIII*, y de pBR322 con *HaeIII*, y se expresan en Kpb.



Es posible determinar la especificidad con la que se amplifican los fragmentos al aumentar la astringencia de la temperatura del alineamiento, esperando que desaparezcan los productos inespecíficos. En este sentido, se observó un efecto contrario al que se esperaba inicialmente cuando se aumentó la astringencia a 42°C durante la amplificación de los DNAs de las Leguminosas y las Gramíneas, al utilizar el par de sondas 399 y 398: no desaparecieron los posibles fragmentos inespecíficos sino que se obtuvo un barrido, particularmente en las Gramíneas (Figura 14). Por lo tanto, el aumento en la temperatura del alineamiento favorece la formación de concatenámeros a partir de los productos originales de la amplificación; estas observaciones no se han publicado con anterioridad. En la figura 14 también se puede observar que el par de sondas 399 y 398 tienen una complementariedad baja con los DNAs de las Leguminosas, ya que no se detectan productos de la amplificación.

Finalmente, en la figura 15 se muestra una sinopsis de los fragmentos que se lograron amplificar mediante procedimientos del tipo PCR para algunas Gramíneas. Los factores que determinaron la eficiencia de la amplificación fueron: a) el uso del par de sondas 399 y 398; b) el uso del fragmento Stoffel de AmpliTaq (Cetus); c) la baja astringencia en la temperatura del alineamiento, a 37°C; y d) el alto número de ciclos de amplificación, 45. El tamaño aproximado de los fragmentos que se amplificaron fue el siguiente (Figura 15): Maíz, 100 y 130 pb; Sorgo, 200 y 430 pb; y Trigo, 300 y 430 pb. Aunque en teoría se esperaban fragmentos con un tamaño molecular mayor, entre 0.8 y 1.2 Kpb (Figura 10), resulta de mucho interés el hacer notar que el tamaño de los fragmentos que se amplificaron, entre los 300 y 430 pb,

Figura 14. Efecto del aumento de la astringencia de la temperatura del alineamiento, a 42°C, en la amplificación por PCR de los DNAs de plantas superiores al utilizar el par de sondas 399 y 398. La separación electroforética fue en un gel de agarosa PE (Cetus) al 3% en TBES.

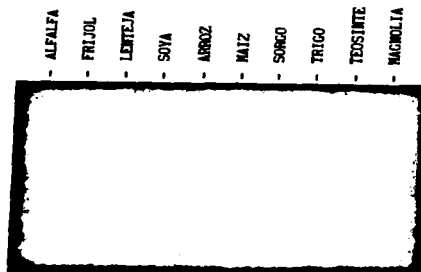
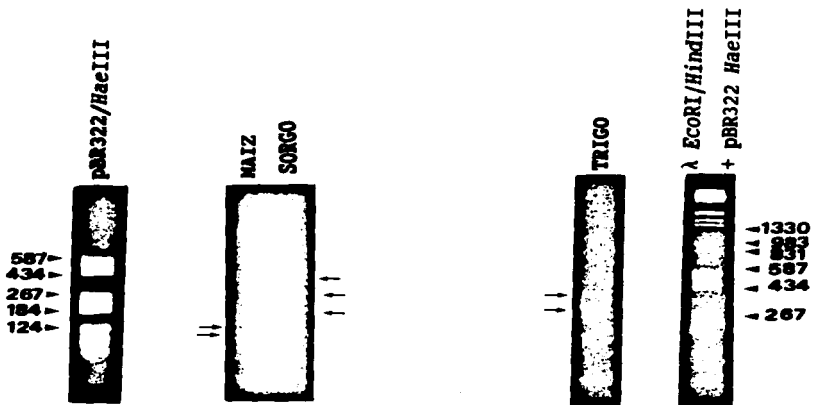


Figura 15. Fragmentos que resultan de la amplificación por PCR de los genes *hb* de las plantas Gramíneas al utilizar el par de sondas 399 y 398. La separación electroforética fue en un gel de agarosa PE (Cetus) al 3% en TBE8. Las flechas indican las bandas más evidentes. Los marcadores del tamaño molecular corresponden a los digeridos de λ con *EcoRI/HindIII*, y pBR322 con *HaeIII*, y se expresan en pb.



corresponden a la longitud aproximada de la región codificante de los genes *hb* en las Dicotiledóneas, lo cual conduce a pensar inevitablemente, que estos genes pudieran carecer de intrones en las Gramíneas. Si esto se puede confirmar posteriormente, la ausencia de intrones en estas plantas resultaría muy significativo desde el punto de vista evolutivo, ya que se ha sugerido que el gen *hb* ancestral del linaje de las plantas contenía la combinación 4 exones/3 intrones (54), lo cual se pondría en duda en el caso de demostrar la ausencia de intrones en las Monocotiledóneas, y que además llevaría a pensar que los genes *hb* en estas plantas pudieron ser el resultado de fenómenos de transferencia genética horizontal (38); el candidato más importante como fuente de estos genes lo serían los insectos ya que se ha encontrado que en estos organismos los genes *hb* carecen de intrones (39). Por supuesto, estas posibilidades son altamente especulativas, y la confirmación vendrá después de secuenciar los fragmentos que se amplifican por PCR para los genes *hb*, y de comparar dichas secuencias con los genes correspondientes en los insectos.

Debido a que las sondas se diseñaron con extremos flanqueantes para *Pst*I y *Sph*I (Tabla III), con el objeto de poder clonar con mayor facilidad los productos de la amplificación, actualmente se desarrollan procedimientos que conduzcan a la obtención de la secuencia de estos fragmentos, lo cual permitirá establecer inequívocamente la presencia de los genes *hb* en las Monocotiledóneas.

V. CONCLUSIONES.

Dentro de los estudios de tipo evolutivo conforme se avanza en el conocimiento de los organismos vivos y de sus antepasados, resultan más evidentes los principios que han permitido edificar una serie de hipótesis con base en la teoría de la evolución de Charles Darwin (87): la unidad y la diversidad. Para cualquier ejercicio mental sano, resulta muy difícil el entender al mundo biológico si no es con base en los términos evolutivos, con su componente directriz: la selección natural. Gracias al crecimiento explosivo que se ha presentado en las últimas décadas en el desarrollo de las ramas experimentales de las ciencias biológicas, tal como la biología molecular, se ha acumulado una gran cantidad de información que ha permitido confirmar sugerencias de tipo evolutivo que se propusieron en el pasado, así como conocer con detalle los mecanismos por los cuales cambian los organismos y sus estructuras, lo que les permite adaptarse al ambiente que les rodea. Tal vez nunca podamos reconstituir fielmente los fenómenos que ocurrieron en el pasado, tal como el origen de nuestro sistema solar, el origen de nuestro planeta, o el origen de la vida en la Tierra, debido a que no estuvimos presentes cuando ocurrieron estos acontecimientos, sin embargo, gracias al entendimiento del mundo orgánico actual lograremos comprender lo que el pasado fue; esto nos dará tranquilidad y satisfacción intelectual en el futuro.

En el presente trabajo se intentó obtener información que contribuya al esclarecimiento del origen y la evolución de los genes *hb* en las plantas. Las Hbs de los animales se han estudiado con mucho detalle (ver 88) y representan uno de los ejemplos más claros en donde la evolución molecular se ha vinculado de manera estrecha con el éxito

adaptativo de los organismos (89). Sin embargo, en el caso de las plantas el estado actual de los conocimientos para los genes *hb* ha permitido alcanzar el nivel de la especulación y de las sugerencias, en tanto no se obtengan datos experimentales adicionales. Por ejemplo, todo parece indicar que los genes *hb* de los animales y de las plantas descienden a partir del mismo ancestro procarionte (ver Introducción), sin embargo, estos genes sólo se habían detectado en *Trema* y en las Dicotiledóneas que nodulan y que fijan el nitrógeno: no se pudieron detectar en otras plantas cuando se utilizaron sondas de genes *hb* específicas o técnicas de espectroscopia de alta sensibilidad (C.A. Appleby, comunicación personal) previamente a este trabajo hasta donde sabemos. Entonces ¿son el resultado de líneas evolutivas horizontales o verticales los genes *hb* de las plantas?

Para el desarrollo de este trabajo se confió totalmente en los presupuestos de las ideas evolutivas y se construyó una secuencia consenso para los genes *hb* de las plantas, con el objeto de detectar las regiones más conservadas, al suponer que estas regiones también deberían estar conservadas en los genes *hb* de otras plantas, en particular aquellas regiones que son esenciales para el funcionamiento de la proteína, como lo es la región SCHBP/H que forma parte del empaque al grupo hemo (32, ver Apéndice 2). La estrategia del uso de las secuencias consenso para sintetizar sondas moleculares se ha utilizado por diversos autores con anterioridad, quienes la han aplicado exitosamente en procedimientos de hibridación y de amplificación por PCR (90, 91).

Así, se seleccionaron las regiones más conservadas después del ejercicio teórico por el que se obtuvo la SCTHBP, y se desarrollaron

procedimientos experimentales que permitieron obtener resultados de los cuales se concluye: a) la sonda SCHBP/H es funcional cuando se hibridiza con los DNAs que provienen de las Leguminosas, y puede ser útil para llevar a cabo estudios moleculares en esta familia de plantas, como lo es la detección de polimorfismos o la determinación de la expresión de los genes *hb* en diversos tejidos de la planta; b) esta sonda hibridiza con los DNAs de las no-Dicotiledóneas, en donde la intensidad mayor de la hibridización es con las Gramíneas, lo que sugiere que estas plantas contienen genes *hb* al igual que las Dicotiledóneas; c) la sonda tiene un "alcance filogenético" amplio, ya que hibridiza con los DNAs de plantas primitivas, como la Magnolia y la Cícada, aunque la intensidad de la señal disminuye en relación con el aumento en el tiempo de divergencia; d) esta sonda se puede utilizar para aislar genes *hb*, o secuencias muy relacionadas, de diversas plantas; y e) esta sonda en conjunto con la que se deriva de la región A, SCHBP/A, son lo suficientemente similares con algunas regiones específicas del DNA de algunas Gramíneas, lo que permite amplificar fragmentos discretos por PCR.

Los resultados de este trabajo sugieren fuertemente que los genes *hb* están presentes en las Gramíneas, y tal vez también en otras plantas vasculares, sin embargo la confirmación definitiva de la identidad de las regiones que hibridizan con la sonda SCHBP/H, y de los fragmentos que se amplifican al utilizar las sondas 399 y 398, tendrá que provenir de la secuenciación de dichos fragmentos. Hasta el momento, todo parece indicar que los genes *hb* son comunes en el reino Plantae y que compartieron el estado ancestral con los genes *hb* de los animales, hace unos 600×10^6 años o más.

Probablemente los genes *hb* de las plantas descienden por línea directa de los genes *hb* de las bacterias. Esta suposición se apoya en el hecho de que ambos genes tienen un valor alto de similitud, particularmente en la región SCHBP/H (32), además de que el contenido de Alanina es alto en la SCTHBP al igual que en las proteínas de los procariontes (70). Tal vez estas dos características, así como la alta afinidad que tienen las Hbs vegetales hacia el O₂ (57), sean vestigios ancestrales que se han conservado en los genes *hb* de las plantas.

Por otro lado, debido a que la región que se utilizó como la sonda SCHBP/H es la más conservada en la SCTHBP, y a que esta sonda (¿universal?) hibridizó con los DNAs de diversas plantas superiores, se propone que se considere a esta región como la "firma" ("signature") para las Hbs de las plantas con la siguiente secuencia: -XAYDXLXXXIK-.

Finalmente, además de la relevancia desde el punto de vista evolutivo, la presencia de los genes *hb* en las monocotiledóneas tiene interés desde una perspectiva fisiológica.

En los últimos años se ha logrado inducir en las raíces de algunas Gramíneas el desarrollo de estructuras semejantes a los nódulos. El procedimiento consistió en tratar al tejido radical de la Avena y del Arróz con celulasas y con polietilén glicol, para infectar posteriormente este tejido con cepas de *Rhizobium leguminosarum* biovar trifolii que se trataron previamente con los flavonoides luteolina y naringenina, que son inductores de los genes *nod* de la bacteria (92); en el segundo caso, se trató a *Azospirillum* con el factor de crecimiento 2,4-D, con el cual se infectaron las raíces del Trigo (93). En ambos casos se encontraron los simbioses en el interior del "tejido nodular", aunque hasta este momento no se ha demostrado que los

microsimbiontes fijan el nitrógeno dentro de dicho tejido.

Por lo tanto, la detección de los genes *hb* en las Gramíneas llevará posteriormente a determinar si estos genes se expresan, y en que tejido lo hacen, de tal manera que el grupo de genes que codifican para las "nodulinas clásicas" se pudieran expresar en estas plantas, y así se pudiera favorecer el establecimiento de una simbiosis que a la postre pudiera ser funcional en estas plantas que son de mucha importancia como fuente de alimento.

VI. BIBLIOGRAFIA.

1. Nagy, B. y L.A. Nagy. (1988). Los Primeros Organismos Vivientes. En: *El Redescubrimiento de la Tierra* (Ed. por Marcorini, E.). CoNaCyT, México. pp. 103-114.
2. Schopf, J.W. (1978). *Inv. y Ciencia*. 26: 58-75.
3. Pace, N.R. (1991). *Cell*. 65: 531-533.
4. Fridovich, I. (1974). *Hor. Biochem. Biophys.* 1: 1-37.
5. Hunt, L.T., S. Hurst-Calderone y M.O. Dayhoff. (1978). Globins. En: *Atlas of Protein Sequence and Structure* (Ed. por Dayhoff, M.O.). Natl. Biomed. Res. Foundation, Washington, USA. pp. 229-250.
6. Perutz, M.F. (1979). *Inv. y Ciencia*. 29: 40-55.
7. Ingram, V.M. (1961). *Nature*. 189: 704-708.
8. Oshino, R., T. Asakura, K. Takio, N. Oshino y B. Chance. (1973). *Eur. J. Biochem.* 39: 581-590.
9. Iwaasa, H., T. Takagi y K. Shikama. (1989). *J. Mol. Biol.* 208: 355-358.
10. Keilin, F.R.S. y J.S. Smith. (1947). *Nature*. 159: 692-694.
11. Ollis, D.L., C.A. Appleby, P.M. Colman, A.E. Cutten, J.M. Guss, M.P. Ventakapa y H.C. Freeman. (1983). *Aust. J. Chem.* 36: 451-468.
12. Appleby, C.A. (1984). *Ann. Rev. Plant Physiol.* 35: 443-478.
13. Appleby, C.A. (1985). Plant Hemoglobin Properties, Function and Genetic Origin. En: *Nitrogen Fixation and CO₂ Metabolism* (Ed. por Ludden, P.W. y J.E. Burris). Elsevier Sci. Pub. Co., NY, USA. pp. 41-51.
14. Arredondo-Peter, R. (1983). *Bol. Educ. Bloq.* 2: 3-7.
15. Arredondo-Peter, R. (1985). *Bol. Educ. Bloq.* 4: 103-107.
16. Bogusz, D., C.A. Appleby, J. Landsman, E.S. Dennis, M.J. Trinick y W.J. Peacock. (1988). *Nature*. 331: 178-180.
17. Appleby, C.A., D. Bogusz, E.S. Dennis y W.J. Peacock. (1988). *Plant Cell Environ.* 11: 359-367.
18. Gó, M. (1981). *Nature*. 291: 90-92.
19. Redmon, J.W., M. Batley, M.A. Djordjevic, R.W. Innes, P.L. Kuempel y B.G. Rolfe. (1986). *Nature*. 323: 632-635.
20. Nap, J.P. y T. Bisseling. (1990). *Science*. 250: 948-954.

21. Dehio, Ch. y F.J. de Bruijn. (1992). *Plant J.* En prensa.
22. de Bruijn, F.J. y J. Schell. (1992). Regulation of Plant Genes Specifically Induced in Developing and Mature Nitrogen Fixing Nodules: *Cis*-acting Elements and *Trans*-acting Factors. En: *Control of Plant Gene Expression* (Ed. por Verma, D.P.S.). The Telford Press, Caldwell, USA. En prensa.
23. Stougaard, J., N.N. Sandal, A. Gron, A. Kühle y K.A. Marker. (1987). *EMBO J.* 6: 3565-3569.
24. Kuhse, J. y A. Pühler. (1987). *Plant Sci.* 49: 137-143.
25. Jensen, E.O., K.A. Marker, J. Schell y F.J. de Bruijn. (1988). *EMBO J.* 7: 1265-1271.
26. Bergman, J.E., E. Preddie, L. Cortes y R. Brousseau. (1991). *Nucl. Acids Res.* 19: 1338.
27. Xue, Z.T., K. Larsen y B.U. Jochimsen. (1991). *Plant Mol. Biol.* 16: 899-906.
28. Jensen, E.O., K.A. Marker e I. S. Villadsen. (1986). *EMBO J.* 5: 843-847.
29. Metz, B.A., P. Welters, H.J. Hoffman, E.O. Jensen, J. Schell y F.J. de Bruijn. (1988). *Mol. Gen. Genet.* 214: 181-191.
30. Stougaard, J., J.E. Jorgensen, T. Christensen, A. Kühle y K.A. Marker. (1990). *Mol. Gen. Genet.* 220: 353-360.
31. Evans, T., G. Fensenfeld y M. Reitman. (1990). *Ann. Rev. Cell Biol.* 6: 95-124.
32. Arredondo-Peter, R. y E. Escamilla. (1991). *Plant Mol. Biol. Rep.* 9: 195-207.
33. Stougaard, J., T.E. Petersen y K.A. Marker. (1987). *Proc. Natl. Acad. Sci. USA.* 84: 5754-5757.
34. Hurrell, J.G.R., K.R. Thulborn, W.J. Bruoghton, M.J. Dilworth y S.J. Leach. (1977). *FEBS Lett.* 84: 244-246.
35. Brown, G.G., J.S. Lee, N. Brisson y D.P.S. Verma. (1984). *J. Mol. Evol.* 21: 19-32.
36. Ohta, T. (1991). *J. Mol. Evol.* 33: 34-41.
37. Lee, J.S. y D.P.S. Verma. (1984). *EMBO J.* 3: 2745-2752.
38. Lewin, R. (1982). *Science.* 217: 42-43.
39. Lewin, R. (1984). *Science.* 226: 328.
40. Davenport, H.E. (1960). *Nature.* 186: 653-654.

41. Coventry, D.R., M.J. Trinick y C.A. Appleby. (1976). *Biochem. Biophys. Acta.* 420: 105-111.
42. Tjepkema, J.D. (1983). *Can. J. Bot.* 61: 2924-2929.
43. Hattori, J. y D.A. Johnson. (1985). *Plant Mol. Biol.* 4: 285-292.
44. Roberts, M.P., S. Jafar y B.C. Mullin. (1985). *Plant Mol. Biol.* 5: 333-337.
45. Appleby, C.A., J.D. Tjepkema y M.J. Trinick. (1983). *Science.* 220: 951-953.
46. Landsman, J., E.S. Dennis, T.J.V. Higgins, C.A. Appleby, A.A. Kortt y W.J. Peacock. (1986). *Nature.* 324: 166-168.
47. McGee, K. (1990/91). *Aust. Nat. Hist.* 23: 554-561.
48. Appleby, C.A., D. Bogusz, E.S. Dennis, A.I. Fleming, J. Landsman y W.J. Peacock. (1988). The Vertical Evolution of Plant Hemoglobin Genes. En: *Nitrogen Fixation: Hundred Years After* (Ed. por Bothe/de Bruijn/Newton). Gustav Fisher, Stuttgart, Alemania. pp. 623-628.
49. Marker, K.A. y N.N. Sandal. (1987). Evolution of the Leghemoglobins. En: *Plant Molecular Biology* (Ed. por von Wettstein, D. y N.H. Chua). Plenum Pub. Corp., NY, USA. pp. 503-507.
50. Bogusz, D., D.J. Llewellyn, S. Craig, E.S. Dennis, C.A. Appleby y W.J. Peacock. (1990). *Plant Cell.* 2: 633-641.
51. Goodman, M., J. Pedwaydon, J. Czelusniak, T. Suzuki, T. Goton, L. Moens, F. Shishikura, D. Walz y S. Vinogradov. (1988). *J. Mol. Evol.* 27: 236-249.
52. Wakabayashi, S., H. Matsubara y D.A. Webster. (1986). *Nature.* 322: 481-483.
53. Koshla, Ch. y J.E. Bailey. (1988). *Mol. Gen. Genet.* 214: 158-161.
54. Riggs, A.F. (1991). *Am. Zool.* 31: 535-545.
55. Dorit, R.L., L. Schoenbach y W. Gilbert. (1990). *Science.* 250: 1377-1382
56. Runnegar, B. (1984). *J. Mol. Evol.* 21: 33-41.
57. Appleby, C.A., E.S. Dennis y W.J. Peacock. (1990). *Aust. Syst. Bot.* 3: 81-89.
58. Walker, J.C.G., C. Klein, M. Schidlowski, W.J. Schopf, D.J. Stevenson y M.R. Walker. (1983). Environmental Evolution of the Archean-early Proterozoic Earth. En: *Earth's Earliest Biosphere: Its Origin and Evolution* (Ed. por Schopf, W.J.). Princeton Univ. Press, Princeton, USA. pp. 187-213.

59. Crane, P.R., E.M. Friis y K.R. Pedersen. (1986). *Science*. **232**: 852-854.
60. Bremer, K. (1985). *Cladistics*. **1**: 369-385.
61. Troitski, A.V., Y.F. Melekhovets, G.M. Rakhimova, V.K. Bobrova, K.M. Valiejo-Roman y A.S. Antonov. (1991). *J. Mol. Evol.* **32**: 253-261.
62. De Benedetto, C., L. De Gara, O. Arrigoni, M. Albrizio y R. Galleroni. (1992). *Plant Sci.* **81**: 75-82.
63. Doyle, J.J. y J.L. Doyle. (1990). *Focus*. **12**: 13-15.
64. Cuniberti, G. y M. Guenza. (1990). *Biophys. Chem.* **38**: 11-22.
65. Sambrook, J., E.F. Fritsch y T. Maniatis. (1989). *Molecular Cloning: A Laboratory Manual*. CSH Lab, CSH, USA. 1659 pp.
66. Birnboim, H.C. y J. Doly. (1979). *Nucl. Acids Res.* **7**: 1513.
67. Davis, B.J. (1964). *Ann. NY Acad. Sci. USA*. **121**: 404-427.
68. Gontijo, N.F., J.C.C. Ribeiro y S.D.J. Pena. (1990). *Focus*. **12**: 55-57.
69. Tungpradabkul, S. y S. Panyim. (1985). Nucleic Acid Hybridization. En: *Application of Genetic Engineering to Research on Tropical Disease Pathogens with Special Reference to Plasmodia* (Ed. por Panyim/Wilairat/Yuthavong). UNDP/World Bank/WHO. pp. 181-218.
70. Doolittle, R.F. (1987). *Of URFs and ORFs: A Primer on How to Analyze Derived Amino Acid Sequences*. University Science Books, Mill Valley, USA. 103 pp.
71. Ayala, F.J. (1978). *Inv. y Ciencia*. **26**: 18-33.
72. Dellaporta, S.L., J. Wood y J.B. Hicks. (1984). Maize DNA Minipreps. En: *Molecular Biology of Plants: A Laboratory Course Manual* (Ed. por Malmberg/Messing/Sussex). CSH Lab, CSH, USA. pp. 36-37.
73. Stern, H. (1968). *Meth. Enzymol.* **12**: 100-112.
74. Mayer, A.M. y E. Harel. (1979). *Phytochemistry*. **18**: 193-215.
75. Brisson, N., A. Pombo-Gentile y D.P.S. Verma. (1982). *Can. J. Biochem.* **60**: 272-278.
76. Beadle, G.W. (1980). *Sci. Am.* **242**: 96-103.
77. Endress, P.K. (1987). *Trends Ecol. Evol.* **2**: 300-304.
78. Briggs, D. y S.M. Walters. (1984). *Plant Variation and Evolution*. Cambridge Univ. Press. Cambridge, Inglaterra. 412 pp.

79. Southern, E.M. (1975). *J. Mol. Biol.* **98**: 503-517.
80. Meinkoth, J. y G. Wahl. (1984). *Anal. Biochem.* **138**: 267-284.
81. Bej, A.K., M.H. Manubani y R.M. Atlas. (1991). *Crit. Rev. Biochem. Mol. Biol.* **26**: 301-334.
82. Compton, T. (1990). Degenerate Primers for DNA Amplification. En: *PCR Protocols - A Guide to Methods and Applications* (Ed. por Innis/Gelfand/Sninsky/White). Academic Press, NY, USA. pp. 39-45.
83. Langeveld, S.A., J.M. Dore, J. Memelink, A.F.L.M. Derks, C.I.M. Van Der Vlugt, C.J. Asjes y J.F. Bol. (1991). *J. Gen. Virol.* **72**: 1531-1541.
84. Perkin Elmer Cetus. (1991). *Biotechnology Catalog*. pp. 8.
85. Tindall, K.R. y T.A. Kundell. (1988). *Biochemistry*. **27**: 6008-6013.
86. Bell, D.A. y D.M. De Marini. (1991). *Nucl. Acids Res.* **19**: 5079.
87. Darwin, Ch. (1959). *El Origen de las Especies*. Edición de 1983. SARPE. Madrid, España. 638 pp.
88. Lesk, A.M. y C. Chothia. (1980). *J. Mol. Biol.* **136**: 225-270.
89. Goodman, M. (1980). Secuencias Proteínicas en la Filogenia. En *Evolucion Molecular* (Ed. por Ayala, F.J.). Omega. Barcelona, España. pp. 146-164.
90. Queralt, R. y R. Oliva. (1991). *Nucl. Acids Res.* **19**: 5786.
91. Snijders, P.J.F., Ch.J.L.M. Meijer y J.M.M. Walboomers. (1991). *J. Gen. Virol.* **72**: 2781-2786.
92. Al-Mallah, M.K., M.R. Davey y E.C. Cocking. (1989). *J. Exp. Bot.* **40**: 473-478.
93. Beardsley, T. (1991). *Sci. Am.* **264**: 14.

A P E N D I C E I

Padilla Acero, J. y R. Arredondo-Peter. (1992). Nodulinas en la Simbiosis *Rhizobium* - Leguminosas. Bol. Educ. Bioq. En prensa.

CONT. BEB V10 No. 1, 1991 = A: BE-10 A

NODULINAS EN LA SIMBIOSIS R h i z o b i u m* - LEGUMINOSA

Jaime Padilla Acerol₁ y Raúl Arredondo Peter₂

1

Departamento de Biología Molecular y Biotecnología de Plantas, Centro de Investigaciones sobre Ingeniería Genética y Biotecnología, UNAM, Apdo. Postal 510-3, 62271 Cuernavaca, Mor.

2

Departamento de Microbiología, Instituto de Fisiología Celular, UNAM, Apdo. Postal 70-242, 04510 México, D.F.

* A lo largo de esta revisión se utiliza el género Rhizobium para referirse a las bacterias fijadoras del nitrógeno que

forman simbiosis con especies de la familia Leguminosae, y que pertenecen a la familia Rhizobiaceae.

RESUMEN

Los procesos morfogenéticos son el resultado de la diferenciación de grupos de células en tejidos especializados. Durante este proceso ocurre la expresión diferencial de determinados genes y, como consecuencia, la síntesis de proteínas específicas. Los nódulos radicales de

las plantas fijadoras del nitrógeno son un modelo interesante de estudio, ya que en ellos se han podido esclarecer muchos procesos morfogénéticos: se han detectado proteínas específicas de estos órganos, su síntesis temporal a lo largo del desarrollo, y los factores que regulan la expresión de los genes que codifican a estas proteínas. En este trabajo los autores intentan presentar un panorama general del conocimiento actual de las proteínas específicas de los nódulos, las nodulinas, así como algunas consideraciones acerca de su función dentro de los tejidos en los que se sintetizan.

GLOSARIO

SDS-PAGE, electroforesis en un gel de poliacrilamida en presencia de agentes desnaturantes como el detergente dodecil sulfato de sodio (SDS).

2D-PAGE, electroforesis en un gel de poliacrilamida que se corre en dos dimensiones: en la primera la separación es por un gradiente de pH hasta que las proteínas alcanzan su punto isoeléctrico (también se le conoce como enfoque isoeléctrico o IEF), y en la segunda la separación es por SDS-PAGE.

CDNA, DNA copia que resulta de la transcripción inversa que utiliza como molde al RNAm, mediante el uso de la enzima transcriptasa reversa.

HIDROPATIA, valor de hidrofobicidad que presenta un residuo de aminoácido en la cadena polipeptídica, de acuerdo con su naturaleza y con la de los demás residuos que interactúan con él. Un perfil de hidropatía muestra las regiones hidrofóbicas e hidrofílicas de una proteína.

GENERALIDADES

Las leguminosas son una familia de plantas con flores, que tienen la capacidad de fijar el nitrógeno, N_2 , al transformarlo en amonio, NH_4^+ , gracias al establecimiento de una relación simbiótica con bacterias de los géneros Rhizobium, Bradyrhizobium, Azorhizobium y Sinorhizobium, que son quienes en realidad llevan a cabo este proceso de fijación. Como resultado de dicha simbiosis, la cual se desarrolla principalmente en el tejido radical de la planta (aunque también se presenta en el tallo), se forman unos órganos especializados que se conocen como "nódulos". Estas estructuras nodulares son el punto culminante de la interacción entre el tejido radical de la planta y el microorganismo; los fenómenos principales que conducen a la formación del nódulo se pueden resumir de la siguiente manera: a) unión de la bacteria a la superficie externa del pelo radical; b) curvamiento del pelo radical; c) formación

del hilo de infección, que es por donde penetran las bacterias al tejido vegetal; d) rediferenciación del tejido radical en una corteza y formación de un primordio de nódulo; e) liberación de las bacterias en el citoplasma de la célula vegetal y formación de una membrana peribacteroidal; f) diferenciación de los rizobia en bacteroides fijadores del nitrógeno.

La formación de los nódulos fijadores del nitrógeno implica la interacción compleja de una gran variedad de factores que se sintetizan, de manera coordinada, por la bacteria y la planta, y que regulan la expresión de genes específicos en ambos simbioses. Por parte de la planta, se han detectado entre 20 y 30 proteínas que se sintetizan exclusivamente en los nódulos radicales, pero no en raíces sin infectar ni en otros tejidos; se conoce como "nodulinas" a los genes, así como a los productos de su expresión, específicos del tejido nodular. Estas nodulinas se expresan de manera diferencial a lo largo del desarrollo del nódulo, por lo que se les divide en nodulinas tempranas y tardías (ENOD, de "Early NODulins", y LNOD, de "Late NODulins", respectivamente).

NOMENCLATURA. Las proteínas específicas de los nódulos pueden ser de origen bacteriano o de la planta. Se conocen como "bacteroidinas" a las que se codifican por el genoma de los rizobia, y "nodulinas" a las de origen vegetal; un tercer grupo son las "simbactinas" que son aquellas que se encuentran en las células de los nódulos radicales pero fuera del bacteroides; éstas se codifican por el genoma de Rhizobium.

Desde el punto de vista experimental, debido a que al extraer estas proteínas de los nódulos se obtienen tanto a las bacteroidinas como a las nodulinas, se han desarrollado estrategias que permiten diferenciar a unas de otras; la más importante de ellas consiste en extraer el RNAm total, y a partir de él, mediante técnicas estándar, obtener el RNA poli(A⁺) característico de las células eucariontes, para sintetizar posteriormente in vitro las proteínas correspondientes, y detectar a las nodulinas entre ellas.

Ya que la caracterización inicial de las nodulinas se realiza por electroforesis en geles desnaturalizantes (SDS-PAGE), se designan con la letra "N" seguida de su peso molecular, de acuerdo con su migración en estos geles. Por ejemplo, la nodulina N45 es una proteína específica de los nódulos con un peso molecular de 45 KDa; esta nomenclatura se sustituye por el nombre adecuado de la proteína, en cuanto se conoce su función. En el caso de que en el nódulo se sintetice una forma específica de alguna enzima que esté presente en algún otro tejido, el nombre se antecede con "n"; por ejemplo, la nodulina N35, que corresponde a la uricasa específica de nódulos, se designa como n-uricasa (2).

TECNICAS DE ESTUDIO.

ENSAYOS INMUNOLOGICOS. Estas fueron las primeras técnicas que se utilizaron para detectar a las proteínas específicas de los nódulos y consisten en obtener anticuerpos, a partir de conejos inmunizados, con las fracciones proteicas del nódulo, para convertirlos, posteriormente, en antisueros específicos del nódulo, por titulación contra las proteínas de raíces no infectadas. El análisis de inmunodetección se realiza mediante el uso de este último suero con las proteínas del nódulo, que se separaron previamente por SDS-PAGE y se transfirieron a filtros de nitrocelulosa (NC).

Este método presenta las siguientes desventajas: a) la existencia de inmunorreactividad cruzada con algunos productos bacterianos que están presentes en el extracto inicial; b) la eliminación de algunas nodulinas que manifiestan isoenzimas o dominios de proteína - epítopes - similares estructuralmente en el extracto de raíces sin infectar; y c) la dificultad para detectar proteínas abundantes que exhiben una capacidad antigénica baja, o las que se encuentran en cantidades pequeñas. Algunos de estos problemas se han resuelto, al obtener anticuerpos dirigidos contra proteínas purificadas del nódulo; de esta manera, ha sido posible obtener antisueros contra la glutamina sintetasa (GS) de nódulos de frijol, la uricasa II, la leghemoglobina (LB), y otras más.

Análisis de los productos de traducción *In Vitro*. Con el fin de establecer la contribución del vegetal en la simbiosis, es necesario utilizar métodos que proporcionen información exclusiva de la expresión genética de la planta en este proceso. Así, se analizan los productos de traducción in vitro provenientes de distintas fracciones del RNAm, que se aísla del nódulo o de la raíz. Se incuba este RNAm en presencia de un sistema reconstituído y funcional de síntesis de proteínas tipo eucarionte, por ejemplo un sistema "libre de células" - como el lisado de reticulocitos de conejo o el de germen de trigo - más aminoácidos (uno de ellos es radiactivo) y cofactores adecuados. Como resultado se obtienen polipéptidos radiactivos que son codificados por los RNAm de la planta, y que se analizan en una placa de autorradiografía, después de su separación por SDS-PAGE o mediante el corrimiento en geles bidimensionales (2D-PAGE). De esta manera es posible detectar a las proteínas específicas de algún tejido, al comparar los patrones de corrimiento entre, por ejemplo, raíz y nódulos, o nódulos radicales y aquellos que se producen en el tallo de la planta (Fig 1) (3).

OJO SUA, AQUI ENTRA LA FIGURA NO. 1.

PIE DE FIGURA:

Figura 1. Autorradiografía de los productos de traducción in vitro del RNA poli(A⁺), aislado de diferentes tejidos de Sesbania rostrata: (a) raíces sin infectar; (b) nódulos radicales de 18 días de desarrollo; (c) tallos sin infectar; y (d) nódulos del tallo de 18 días de desarrollo. La separación de las proteínas fue por 2D-PAGE. Símbolos: cuadros, polipéptidos específicos del nódulo; círculos, polipéptidos estimulados en el nódulo; flechas, polipéptidos estimulados en la raíz o el tallo sin infectar, en relación con otros tejidos de la planta. (Tomado de 3).

Las limitaciones de esta estrategia provienen de la labilidad de ciertos RNAm bajo ciertas condiciones, la escasez de ciertos RNAm, y, dado que la traducción se realiza generalmente en presencia de [³⁵S]-Met, el contenido de este aminoácido en la proteína determina la intensidad de la señal en la placa fotográfica, y no refleja necesariamente su abundancia in vivo.

CLONACION DEL cDNA DE LAS NODULINAS. Otra aproximación para el estudio de la expresión de nodulinas es por el análisis tipo "Northern", esto es, por el reconocimiento de la identidad de los RNAm y las secuencias genéticas de nodulinas que los codifican. Este método requiere del uso de sondas de hibridación contra el RNAm que se ha sometido a electroforesis; para hacer esto, se requiere de la construcción de bancos de cDNA a partir del RNAm poli(A⁺) de los nódulos, raíces u otros tejidos. Estos cDNA se clonan en vehículos apropiados, como plásmidos bacterianos, y se seleccionan los individuos, o clonas, que poseen una secuencia particular, integrada mediante una hibridación diferencial, es decir, se seleccionan las clonas con el cDNA que hibridiza con el cDNA de los nódulos, pero no con el cDNA de raíces sin infectar, o con el de otros tejidos (4).

PLANTAS TRANSGENICAS. Estas técnicas son de gran ayuda en el entendimiento de la organización genética y de la regulación de los genes de las nodulinas, aunque se han estudiado con detalle otros sistemas genéticos mediante esta herramienta. Las plantas transgénicas se generan por la infección de la bacteria Agrobacterium tumefaciens en una planta huésped determinada, generalmente el tabaco (Nicotiana tabacum), en donde la bacteria puede transferir parte de su DNA a la planta receptora mediante el plásmido Ti. Este plásmido se ha manipulado genéticamente de tal manera que se puede utilizar como un vector general para transferir genes de interés a plantas específicas (5). Con base en este sistema ha sido posible transformar plantas de Lotus corniculatus (pata de pájaro) con genes quiméricos de las nodulinas, es decir, fusiones génicas de la secuencia de interés con genes cuya actividad se puede seguir colorimétricamente (β-Galactosidasa, β-Ga; Cloranfenicol Acetil Transferasa, CAT; -β-Ga

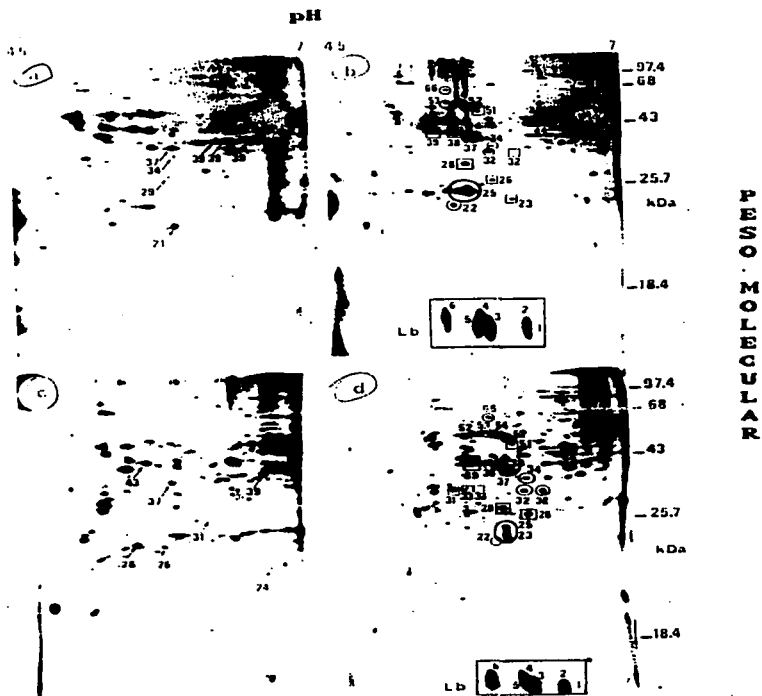


Figura 1.

Glucuronidasa, Gus, entre otros); así, ha sido posible estudiar los elementos genéticos de control que actúan en forma cis (sobre la misma cadena del DNA) y los factores que actúan en trans (proteínas que se unen al DNA) (6), y su expresión en diversos tejidos de la planta.

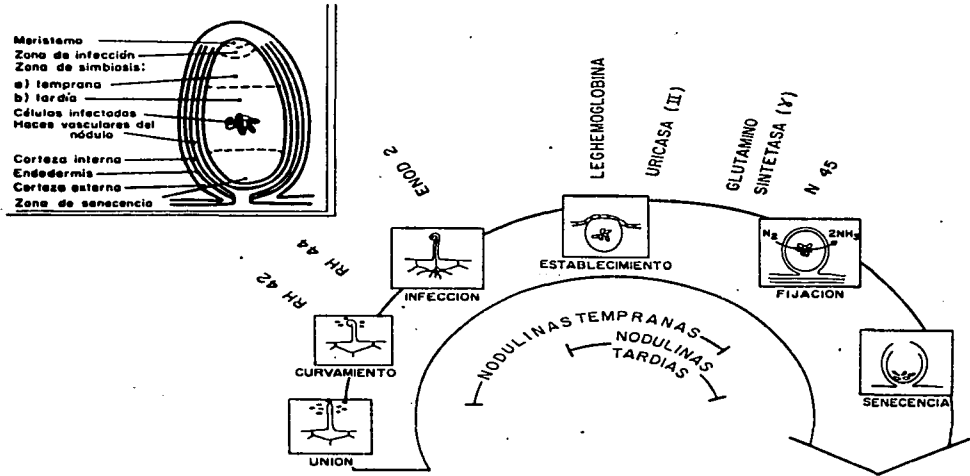
CLASIFICACION. Se ha observado que el programa de desarrollo del nódulo, desde el reconocimiento de la bacteria por el tejido radical, hasta su formación completa, se caracteriza claramente por la expresión de dos grupos de nodulinas (1): a) las nodulinas tempranas, o ENOD, que se expresan durante el proceso de infección, a lo largo de la morfogénesis del nódulo, y se les encuentra específicamente en la porción no infectada de este órgano, que incluye a la corteza; y b) las nodulinas tardías, o LNOD, cuya expresión se detecta alrededor del inicio del proceso de la fijación del nitrógeno, lo cual ocurre cerca del día 11 después de la inoculación de plantas como el frijol o soya, y se les encuentra con mayor abundancia en las células infectadas de la planta (7) - en el recuadro de la figura 2 se ilustran los tejidos que forman al nódulo radical. Entre las nodulinas tempranas que mejor se han estudiado, se encuentran las específicas de los pelos radicales, o proteínas RH (de "Root Hair"), la ENOD2, ENOD13, ENOD55, N75, N75' y N38; y entre las tardías, se encuentran la Lb, la uricasa II (N35), la subunidad α de la glutamino sintetasa (n-GS) y la N45, entre otras (Fig. 2). En la siguiente sección se hablará con mayor detalle de algunas de ellas. En la figura 3 se muestra la expresión diferencial de estas nodulinas durante el desarrollo del nódulo, y su distribución en células infectadas y no infectadas.

OJO SUA: AQUI ENTRAN LAS FIGURAS 2 Y 3. PIES DE FIGURA:

FIGURA 2: Secuencias reguladoras y regiones promotoras de los genes N23 y Lbc3 de la soya. (Tomado de 6).

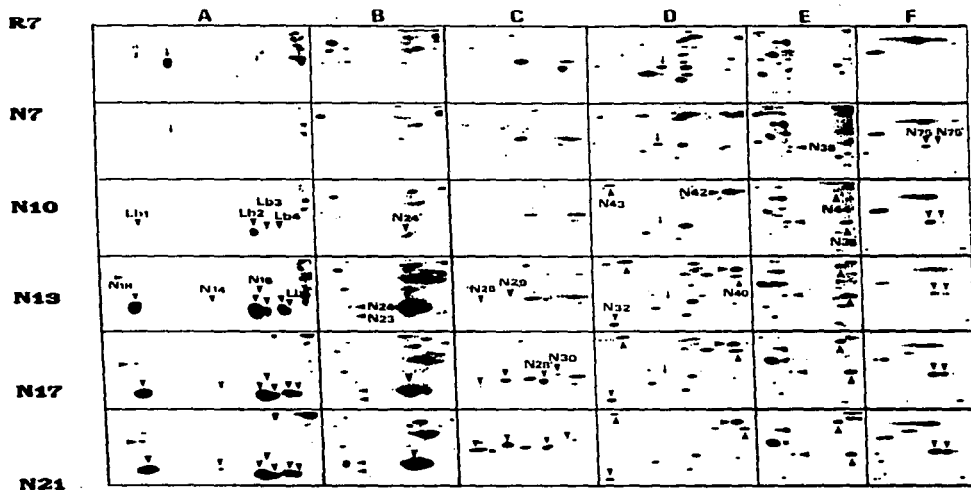
Editor:
Este es el
pie de figura
de la
Fig. 8

FIGURA 3: Expresión de los genes de nodulinas en la soya. (A), durante el desarrollo del nódulo, de 7 a 21 días; R7, raíz sin infectar. (B), en protoplastos de células infectadas, CI; no infectadas, CNI, y en la corteza, CO, en comparación con nódulos totales de 27 días de desarrollo, N. Las columnas se refieren a porciones específicas de los genes después de la separación de las proteínas por 2D-PAGE. (Tomado de 7).



Reverencia a los autores de las primeras figuras,
que sirva de guía a la descripción
en el texto.

Figura 2.



(B)

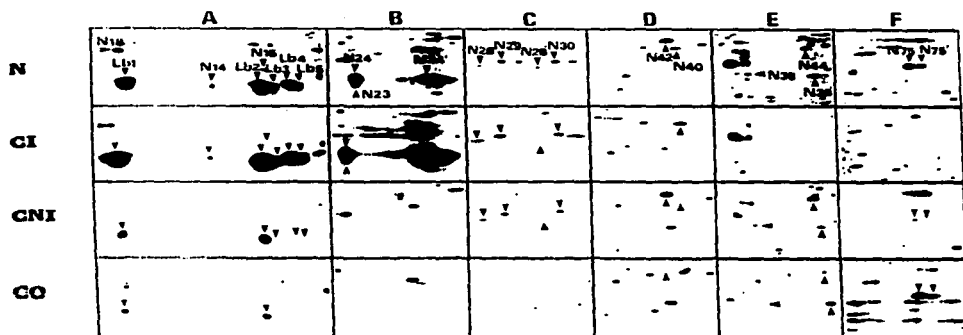


Figura 3.

Además de las 20 a 30 nodulinas que se han detectado gracias a estos estudios, que comprenden el análisis de los productos de traducción in vitro por 2D-PAGE, también ha sido posible encontrar proteínas específicas de la raíz, y otras cuya síntesis se estimula durante el proceso de rediferenciación de la raíz en nódulo (Fig. 3A), así como aquellas nodulinas que se encuentran en los diferentes tipos de células y tejidos que forman estas estructuras especializadas (Fig. 3B) (7).

PRINCIPALES NODULINAS.

NODULINAS TEMPRANAS:

a) Proteínas específicas de los pelos radicales (RH). Estas proteínas se encuentran en la superficie externa de la membrana plasmática de la célula vegetal, y tienen como característica el presentar una naturaleza ácida debido a la abundancia de residuos de Asp y Glu. La síntesis de las proteínas RH-42 y RH-44 aumenta significativamente después de la inoculación de las plantas de chícharo con Rhizobium; mediante el uso de plantas con mutaciones en estos genes, se ha podido relacionar a RH-44 con la deformación del pelo radical, y a RH-42 con el curvamiento de este mismo pelo (1).

b) ENOD2. Esta proteína es rica en residuos de Pro y está compuesta de dos pentapéptidos repetitivos de Pro-Pro-His-Glu-Lis y Pro-Pro-Glu-Tir-Gln, tal vez como resultado de fenómenos de duplicación genética. La secuencia de residuos de aminoácidos de ENOD2 es similar en alto grado a la de la proteína 1A10 de la soya, que se encuentra en la pared celular; como característica importante, los residuos de Pro están hidroxilados a hidroxiprolina en ambas proteínas, por lo que se ha sugerido que ENOD2 corresponde a una proteína de la pared celular (8). De manera análoga, algunas proteínas de origen animal, como la colágena, también son hidroxiladas en los residuos de Pro, entre otros procesos de modificación post-traduccional; sin embargo, a diferencia del grupo de las plantas, la proil hidroxilasa de los tejidos animales es específica hacia el triplete -X-Pro-Gli- cuando hidroxila al residuo de Pro (9).

Como resultado de los estudios de hibridación in situ de las preparaciones de nódulos con [³⁵S]-RNA antisentido que contiene al gen de ENOD2, fue posible detectar con detalle el tejido donde se expresan los genes que codifican para esta proteína (Fig. 4), además de poder estudiar su expresión a lo largo del desarrollo del nódulo. En la corteza interna del nódulo, ENOD2 se encuentra a partir de estados tempranos del desarrollo: los días 6 y 7 después de inocular plantas de soya y de chícharo, respectivamente. Además, ENOD2 se expresa en "nódulos vacíos", es decir, algunas cepas mutantes

de Rhizobium inducen la rediferenciación de la raíz en nódulos sin que exista infección en el interior del tejido; esto sugiere que los genes ENOD2 se expresan específicamente a lo largo de la diferenciación del meristemo nodular en la formación de una corteza interna (8).

OJO SUA: AQUI ENTRA LA FIGURA 4. PIE DE FIGURA:

FIGURA 4. Localización de ENOD2 en nódulos radicales del chícharo (A y B) y de la soya (C y D) por hibridización in situ con una sonda de [S^{35}]-RNA antisentido (pGmENOD2). En las micrografías de campo claro (A y C) la marca radioactiva se detecta como regiones oscuras, y en las de campo oscuro (B y D) como regiones claras. M, meristemo apical; STM, tejido simbiótico temprano; STR, tejido simbiótico tardío; CI, corteza interna; CE, corteza externa; HVN, haces vasculares del nódulo; HVC, haces vasculares centrales; CR, corteza de la raíz; y E, endodermis. (Tomado de 8).

NODULINAS TARDIAS:

a) Leghemoglobina (Lb). Esta es la nodulina mejor caracterizada, sin embargo, debido a que recientemente se le ha detectado en tejido radical no infectado, los autores prefieren considerarla como una proteína "nódulo estimulada", ya que no cumple con los postulados que definen a una nodulina sensu strictu.

b) Uricasa II. Durante mucho tiempo se le conoció como N35, la que corresponde a la subunidad II de la uricasa específica de los nódulos. Su actividad se relaciona con el metabolismo de los ureídos, ya que cataliza la formación de la alantoína a partir del ácido úrico; las Leguminosas tropicales, como el frijol, transportan principalmente ureídos como la alantoína y el ácido alantoico, que son algunos de los productos iniciales del proceso de aminación primaria del nitrógeno fijado; en contraste, las leguminosas de climas templados transportan predominantemente amidas, como la Asn o Gln. En nódulos de frijol, se puede detectar la actividad de la uricasa II a partir del día 11 después de la inoculación (Fig. 5), unos 3 a 4 días antes de que se detecte la fijación de nitrógeno (10).

OJO SUA: AQUI ENTRA LA FIGURA 5. PIE DE FIGURA:

FIGURA 5. Expresión de los genes de uricasa II durante el desarrollo de nódulos del frijol. N, nódulos de diferentes días de desarrollo; R, raíz sin infectar; U, uricasa pura. Detección específica en filtros de NC con suero anti-uricasa. (Tomado de 10).

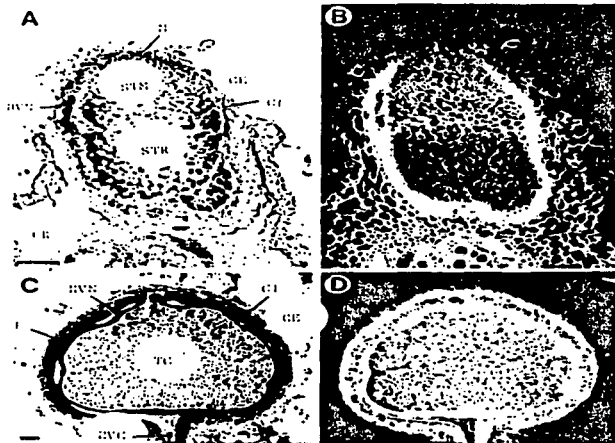


Figura 4.

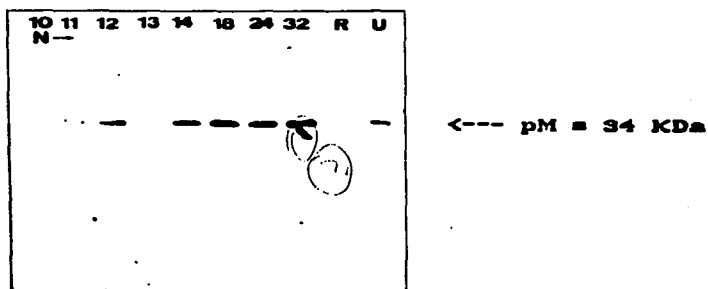


Figura 5.

c) Glutamino Sintetasa. Esta enzima tiene como función el asimilar el amonio fijado y secretado por la bacteria al citoplasma de la célula vegetal, mediante la ruta glutamino sintetasa/glutamato sintasa; es un octámero de unos 320 a 380 KDa de peso molecular. Se ha encontrado una isoforma específica en los nódulos radicales, la GS α -1, que contiene un monómero específico de este tejido; la GS γ se detecta a partir del día 10 después de la inoculación, en plantas del frijol, y su concentración aumenta hasta el día 32 (Fig. 6).

OJO SUA: AQUI ENTRA LA FIGURA 6. PIE DE FIGURA:

FIGURA 6. Expresión de los diferentes componentes de la glutamino sintetasa α , β y γ durante el desarrollo de los nódulos del frijol (10 a 32 días), en comparación con hipocótilos, H2. Separación de los productos de traducción in vitro por 2D-PAGE. (Tomado de 11).

d) N45. Esta nodulina tiene un peso molecular de 45 KDa, y se le detecta a partir del día 11 después de inocular plantas de lupino (Fig. 7). A nivel de la secuencia de nucleótidos, contiene dos regiones que están reiteradas, con alrededor del 80% de similitud entre ellas: los nucleótidos 424 a 484, y 542 a 731, contra los nucleótidos 757 a 817, y 878 a 1067. Mediante el análisis de hidropatía de la secuencia de residuos de aminoácidos, se encontró que N45 es una proteína hidrofílica que contiene un péptido señal hidrofóbico para su translocación a través de la membrana, y sitios con residuos de Asn para glucosilación (12). Se desconoce su función, ya que no se ha encontrado similitud con otras secuencias proteicas, mediante la búsqueda en bancos de datos.

OJO SUA: AQUI ENTRA LA FIGURA 7. PIE DE FIGURA:

FIGURA 7. Expresión de N45 durante el desarrollo de nódulos del lupino (3 a 39 días); R, raíz sin infectar. Detección tipo "Northern": la población del RNA poli(A⁺) se separa por electroforesis y se hibridiza contra la sonda de cDNA que contiene genes de N45. (Tomado de 12).

REGULACION DE LOS GENES DE NODULINAS. Se han detectado factores ambientales que regulan la expresión de los genes de las nodulinas, como la pPO₂ (oxígeno disponible) o la concentración del nitrógeno combinado (NH₄⁺, NO₃⁻), así como factores intrínsecos del tejido vegetal, como elementos que actúan en cis o factores que actúan en trans. El caso mejor estudiado es el de la Lb, pero recientemente ha llamado mucho la atención otra familia de genes que codifica para N20, N22, N23 y N44.

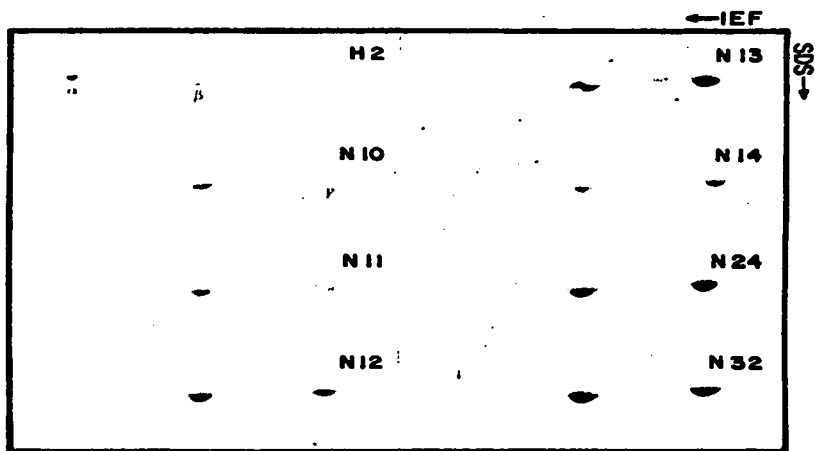


Figura 6.

R 3 7 11 13 18 21 29 39 .

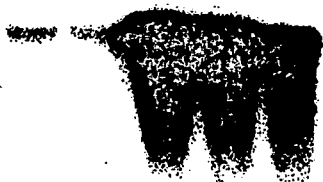


Figura 7.

Mediante la construcción de genes quiméricos, que consisten en la fusión de las regiones 5' y 3' que flanquean al gen de Lbc₁₃ de la soya, con el gen de CAT, y la transferencia a la planta heteróloga Lotus corniculatus, se encontró que el gen se expresa únicamente en los nódulos de las plantas regeneradas y no en otros tejidos; la secuencia que confiere la especificidad en la expresión se encuentra en el segmento 5'. En el caso de N23, las deleciones en la región 5' flanqueante con la enzima Bal31 de 97 pares de bases (pb) en la posición -247, reduce la actividad de CAT en un 15%, y la eliminación de otras 82 pb en -165 nulifica la actividad de CAT. De esta manera, se reconocen dos secuencias que se requieren para la expresión específica de N23 en los nódulos: -344 a -247, y -247 a -165. En el caso de Lbc₁₃ de la soya, se han detectado varios elementos reguladores: -1100 a -950, en donde se encuentra un elemento positivo fuerte; -230 a -170, un elemento positivo débil; -139 a -102, un elemento órgano específico; y -102 a -149, un elemento negativo (Fig.8) (6).

OJO SUA: AQUI ENTRA LA FIGURA 8. PIE DE FIGURA:

FIGURA 8. Representación general de los eventos que conducen a la formación del nódulo, y la expresión de algunas nodulinas tempranas y tardías. Recuadro: principales tejidos que forman al nódulo radical. (Modificado de 1).

Editor:
Este es el
Pie de figura
de la
Fig. 2

Además, mediante ensayos por deleciones con Bal31 y experimentos por retardamiento en geles, en donde se incubaba la secuencia de interés con extractos del tejido y se determina la alteración en la migración electroforética por la posible interacción DNA-proteína, se han detectado factores que actúan en trans y las regiones promotoras de Lbc₁₃ de la soya que reconocen: -342 a -191, y -190 a -125. Estas regiones contienen secuencias ricas en A y T, de 24 y 16 pb para la primera y la segunda, respectivamente. El factor que actúa en trans para la Lbc₁₃ de la soya también se ha encontrado en extractos de nódulos de Sesbania rostrata (dormilona) y de Medicago sativa (alfalfa), éstos con la capacidad de reconocer a las regiones reguladoras de Lbc₁₃ de la soya (6), por lo que se piensa que las señales que regulan la actividad de estos genes están conservadas en las leguminosas y responden a lo que pudiera ser un factor universal.

AGRADECIMIENTOS. Los autores desean agradecer al Sr. Arturo Franco T. el trabajo de ilustración y fotografía.

REFERENCIAS.

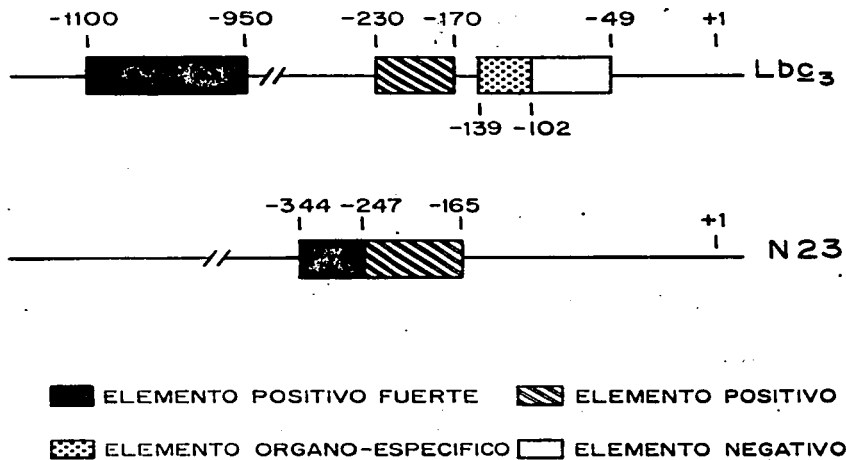


Figura 8.

1. Gloude-mans, T.y T. Bisseling. (1989). Plant Gene Expression in Early Stages of Rhizobium-Legume Symbiosis. Plant Sci. 65: 1-4.
2. van Kammen, A. (1984). Suggested Nomenclature for Plant Genes Involved in Nodulation and Symbiosis. Plant Mol. Biol. Rep. 2(2): 43-45.
3. de Lajudie, P.y T. Huget. (1988). Plant Gene Expression During Effective and Ineffective Nodule Development of the Tropical Stem-Nodulated Sesbania rostrata. Plant Mol. Biol. 10: 537-548.
4. Conkling, M.A., Ch. Cheng, Y.T. Yamamoto y H.M. Goodman. (1990). Isolation of Transcriptionally Regulated Root-Specific Genes from Tobacco. Plant Physiol. 93: 1203-1211.
5. Schell, J.S. (1987). Transgenic Plants as Tools to Study the Molecular Organization of Plant Genes. Science 237: 1176-1182.
6. Jensen, E.O., J.S. Stougaard, J.E. Jorgensen, N. Sandal, F.J. de Bruijn, J. Schell y K.A. Mareker. (1988). Regulation of Nodule-Specific Plant Genes. En: Nitrogen Fixation, Hundred Years After (editado por Bothe/Newton). Gustav Fisher, NY. pp. 605-609.
7. Kouchi, H., M. Tsukamoto y S. Tajima. (1989). Differential Expression of Nodule-Specific (Nodulin) Genes in the Infected, Uninfected and Cortical Cells of Soybean (Glycine max) Root Nodules. J. Plant Physiol. 135: 608-617.
8. van de Wiel, C., B. Scheres, H. Franssen, M.J. van Lierop, A. van Lammeren, A. van Kammen y T. Bisseling. (1990). The Early Nodulin Transcript ENOD2 is Located in the Parenchyma (Inner Cortex) of Pea and Soybean Root Nodules. EMBO J. 2(1): 1-7.
9. Kivirikko, K.I. (1980). Post-translational Modifications of Collagen. En: Gene Families of Collagen and Other Proteins (editado por Prockop/Champe). Elsevier North Holland, NY. pp. 107-119.
10. Sánchez, F., F. Campos, J. Padilla, J.M. Bonneville, C. Enríquez y D. Caput. (1987). Purification, cDNA Cloning, and Developmental Expression of the Nodule-Specific Uricase from Phaseolus vulgaris L. Plant Physiol. 84: 1143-1147.
11. Padilla, J.E., F. Campos, V. Conde, M. Lara y F. Sánchez. (1987). Nodule-Specific Glutamine Synthetase

12. Szczyglowski, K., L. Boron, U. Szybiak-Strozycka y A.B. Legocki. (1989). Characterization of cDNA Clone Coding for Nodulin-45 from Yellow Lupine (Lupinus luteus). Plant Sci. 65: 87-95.

A P E N D I C E 2

Arredondo-Peter, R. y E. Escamilla. (1991). A Consensus Sequence of Plant Hemoglobins. *Plant Mol. Biol. Rep.* 9: 195-207.

Genetic Resources

A Consensus Sequence of Plant Hemoglobins

R. Arredondo-Peter and E. Escamilla

Departamento de Microbiología, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Apdo. Postal 70-242, 04510 México, D. F., México

Key Words: Plant hemoglobins, leghemoglobin, consensus sequence, computer methods

Abstract: From the study of plant hemoglobin protein and gene sequences, a consensus sequence was constructed and analysed by computer methods; analysis comprised the amino acid composition, hydropathy profile, and match degree with other plant and non-plant hemoglobin sequences. Resulting consensus sequence shows the main features of hemoglobin genes and proteins in plants.

Accession no. D64927

Hemoglobins (Hb) are widespread in the animal kingdom functioning as oxygen carriers and as storage source of oxygen. As part of the symbiotic nitrogen fixation process, it is found also in the plant kingdom, where it has similar functions to those described for animal Hb. This plant Hb is known as leghemoglobin (Lb) since it was first detected in root nodule tissues of legumes (Keilin and Smith, 1947). Although the secondary and tertiary structures of animal and plant Hb are highly conserved (Lesk and Choithia, 1980), variation at the level of primary structure is commonly observed.

Hb genes and transcripts have been detected in several plant species belonging to families other than *Leguminosae*; most of them are able to nodulate and thus to fix dinitrogen (Tjepkema, 1983). At the genetic level, the use of molecular probes coding for specific Lbs has been an essential tool in the search for Hb sequences in several plants, but due to

Abbreviations: IVS, intervening sequence; Lb, leghemoglobin; Hb, hemoglobin.

divergence events (variation among residues), these probes have been limited in their use (C. A. Appleby, personal communication). Cloning of the Hb gene from *Parasponia* (an *Ulmaceae* nodulated by *Rhizobium* species) led to the detection of these genes in dicot plants which are not nodulated and therefore unable to fix dinitrogen (Appleby et al., 1988). Jeffreys (in Lewin, 1982) previously proposed that plant Hb resulted from horizontal gene transfer from a phytophagous insect, mediated by a viral vector; analysis of insect Hb genes showed, however, that they lack introns (Lewin, 1984). A variety of genes for vertebrate and plant Hb are interrupted by two and three introns, respectively. Today it is widely accepted that animal and plant Hb are descended from a common ancestor, probably a prokaryote, more than 600 my ago; the bacterium *Vitreoscilla* is able to synthesize a dimeric Hb (Wakabayashi et al., 1986).

This work reports the construction of consensus sequences for genes and proteins of Hb from plant sources. The resulting consensus shows the more relevant features of Hb in the plant kingdom.

Materials and Methods

We studied sequences of 16 plant Hb protein and six gene sequences (Table I). Sequence alignment was by the MATCH program (Johnson and Doolittle, 1986) from the NEWAT'85 program, kindly donated by Dr. R. Doolittle (University of California, San Diego). From this comparison, consensus sequences at the protein and DNA levels were separately constructed. This procedure was useful in determining with certainty which residues occupied hypervariable positions. Due to divergence between plant sources, different residues could be positioned at specific positions; for example, in the protein consensus either P (Pro) or D (Asp) residues could be placed at position 59, and the corresponding triplet in the gene consensus might be C or T (position 178), C or T (179), or C or G (180) (see Figs. 1 and 2). Because only one residue must occupy a position, by correlating the consensus sequences of the proteins and genes, specific residues were selected. In this example, P is coded by CCC and D by GAC; CCC can occur as several possibilities: C-T, C-T, or C-G combinations, but not GAC. Thus, P was selected for the protein consensus (position 59) and codon CCC for the gene consensus (positions 178-180). Conserved and variable residues were simply selected by placing the most abundant residues at specific positions. The NEWAT'85 program was also used for tallying the amino acid composition of the

Table I. Hemoglobin proteins and genes from legumes and non-legumes.

Plant Source	Protein	Reference
Nonlegumes		
<i>Trema tomentosa</i> Hb ¹	TrHB	Bogusz et al., 1988
<i>Parasponia andersonii</i> Hb ²	PaHB	Appleby et al., 1988a
<i>Casuarina glauca</i> Hb ²	CaHB	Kortt et al., 1988
Legumes		
<i>Lupinus luteus</i> LbI ³	LuLBI	Konieczny, 1987
<i>L. luteus</i> LbII ³	LuLBI	Konieczny et al., 1987
<i>Sesbania rostrata</i> LbII	SeLBI	Kortt et al., 1987
<i>S. rostrata</i> LbVII ¹	SELBVII	Welters et al., 1989
<i>Medicago sativa</i> LbI	MELBI	Davidowitz et al., 1989
<i>M. sativa</i> LbIII	MELBIII	Kiss et al., 1987
<i>Phaseolus vulgaris</i> Lba	PHLb _a	Lehtovaara et al., 1980
<i>Glycine max</i> Lba ²	GLLb _a	Marker & Sandal, 1987
<i>G. max</i> Lbc ²	GLLbc1	Marker & Sandal, 1987
<i>G. max</i> Lbc ²	GLLbc2	Marker & Sandal, 1987
<i>G. max</i> Lbc ²	GLLbc3	Marker & Sandal, 1987
<i>Vicia faba</i> LbI	VILBI	Richardson et al., 1975
<i>Pisum sativum</i> LbI ³	PILBI	Lehtovaara et al., 1980
Plant Source	Gene²	Reference
Nonlegumes		
<i>Trema tomentosa</i> Hb	<i>trhb</i>	Bogusz et al., 1988
<i>Parasponia andersonii</i> Hb	<i>pahb</i>	Appleby et al., 1988a
Legumes		
<i>Sesbania rostrata</i> LbVII	<i>selbvII</i>	Welters et al., 1989
<i>Lupinus luteus</i> LbI	<i>lulbi</i>	Konieczny, 1987
<i>Medicago sativa</i> LbI	<i>melbi</i>	Davidowitz et al., 1989
<i>M. sativa</i> LbIII	<i>melbIII</i>	Kiss et al., 1987

¹Sequences predicted from gene sequences only; proteins were not sequenced.
²A gene sequence was used to correct error-ridden protein sequences.
³Editor's Note: The nomenclature for hemoglobin genes used in this article is tentative and may be revised subsequently to conform to the recommendations of the ISPMB Commission on Plant Gene Nomenclature.

Fig. 1. Plant Hb protein sequences. Symbols key: *, conserved residue; raised dot (·), variable residue; not underlined, hypervariable residue; #, conserved residue between plant and animal Hbs; and ^, intron. One-letter amino acid code is used. Names of proteins are listed in Table I.

```

1
TRHB      S S S E V D K V F T E E Q E A L V V K S W A V M K K N S
PAHB      S S S E V N K V F T E E Q E A L V V K A W A V M K K N S
CAHB      - - - - - - A L T E K Q E A L L K Q S W E V L K Q N I
LULBI     - - - - - - G V L T D V Q V A L V K S S F E E F N A N I
LULBII    - - - - - - G A L T E S Q A A L V K S S W E E F N A N I
SELBII    - - - - - - G - F T D K Q E A L V N A S Y E A F K K N L
SELBVII   - - - - - - G - F T E K Q E A L V N A S Y E A F K Q N L
MELBI     - - - - - - M S F T D K Q E A L V N S S W E A F K Q N L
MELBIII   - - - - - - M G - F T D K Q E A L V N S S W E S F K Q N -
PHLba     - - - - - - G A F T E K Q E A L V N S S W E A F K G N I
GLLba     - - - - - - V A F T E K Q E A L V S S S F E A F K A N I
GLLbc1    - - - - - - G A F T E K Q E A L V S S S F E A F K A N I
GLLbc2    - - - - - - G A F T E K Q E A L V S S S F E A F K A N I
GLLbc3    - - - - - - G A F T D K Q E A L V S S S F E A F K T N I
VILBI     - - - - - - G - F T E K Q E A L V N S S S Q L F K Q N P
PILBI     - - - - - - G - F T D K Q E A L V N S S S E - F K Q N L

      . . . . .
CONSENSUS      G A F T E K Q E A L V N S S M F E A F K Q N I

29
A E L G L K F F L K I F E I A P S A K N L F S Y L K D S P I P L E Q N P P K
A E L G L Q F F L K I F E I A P S A K N L F S Y L K D S P V P L E Q N P P K
P A H S L R L F A L I I E A P E S K Y V F S L K D S N E I P E N N P K
P K N T H R F F T L V L E I A P G A K D L F S F L K G S S E V P Q N N P D
P K H T H R F F I L V L E I A P A A K D L F S F L K G T S E V P Q N N P E
P G H S V L F Y S F I L E K E P A A K G L F S F L K D S D G V P Q N N P S
P G N S V L F Y S F I L E K A P A A K G M F S F L K D F D E V P Q N N P S
P R Y S V F F Y T V V L E K A P A A K G L F S F L K N S A E V - Q D S P Q
P G N S V L F Y T I I L E K A P A A K G M F S F L K D S A G V - Q D S P K
P Q Y S V V F Y T S I L E K A P A A K N L F S F L - - A N G V D P T N P K
P Q Y S V V F Y T S I L E K A P A A K D L F S F L - - A N G V D P T N P K
P Q Y S V V F Y N S I L E K A P A A K D L F S F L - - A N G V D P T N P K
P Q Y S V V F Y T S I L E K A P A A K D L F S F L - - S N G V D P S N P K
P Q Y S V V F Y T S I L E K A P V A K D L F S F L - - A N G V D P T N P K
S N Y S V L F Y T I I L Q K A P T A K A M F S F L K D S A G V D S - P K
F G Y S V L F Y T I I L E K A P A A K G L F S F L K D T A G V E D S - P K

      . . . . .
P Q Y S V V L F Y T S I L E K A P A A K N L F S F L K D S N G V P Q N N P K
G      ^      ^      ^      ^      ^      ^      ^      ^      ^      ^      ^

```


Fig. 2, continued.

100
GCT GAA CTG GGT CTT AAA TTC TTC CTC AAG ATA TTT GAG ATT GCA CCG TCT GCC
GCT GAA CTG GGT CTT CAA TTC TTC CTC AAG ATA TTT GAG ATT GCA CCG TCT GCC
CCT GGC AAC AGT GTT CTG TTT TAC AGC TTG ATA TTG GAG AAA GCA CCA GCA GCA
CCT AAA AAC ACC CAT CGT TTC TTC ACC TTG GTA GTA GAG ATT GCA CCA GCA GCA
CCT GGT AAT AGT GTT TTG TTC TAC ACT ATA ATA TTG GAG AAA GCC CCT GCA GCA
CCT CGT TAT AGT GTT TTT TTC TAC ACT GTC GTA TTA GAG AAA GCA CCT GCA GCA
G A A T C I T C A G T
C AAT CAC A G TAA A ATC TTA A A AA A G A .A
GCT CCC TTG GGT CTT AGG TTC TTC CGT GAC GTA TTG GAG ATT GCA CCT TCT GCC

150
AAG AAC TTG TTC TCG TAT TTG AAG GAC TCT CCG ATT CCT TTG GAG CAG AAC CCA
AAG AAC TTG TTC TCT TAT TTG AAG GAC TCT CCG GTT CCT TTG GAG CAG AAC CCA
AAG GGC ATG TTC TCT TTT CTA AAG GAC TTT GAT GAA GTG CCC CAA AAT AAT CCC
AAG GAT TTG TTC TCA TTT TTG AAG GGA TCT AGT GAA GTA CCC CAG AAT AAT CCT
AAG GCG ATG TTC TCT TTT CTT AAG GAC TCA GCT CCA GTA --- CAA GAT ACC CCT
AAA GGT TTG TTC TCT TTT CTT AAG AAC TCA GCT GAA GTA --- CAA GAT AGT CCT
G G C T A T T G G E A A G T
AAG AAT ATG TTC TCG TAT CTA AAG GAC TCA AAG GGT TA GTT CCC C G A T AT A

200
AAG CTG AAG CCC CAT GCT ATG ACT GTC TTC GTT ATG ACG TGT GAA TCT GCG GTT
AAG CTC AAG CCC CAT GCT ACG ACT GTC TTC GTT ATG ACA TGT GAG TCT GCG GTT
AGT CTC CAG GCC CAT GCT GAA AAG GTT TTT GCA CTG GTG CGT GAT TCC GCC GCT
GAT CTT CAG GCC CAT GCT GAA AAG GTT TTT AAG TTG ACT TAC GAA GCA GCA ATT
AAA CTC CAA AGC CAT GCT GAA AAA GTT TTT GGA ATG GTG CGC GAT TCA GCT G C
CAA CTC CAA GCT CAC GCT GAA AAA GTT TTT GGA CTG GTG CGT GAC TCA GCT GTT
A A C G A A G A A
G T T C G C G A A G T T T T C G T T C T T
CAG CTG AAA ACC CAT GCT AAG ACT GTC TTC GAG TTG ACA TGC GA TCC GC GTT

250
300
CAA CTT CCG AAA GCC GGA AAA GTA ACA GTG AGA GAA TCA AAC TTG AAA AGA CTT
CAA CTT CCG AAA GCC GGA AAA GTG ACA GTG AAA GAA TCA GAC TTG AAA AGA ATT
CAA CTC CGA GCA ACC GGA GTA GTG GTA TTA GCA GAT GCA AGT --- --- TTA
CAA CTT CAA GTG AAT GGA GCA GTG GCT TCA --- GAT GCC ACG TTG AAA AGT TTG
TCA ACT CCG AGC AAC AGG AGG AGT AGT TTT GGG AGA GCT ACA --- --- TTG
CAA CTT CG AGC AAC AGG GGG GGT AGT TTT GGG AGA GCT ACG --- --- TTG
G A G C T G A C
CG A C AA G A A A TT AHT T A CCA GAA G T A
CAA CTT CAA G G GCC AGG G G GGA GTA GTG AAG AGT TCC AG TTG AAA AGA CTA

Fig. 2, continued.

350

GGG GCT ATC CAC TTC AAA AAT GGC GTA GTT AAT GAA CAT TTT GAG GTC ACA AGG
 GGG GCT ATC CAC TTC AAA ACT GGC GTA GTT AAT GAA CAT TTT GAG GTC ACA AGG
 GGT TCT GTC CAT GTC CAA AAA GGA GTC CTT GAT CCT CAC TTT GIG GTG GTT AAA
 GGT TCT GTC CAT GTC TCA AAA GGA GTC GTT GAT GCC CAT TTT CGG GTG GTG AAG
 GGT GCA ATC CAC ATT CAG AAA GGA GTT GTT GAT CCT CAT TTT CCG GTG GTT AAA
 GGT GCT ATC CAC GTT CGA AAG GGA GTT GTT GAT CCT CAT TTT GTG GTG GTT AAA
 G CA A T T T
 T G A C T C AG T A C G GCA C G GTA AA
 GGG TCT GTC CAT ATT TAA AAG GGC GTA GTT AAT CAC CAT TTT GAG GTC ACG AGG

400

TTT GCA CTT TTG GAG ACC ATA AAG GAA GCA GTA CCA GAA ATG TGG TCA CCT GAG
 TTT GCA CTT TTG GAG ACC ATA AAG GAA GCA GTA CCA GAA ATG TGG TCA CCT GAG
 GAA GCA TTG TTG AAA ACA CTA AAG GAA GGC GCA GGG GCC ACT TGG AGC GAT GAA
 GAA GCA ATC CTG AAA ACA ATA AAG GAA GTG GTG GGA GAC AAA TGG AGC GAG GAA
 GAA GCT CTG CTG AAA ACA ATA AAG GAA GTA TCA GGA GAT AAA TGG AGC GAA GAA
 GAG GCT TTG CTC AAA ACA ATA AAA GAA GCA GCT GGA GAC AAA TGG AGT GAA GAG
 A C G A A C AA C T
 GAT A T C A A A CG CG A TG AGA GAA G
 TTG GCT ATC TTG GAG ACC ATA AAG GAA GTC GTT CCA GAT ACT TGG TCT CCG GAA

450

ATG AAG AAC GCA TGG GGA GAA GCT TAT GAT CAG TTG GTT GCT GCT ATC AAG TCC
 ATG AAG AAC GCA TGG GGA GTA GCT TAT GAT CAG TTG GTT GCT GCC ATC AAG TTC
 CTG AGC AAT GCT TGG GAA GTA GCC TAT GAT GAA TTG TCA GCT GCA ATT AAG AAG
 CTG AAC ACT GCT TGG ACC ATA GCC TAT GAC GAA TTG GCA ATT ATA ATT AAG AAG
 TTG AAC ACT GCT TGG GAA GTA GCC TAT GAT GCA TTG GCA ACT GCA ATT AAG AAG
 CTT AAC ACT GCT TGG GAA GTA GCT TAT GAT GCA CTG GCA ACT GCA ATT AAA AAA
 C A A A AG
 A C CT T G C GAA CA A C T ATC
 TTG AAG AAC GCA TGG GCA GTA GCT TAT GAT CCG TTG GTT GCT GCT ATC AAG TCA

470

GAA ATG AAA CCC TCC AGT ACT
 GAA ATG AAA CCC TCC AGT ACT
 GCA ATG AGT
 GAG ATG AAG GAT GCT GCT
 GCA ATG GTT
 GCA ATG AGT
 AT
 C GA CCC T C AG
 GAA ATG ATG GAT GCT GCT ACT

protein consensus sequence (COMP program), as well as for plotting the hydropathy profile (GREASE program) and for searching homologies against the protein sequence database collection of NEWAT'85 (SCANA program). These procedures made it possible to determine the following features of the consensus sequence: conserved, variable and hypervariable regions; residues that take part in α -helices, as well as those that are in the vicinity of the heme moiety; intron (IVS) positions and their genetic flanking sequences; and presence of uncommon amino acid residues in various plant Hb.

Results and Discussion

Figure 1 shows the alignment of protein sequences from several plant sources and the resulting consensus sequence; Figure 2 shows the alignment of six plant genes coding for Hb and the resulting consensus. (These still share hypervariable residues at many positions). The combined consensus sequence for DNA and protein are shown in Figure 3, where variable residues have been determined by procedures described in Materials and Methods. The full consensus sequence corresponds to 483 nucleotides or 161 amino acids. The most abundant amino acid in the consensus sequence is Ala (13.04%) and the estimated molecular weight is 17,481 Da, which is in good agreement with that of various monomeric plant Hb. Figure 3 shows those regions that participate in the seven α -helices. (Plant Hb lacks the authentic D-helix of animal Hb). The most conserved amino acid residues are at 9 to 17, helix A; 43 to 54, a fragment into helix C; and 123 to 153, helix H. By analyzing the frequency of conserved residues, we conclude that helix H is a highly conserved region in the structure of plant Hb. The hydropathy profile of the consensus (Figure 4) shows that the most conserved segment of helix A is hydrophilic, in contrast with those of helices C and H, which are hydrophobic.

Introns (IVS) of plant Hb are highly conserved in their location (Appleby et al., 1988) but flanking sequences are variable; the proposed consensus for three IVS are: IVS-I, ATG ATA; IVS-II, CTG GTG; and IVS-III, GTG GTG (variable and hypervariable nucleotides are bold). By the analysis of amino-acid substitution rates, others have concluded (Lee et al., 1983) that plant Hb has evolved at a higher rate than animal Hb. These genes are conserved, however, among plants, reaching up to 54.9% residues from the whole consensus. The most conserved residues

Fig. 3. Proposed consensus for gene and protein sequences of plant Hb.
 Nucleotide key: Y, variable; X, hypervariable. Amino acid key: bold, conserved;
 underlined, variable; hypervariable are not underlined; ^, intron location; #,
 conserved amino acid residues between plant and animal Hb.

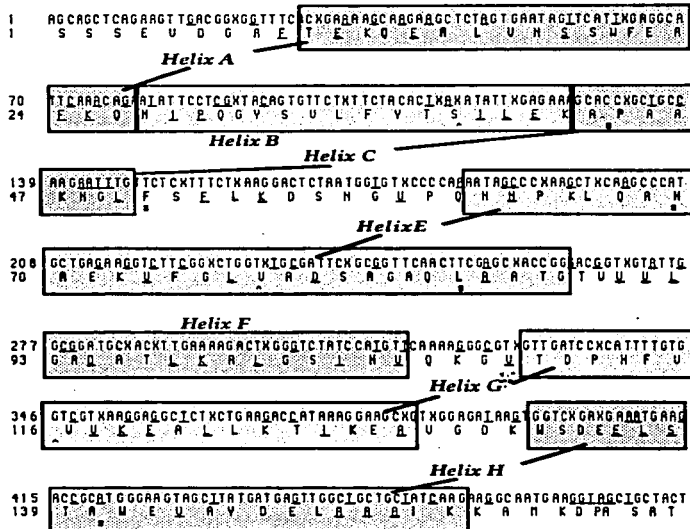
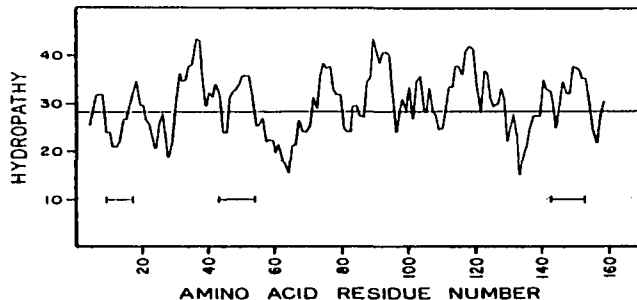


Fig. 4. Hydropathy profile of the plant Hb consensus sequence. Bars indicate the most conserved regions.



between plant and animal Hb are those in the vicinity of the heme moiety (F50, H69, and F114), and P44, L84 and A139 (Figure 3).

Interestingly, the consensus sequence of plant Hb is more homologous to *Vitreoscilla* Hb than to animal Hb, which is significant from the evolutionary point of view. Table II shows the protein Hb sequences that share matches with the consensus sequence from the database collection of NEWAT'85.

From a comparison of protein sequences of 16 plant Hb, it is seen that only three (those from *Trema*, *Parasponia*, and *Casuarina*—all nonlegumes) share C (Cys) residues (position 77 in consensus sequence); these cysteines might be substituted by R (Arg) residues in the legume lineage during evolution. These cysteine residues are, moreover, adjacent to IVS7, which is absent from genes for animal Hb.

H

Conclusions

A limiting factor in the detection of Hb coding sequences in a variety of plant species has been the non-availability of appropriate DNA probes. The consensus sequence proposed in this work might be useful for the synthesis of probes for gene searches for Hb in diverse plant

Table II. Homology between the consensus sequence of plant Hb proteins and that of Hb from other organisms.

<u>Source</u>	<u>Globin¹</u>	<u>Homology Value^{2,3}</u>
Bacterial		
<i>Vitreoscilla</i>	Hemoglobin	20
Plant		
<i>Parasponia andersonii</i>	Hemoglobin I	70
<i>Lupinus luteus</i>	Leghemoglobin	94
<i>Glycine max</i>	Leghemoglobin A	102
Animal		
Lace monitor lizard	Myoglobin	3
Sperm whale	Myoglobin	4
Horse	Myoglobin	5
Human	Myoglobin	4
Rabbit	Hemoglobin β chain	6
Human	Hemoglobin β chain	6
Human	Hemoglobin γ chain	7

¹From NEWAT '85 data base collection.

²Number of 15 segments with at least 6 matches (SCANA program).

³Perfect fit homology value would be 161 from matching consensus against itself.

species. This is an important goal in so far as it could offer support for the hypothesis of vertical evolution of plant and animal Hb (Appleby et al., 1988a). Since, moreover, it has recently become possible to induce root nodule-like structures on roots of monocot species (Al-Mallah et al., 1989), the prior detection of Hb gene sequences in such monocots might facilitate the selection of potential monocot hosts for the induction of nitrogen-fixing symbiosis.

Acknowledgments: To Dr. C. A. Appleby (CSIRO, Canberra) for his discussion and interest in this work. Financial support was from DGAPA-IN02044789UNAM. RAP is a recipient fellowship of CoNaCYT (register number 49199).

References

- Al-Mallah, M. K. M., R. Davey and E. C. Cocking. 1989. Formation of nodular structures on rice seedlings by rhizobia. *J. Exp. Bot.* 40:473-478.
- Appleby, C. A. D., Bogusz, E. S., Dennis, E. S., and W. J. Peacock. 1988. A role for hemoglobin in all plant roots? *Plant Cell. Environ.* 11:359-367.
- Appleby, C. A. D., Bogusz, E. S., Dennis, A. I., Fleming and W. J. Peacock. 1988a. The vertical evolution of plant hemoglobin genes. In: *Nitrogen Fixation, One Hundred Years After.* (eds. Bothede BruijnNewton), pp. 623-628. Gustav Fisher, New York.
- Bogusz, D., Appleby, J., Landsman, E. S., Dennis, M. J., Trinick and W. J. Peacock. 1988. Functioning hemoglobin genes in non-nodulating plants. *Nature.* 331:378-380.
- Davidowitz, E. J. and N. Lang-Utshch. 1989. Nucleotide sequence of a cDNA clone encoding a leghemoglobin from *Medicago sativa*. *Nucl. Acids Res.* 17:3307.
- Johnson, M. S. and R. F. Doolittle. 1986. A method for the simultaneous alignment of three or more amino acid sequences. *J. Mol. Evol.* 25:267-278.
- Kellin, F. R. S. and J. S. Smith. 1947. Hemoglobin and nitrogen fixation in the root nodules of leguminous plants. *Nature.* 159:692-694.
- Kiss, C. B., Z. Végh and E. Vincze. 1987. Nucleotide sequence of a cDNA clone encoding leghemoglobin III (LbIII) from *Medicago sativa*. *Nucl. Acids Res.* 15:3620.
- Konieczny, A. 1987. Nucleotide sequence of lupin leghemoglobin I cDNA. *Nucl. Acids Res.* 15:6742.
- Konieczny, A. E., O. Jensen, K. A. Marcker and A. B. Legocki. 1987. Molecular cloning of lupin leghemoglobin cDNA. *Mol. Biol. Rep.* 12:61-66.
- Kortt, A. A. P., M. Strike, D. Bogusz and C. A. Appleby. 1987. The amino acid sequence of leghemoglobin III from *Sesbania rostrata* stem nodules. *Biochem. Int.* 15:509-516.
- Kortt, A. A. P., S. Inglis, A. I. Fleming and C. A. Appleby. 1988. Amino acid sequence of hemoglobin I from root nodules of the non-leguminous *Cesuaria glauca-Frankia symbiosis*. *FEBS Lett.* 231:341-346.
- Lee, J., S. G. C. Brown and D. P. S. Verma. 1983. Chromosomal arrangement of leghemoglobin genes in soybean. *Nucl. Acids Res.* 11:5541-5553.
- Lehtovaara, P. A., Lappalainen and N. Elfolk. 1980. The amino acid sequence of pea (*Pisum sativum*) leghemoglobin. *Biochim. Biophys. Acta.* 623:98-106.
- Lesk, A. M. and C. Chothia. 1980. How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins. *J. Mol. Biol.* 136:225-270.
- Lewin, R. 1982. Can genes jump between eukaryotic species? *Science.* 217:42-43.
- Lewin, R. 1984. Surprise finding with insect globin genes. *Science.* 226:328.
- Marcker, K. A. and N. N. Sandal. 1987. Evolution of the leghemoglobins. In: *Plant Molecular Biology.* Von Wettstein, D and N. H. Chua (eds.) Plenum Press, NY pp. 503-508.
- Richardson, M. M. J., Dilworth and M. D. Scawen. 1975. The amino acid sequence of leghemoglobin I from root nodules of broad bean (*Vicia faba*). *FEBS Lett.* 51:33-37.
- Tjepkema, J. 1983. Hemoglobins in the nitrogen-fixing root nodules of actinorhizal plants. *Can. J. Bot.* 61:2924-2929.
- Wakabayashi, S. S., Matsuura, and D. A. Webster. 1986. Primary sequence of a dimeric bacterial hemoglobin from *Vitreoscilla*. *Nature.* 322:481-483.
- Welters, P. B., A. Metz, J. Schell and F. J. de Bruijn. 1989. Nucleotide sequence of the *Sesbania rostrata* leghemoglobin (*srgIb3*) gene. *Nucl. Acids Res.* 17:1253.