Sequence-Function Relationships of Prokaryotic and Eukaryotic Galactosyltransferases¹

Christelle Breton,^{*,2} Emmanuel Bettler,^{*} David H. Joziasse,[†] Roberto A. Geremia,^{*} and Anne Imberty^{*}

*Centre de Recherches sur les Macromolécules Végétales,³ CNRS, BP 53, F-38041 Grenoble cedex 9, France; and †Department of Medical Chemistry, Vrije Universiteit, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

Received for publication, December 26, 1997

Galactosyltransferases are enzymes which transfer galactose from UDP-Gal to various acceptors with either retention of the anomeric configuration to form $\alpha 1, 2$ -, $\alpha 1, 3$ -, $\alpha 1, 4$ -, and $\alpha 1, 6$ -linkages, or inversion of the anomeric configuration to form $\beta 1, 3$ -, $\beta 1, 4$ -, and $\beta 1$ -ceramide linkages. During the last few years, several (c)DNA sequences coding for galactosyltransferases became available. We have retrieved these sequences and conducted sequence similarity studies. On the basis of both the nature of the reaction catalyzed and the protein sequence identity, these enzymes can be classified into twelve groups. Using a sensitive graphics method for protein comparison, conserved structural features were found in some of the galactosyltransferase groups, and other classes of glycosyltransferases, resulting in the definition of five families. The lengths and locations of the conserved regions as well as the invariant residues are described for each family. In addition, the DxD motif that may be important for substrate recognition and/or catalysis is demonstrated to occur in all families but one.

Key words: families, galactosyltransferase, glycosyltransferase, peptide motif, sequence analysis.

Glycosylation reactions are of great biological importance for both prokaryotes and eukaryotes (1), and are dependent on a class of enzymes, the glycosyltransferases. These enzymes form a large family that participates in a concerted fashion in the biosynthesis of polysaccharides, and of the carbohydrate moieties of glycoproteins and glycolipids. In eukaryotes, most glycosyltransferases are located in the endoplasmic reticulum and Golgi apparatus. The Golgi enzymes share the same domain structure: they are type II membrane proteins, consisting of a short N-terminal cytoplasmic tail, a transmembrane domain, a stem region of variable length, and a large C-terminal globular catalytic domain (2, 3). In contrast, bacterial glycosyltransferases exhibit various topologies: some of them have several transmembrane domains, whereas others bind to membranes even though no transmembrane domains were predicted (4, 5).

© 1998 by The Japanese Biochemical Society.

In the present study, we focus on the genes that encode galactosyltransferases (GalTs). Galactosyltransferases are involved in the synthesis of many oligosaccharides of biological importance in eukaryotes (N-glycans, O-glycans, histo-blood group antigens, glycolipids, lactose...), and in the synthesis of cell wall polysaccharides in prokaryotes. The reaction catalyzed by these enzymes is the transfer of galactose from UDP- α -D-galactose to an acceptor (sugar or aglycone) with either an α - or β -linkage. Eukaryotic UDP-Gal:GlcNAc β -R β 1 \rightarrow 4-galactosyltransferase (β 4-GalT, EC 2.4.1.38), that catalyzes the transfer of galactose to N-acetylglucosamine (GlcNAc), is one of the most extensively studied glycosyltransferases. It is involved in the biosynthesis of glycoproteins and glycolipid glycans, and in the synthesis of lactose in lactating mammary glands. The existence of a family of homologous β 4-GalTs with related functions is now well established (6-8). Except for that of human β 4-GalT3 (8), the acceptor specificity of these enzymes is modulated by α -lactalbumin (7, 9). The eukaryotic UDP-Gal:Gal β 1-4GlcNAc-R α 1 \rightarrow 3-galactosyltransferases (α 3-GalT, EC 2.4.1.151) constitute another important family. They catalyze a reaction in which galactose is transferred to N-acetyllactosamine to form the non-reducing terminal element, $Gal\alpha 1, 3Gal\beta 1$, 4GlcNAc-R (α Gal epitope), which is responsible for the hyperacute graft rejection in xenotransplantation (10, 11). Except for two short peptide motifs which appear to be shared by the α 3- and β 4-GalTs (12), these enzymes exhibit no significant primary sequence identity. In prokaryotes, galactosyltransferases are involved in the biosyn-

¹This work was supported by the following grants: Programme Physique et Chimie du Vivant-CNRS, Immunology Concerted Action 3026PL950004, and Xenotransplantation Project BIO4CT972242 of the BIOTECH program from the European Union.

² To whom correspondence should be addressed. Tel: +33 4 76 03 76 35, Fax: +33 4 76 54 72 03, E-mail: breton@cermav.cnrs.fr ³ Associated with Université Joseph Fourier.

Abbreviations: GalT, galactosyltransferase; Gal, galactose; Glc, glucose; GlcNAc, *N*-acetylglucosamine; GalNAc, *N*-acetylgalactosamine; GlcA, glucuronic acid; EPS, exopolysaccharide; LPS, lipopolysaccharide; LOS, lipooligosaccharide; HCA, hydrophobic cluster analysis.

thesis of O-antigens, the outer core of lipopolysaccharides, lipooligosaccharides, exopolysaccharides and capsular polysaccharides. The sugar donor is that in eukaryotes, UDP- α -D-galactose, and the acceptors are commonly oligosaccharides coupled to lipids.

It is the aim of the present work to identify a sequence signature that is characteristic of galactosyltransferases. In some cases, it has been demonstrated that enzymes exhibiting the same acceptor specificity share regions of structural homology: similarities have been found between the blood group A and B transferases, which are α 3-GalNAc- and α 3-Gal-transferases, respectively (13), and also between a β 4-GlcNAc-transferase from snail and the mammalian β 4-GalTs (14, 15). On the other hand, using bio-computing methods we were able to find significant similarities between different families of fucosyltransferases of prokaryotic and eukaryotic origin, and we identified a few strictly conserved residues which could be of importance in the catalytic process (16). The same approach was used in the present study in order to extract structural information on the primary sequences of galactosyltransferases and to provide insights into the sequence-function relationships of this class of enzyme.

Definition of groups of galactosyltransferase sequences

A computer-assisted systematic search using the key word "galactosyltransferase" of nucleotide and protein databases GenBank, NBRF-PIR, and Swiss-Prot was performed, which vielded more than 500 entries for eukaryotes and prokaryotes. Further searches of the databases were performed on the basis of sequence similarities using the BLAST program (17). Redundant sequences, partial sequences, ESTs, pseudogenes or 5'- and 3'-UTR regions were searched for and excluded from the study, as were Gal-1-P transferases, which catalyze a rather different reaction. When two sequences were almost identical (>95% sequence identity), only one was taken into account in the sequence analysis. In the first step we only considered galactosyltransferase sequences whose functions have been experimentally determined. At the time of writing, we had identified 35 distinct GalTs, which constituted the background of this study (Table I). These enzymes catalyze the transfer of galactose from UDP- α -D-Gal to various acceptors with either retention of the anomeric configuration to

TABLE I. Summary of galactosyltransferase sequences of known function.

Group	No. of sequences	Origin	Name	Type of linkage
a.Galacte	osyltransfe	rases		
I	2	Yeast	Mnn10/Gma12	$\alpha 1 \rightarrow 2$
II	5	Animal	α 3-GalTs/B transferase	α 1→3
III	1	Bacterial	RfaI	α 1→3
IV	5	Bacterial	RfbF/RfpB	<i>α</i> 1→3
V	2	Bacterial	LgtC	<i>α</i> 1→4
VI	1	Bacterial	LpcA	<i>α</i> 1→6
VII	2	Bacterial	RfaB/AmsD	<i>α</i> 1→6
β-Galacto	syltransfe	rases		
VIII	2	Animal	β3-GalTs	β1→3
IX	7	Animal	β4-GalTs/CKI/CKII	$\beta 1 \rightarrow 4$
Х	4	Bacterial	LgtB/LgtE	$\beta 1 \rightarrow 4$
XI	1	Bacterial	Cps14J	$\beta 1 \rightarrow 4$
XII	3	Animal	CGTs	β1→cer

form $\alpha 1,2$ -, $\alpha 1,3$ -, $\alpha 1,4$ -, and $\alpha 1,6$ -linkages, or inversion of the anomeric configuration to form $\beta 1$ -Cer, $\beta 1,3$ -, and $\beta 1,4$ -linkages. On the basis of both the nature of the catalyzed reaction and the protein sequence identity, these enzymes can be classified into twelve groups (I-XII, Table I). Typically, proteins belonging to the same group exhibit at least 25% sequence identity. This classification based on sequence identity may reflect overall structural similarities in the protein folds.

Retaining galactosyltransferases, *i.e.* those catalyzing the formation of α -linkages, are spread over seven groups. Group I contains two yeast α 2-GalTs which exhibit low overall sequence identity (26%). The enzymes belonging to groups II, III, and IV are all α 3-GalTs. Group II comprises all the highly homologous animal α 3-GalT sequences (75 to 85% identity), which are involved in the biosynthesis of the α Gal epitope (Gal α 1,3Gal β -R), and the human blood group B α 3-GalT, which exhibits about 40% sequence identity with the animal α 3-GalTs. Group III contains only one sequence, a bacterial α 3-GalT (RfaI) involved in the biosynthesis of LPS. The degree of identity between the various RfbF and RfpB sequences that constitute group IV is extremely variable, ranging from 25 to 79%. Group V contains two bacterial α 4-GalTs (LgtC), which are also involved in LPS biosynthesis. Groups VI and VII comprise bacterial α 6-GalTs.

All β -galactosyltransferases (inverting enzymes) can be classified into five groups (VIII-XII). Group VIII contains a sequence corresponding to the recently cloned rat β 3-GalT involved in the biosynthesis of gangliosides and a patented human sequence. All the highly homologous animal β 4-GalTs (50 to 85% identity) involved in the biosynthesis of N-glycans and in lactose synthesis constitute group IX. These enzymes use a GlcNAc residue as an acceptor (or Glc in the presence of α -lactal burnin). The bacterial β 4-GalTs which constitute homogeneous group X (70 to 90% identity) participate in the biosynthesis of LOS. Group XI contains only one bacterial β 4-GalT. Finally, group XII contains the mammalian β -GalTs (CGTs) involved in the biosynthesis of glycosphingolipids. The CGTs are enzymes residing in the endoplasmic reticulum with a C-terminal transmembrane domain in contrast to the other eukaryotic galactosyltransferases, which reside in the Golgi apparatus and are anchored by an N-terminal transmembrane domain.

The twelve groups described here do not exhibit any significant sequence similarity, generally less than 20% identity, with the exception of the LpcA protein of group VI which exhibits low but significant sequence identity (\sim 27%) with the LgtC proteins of group V. One of the most recently cloned sequences of mammalian β 4-galactosyltransferase (18) does not appear to be related to any enzyme studied here and thus was not considered further.

Definition of families

The reaction catalyzed by galactosyltransferases involves two substrates, a nucleotide sugar donor (UDP-Gal) and an acceptor. Since these enzymes utilize the same donor, it was of interest to search for local homology which could account for the nucleotide sugar binding domain. Similarly, enzymes exhibiting the same acceptor specificity could also share regions of structural homology. Because of its known sensitivity as to very low sequence identity, the HCA method (19) was used to compare the protein se-

TABLE II.	Protein sequences	belonging to fa	milies A-E.		
Acc No.ª	Source	Name	Linkage	aa	Mech. ^b
Family A					
104989	bovine	bα3-GalT	Galα1→3Galβ-R	368	R
X01871	human	B transferase	$Gal\alpha 1 \rightarrow 3[Fuc\alpha 1, 2]Gal\beta - R$	338	R
\$71333	marmoset	maa3-GalT	Galα1→3Galβ-R	376	R
M85153	manneset	ma3-GalT	Galα1→3Galβ-R	371	R
126152	nia	ng 3-GalT	Galα1→3Galβ-R	371	R
L30132	ръ				
1166110	dog	Forceman synth	GalNAca1→3GalNAcβ-R	347	R
106175	human	A transferase	GalNAca1 \rightarrow 3[Fuca1.2]GalB-R	354	R
JU5175	numan	A transierase	Gantzieur ofraeuristomy :		
V14669	having	b84-GalT	Gal81→4GlcNAc8-R	402	I
A14330	obiokan	CKI	GalB1→4GlcNAcβ-R	362	I
019890	chicken	CKI	$GaiB1 \rightarrow 4GlcNAcB-R$	373	Ī
U19889	Lumon	B4 ColT	$GalB1 \rightarrow 4GlcNAcB-R$	398	1
X14085	numan	04 CalT2	CalB1 \rightarrow 4GlcNAcB-R	397	ī
¥12509	numan	04-0a112	Call $\rightarrow 4$ GlcNAcB-R	393	ī
¥12510	numan	p4-Gal15	$Galp1 \rightarrow 4GleNAcB-R$	309	ī
103880	mouse	mp4-Gall	Gaip I-40kNAcp-K	577	
X80228	L. stagnalis	β4GnT	GlcNAcβ1→4GlcNAcβ-R	490	I
	<i>a</i> ,	D10511 4	unknown	280	
Z29095	C elegans	RIOETI 4		387	
X98132	C. elegans		unknown		
Family B	-***				
•					
P19816	S. typhimurium	Rfal	Gala1→3Glca-R	337	R
				206	Ď
U14554	N gonorrheae	LgtC	$Gala I \rightarrow 4Gala - R$	211	r D
U65788	N meningitidis	LgtC	Gala I→4Gala-K	511	K
X94963	R leguminosarum	LpcA	Gala1→6Mana-R	288	R
				226	n
X53847	S typhimurium	RfaJ	$Gica1 \rightarrow 2Gala - R$	330	к п
M80599	E. coli	RfaJ	Glca1→2Glca-R	338	ĸ
				220	n
M80599	E colı	Rfal	Gica1→3Gica-R	339	n
S54723	D. melanogaster	DUGT	$Gica I \rightarrow 3Mana - R$	1346	n n
U38417	S. pombe	Gpt1	Glea I→ 3Mana-R	1447	ĸ
	·	DA O		635	
L31762	K. pneumoniae	RIDU	unknown	346	
AF001308	A. thaliana	110M13.14	unknown	330	
U32711	H influenzae	H10258	unknown	550	
Family C					
j					_
EO7739	human	hβ3-GalT	Galβ1→3GlcNAcβ-R	326	I
AB003478	rat	rβ3-GalT	Gal β1→3GalNAcβ-R	371	I
					-
U14554	N. gonorrheae	LgtB	Galβ1→4GlcNAcβ-R	279	1
U14554	N. gonorrheae	LgtE	Galβ1→4Glcβ-R	280	1
U25839	N meningitidis	LgtB	Galβ1→4GlcNAcβ-R	275	I
U25839	N. meningitidis	LgtE	Galβ1→4Glcβ-R	276	I
AC000318	A thahana	T7N9 14	unknown	677	
7.71178	C. elegans	B0024.3	unknown	507	
L35770	D. melanogaster	Fringe	unknown	412	
[11]149	D melanogaster	Brainiac	unknown	325	
Q03974	H. influenzae	Lex1	unknown	302	
Family D					
V01001	C kunilai	R#F	Gala1→3GlcNAc8-R	376	R
721001 1.11410	C. nyoner K province	RfbF	Galα1→3GleNAcβ-R	377	R
121762	K preumoniae	DAF	Galor 1→3GlcNAcβ-R	377	R.I
L31102	K. pneumon. OCK	1/101	Gal/B1→3Galα-R	<u> </u>	
131167	S marcascans	RfbF	Gala1→3GlcNAcB-R	380	R
L3+10/	S. marcescens	RfnR	Gala1→3GlcNAcβ-R	377	R
513323	5. uvsemeriae	афо	Janua Politikapik		

TABLE II.	(continued)				
Acc No.ª	Source	Name	Linkage	aa	Mech. ^b
X77921	E. amylovora	AmsD	Galα1→6Galβ-R	351	R
M80599	E. coli	RfaB	Galα1→6Glcα-R	369	R
P13484	B. subtilis	Tage	Glca1→Glycerol-P	673	R
U58765	N. meningitidis	RfaK	GlcNAca1→2Hep(II)α-R	354	R
Y13970	N. meningitidis	SiaD	(poly)NeuAcα2→8NeuAcα-R	1037	1
Z83335	S. pneumoniae	CaplE	GalAα1→3	378	R
U67549	M jannaschu	prot M	unknown	406	
Y07786	V cholerae	ORF41x3	unknown	377	
Z47767	Y. enterocol. 0:3	TrsD(WbcM)	unknown	358	
U46859	Y enterocol 0:8	WbcI	unknown	358	
Family E					
X85787	S. pneumoniae	Cps14J	Galβ1→4GlcNAcβ-R	318	I
U14554	N gonorrheae	LgtA	GlcNAcβ1→3Galβ-R	348	I
U25839	N meningitidis	LgtA	GlcNAcβ1→3Galβ-R	333	I
X85787	S. pneumoniae	Cps14I	GlcNAcβ1→3Galβ-R	306	I
U14554	N gonorrheae	LgtD	GalNAc β1→3Galβ-R	337	I
P33697	R. mehloti	ExoO	Glc β1→6Glcβ-R	334	I
X77617	E. coli	Kfic	GlcAβ1→4GlcNAcα-R GlcNAcα1→4GlcAβ-R	520	R.I
X85018	human	GalNAcT1	GalNAcα1→Ser/Thr	559	R
^a Accordion N	Jo are for either th	o ConBank or	Swiss Prot databank Mach	machanism of	action · R

^aAccession No. are for either the GenBank or Swiss-Prot databank. ^bMech., mechanism of action; R (retention) and I (inversion) of anomeric configuration.

quences belonging to different groups. The effectiveness of HCA comes from its ability to significantly detect the regular secondary structural elements constituting the hydrophobic cores of globular proteins (20). Combining BLAST and HCA, we searched for local homologies which could be present in some regions of the catalytic domains of galactosyltransferases belonging to different groups as well as in other classes of glycosyltransferases, thus allowing the classification of families. The results of the present study allowed the grouping of the 12 groups of galactosyltransferases into 5 families.

Family A. This family comprises the eukaryotic α 3- and β 4-GalTs of groups II and IX. Other sequences were also included in this family (Table II), such as the human blood group A transferase and canine Forssman synthetase. which are both α 3-GalNAc-transferases. The blood group B and A transferases differ in only four amino acids, which were assumed to be responsible for the differences in nucleotide-sugar specificity (21). These enzymes transfer galactose (or GalNAc) to the galactose residue of the H structure, [Fuc α 1,2]Gal β 1-R, present in glycoproteins and glycolipids. The Forssman synthetase is homologous to the blood group transferases and animal α 3-GalTs (42 and 35% sequence identity, respectively), but exhibits a different acceptor specificity, as it is involved in the biosynthesis of the non-reducing end of the Forssman antigen, GalNAc α 1,3GalNAc β 1,3Gal α 1,4Gal β 1,4Glc-Cer (22). Similarly, a β 4-GlcNAc-transferase from snail (Lymnaea stagnalis), which is thought to be involved in a variant pathway for the synthesis of complex type N-glycans (14) and which is homologous with the eukaryotic β 4-GalTs (27-37% identity), was also included in family A.

The glycosyltransferases belonging to this family catalyze rather different reactions with regard to the nature of the donor (UDP-Gal, UDP-GalNAc, or UDP-GlcNAc) and of the acceptor (Gal, [Fuc $\alpha 1,2$]Gal, GalNAc, GlcNAc, or Glc), as well as to the stereochemistry of the reaction. The presence of two short conserved peptide motifs (DVD and DKKN) in both the eukaryotic α 3- and β 4-GalTs has already been described (12). Except for these two motifs. no homology between these two classes of galactosyltransferases was detected using the classical linear sequence alignment methods. However, with HCA, significant homology was found in the catalytic domains of all enzymes belonging to family A. Multiple sequence alignment with the most representative protein sequences belonging to this family was performed using the HCA results and is shown in Fig. 1. A number of aromatic, basic and acidic residues were found to be highly conserved. One of the best conserved regions corresponds to the DVD motif (region I). Four cysteine residues are conserved in all the β 4-GalTs and in the snail β 4-GlcNAcT (residues 129, 171, 245, and 264 in the human β 4-GalT), but no cysteines were found at equivalent positions in the α 3-GalTs, even though a Cys residue close to the DVD motif, like Cvs245 in the β 4-GalT, is also present in the α 3-GalTs. Two other regions, corresponding to the acidic segment (GGEDDD) characteristic of the β 4-GalTs and to the DKKN motif (regions II and III, respectively), were also found to be conserved in the α 3and β 4-GalTs. Judging from our alignment, only the aspartate residue of the DKKN motif is actually strictly conserved, thus stressing the importance of this amino acid. Several residues in these two regions were suggested to interact with UDP-Gal or to be close to the nucleotide

1004

Family A		_	Ι		
ba3GalT ma3GalT pa3GalT	132: 135: 135:	: TVFAVGEVTEHYLEEFLTSONK-HEMVGHEVTFYIMVDDYSEMPLIELGPLRSFKUFELKEEKENODISMEM : TVFAVGEVTFHYLEDFLESOM-YFMVG - HEVTFYVHIDDTSEMPVHENPIHSLOVPEIRSEKENODISMEM : TVFAVGEVTEHYLEEFLISONI-YFMVG - HEVIEYIMVDDISEMPLIELGPIRSFKUFEIRSEKENODISMEM	ктисен і Vahioh-Б Ктисен Паніон-Б Ктисен Паніон-Б	VDEIÊCMDVDQUEQDKEGŸ VDÊÎÊCMDVDQVFQDNEGV VDFLFCIDVDQVFQNNEGV	:236 :239 :239
maα3GalT A transf Forssman	140: 119: 114:	: TOFAIGMTERTIEFVTS/NR-YFROG IKVTFYOR/DOVKAPEBLIPTESTATEVDEKKRODISEN TVFAIKOV-AFIKIPLETELTE : TVFAVGMT-REVQHTESTEQ-FFROG YOVYY1FINDPACIPROPLGPGRLLSEIPIGRKESISTENEEISTRA : TVFAVGMT-REVQHTESTEQ-FFROG YOVYY1FINDPACIPROPLGPGRLLSEIPIGRKEST	ETISENIAQEAHR - E	VDYLYCVDVDVFRDHVGV VDYLYCVDVDVFRDHVGV VDYLYCVDVDVFRNPWCP	:222
bβ4GalT mβ4GalT β4GalT	164: 161: 159:	: ҚИЛССЕЙТҮРИРОСТЬРНКУЛІІЛІРИКСЕШСКҮМ-УҮЛНРИЧОВ ООДУУС!ҮЦТНОА(ЕЫМЕХА) : PIRTGGANSPKUCVSPHKVAIIIIPKNROPHLKYMLYIHPILOR OODYG!YUTHOA(ШТНРИА) - NYMCETNAPROCVSPHKVAIIIPEKNROPHLKYMLYIHPVIOR OODYG-IYUTYINOAGUTHFINA	K-ILLIV(@KEALK-D K-ILLIV(@VEALK-D K-ILLIV(@VEALK-D	YDYNCEVES-DVDEIPHIDHNT YDYNCEVES-DVDEIPHDHNN YDYTCEVES-DVDEIPHIDHNN	:263 :260 :261
β4GalT2 β4GalT3	152: 108:	CALMAGRYTPPDCTPAQTY VI IPPRHREHELRYWLHYLHP1LRA-ORDRYGVYUINOH -GZYIAYA RYEPKORYRPAGCEPRSRT IVPHRAREHELRYWLHYLHP1LOR-QQUAYGIYUHQA GYGIYAYA	K-DINVGFLEALKED K-DINVGVREALR-D	AAYDCFIFS-DVDLVPHDRAL EEWIXLFLH DVDLEPENDANL	:252 :206
CKI CKII Lsβ4GnT	123: 128: 172:	: Evregozdapkockalokvili i ipyrnrpphikywlyymedior ocidygvywinododłeinija : dysloskyppdclerokvyti.tpernrehillkywlythedira-okyaygywinogGedenija : dyodgomytphyctparktyli i ipyrnrchilytillenlipmimr orydetifyiRottpetfnyg	K LINVCFIFALK- K-LINVCFLFALKDJ I IFNAGYLFALKDJ	YDYDCFVES-DYDLIPHDDRNF EEYDCEIFS-DYDLIPHDDRNL -NYDCFILH-DYDHIPIDRNM * *	:222 :228 :271
ba3GalT ma3GalT	266: 269:	: SAN IPÉGSOPYYHÁLIFOGTPTQVINITQECFKGILK-DŘÍNDIEAQMHDESHLNKYFLINKPTKLISP :3 : San IPÉGBOPYYHÁLIFOGTPTHILNI, TREÇFKGILQ-DŘÍNHDISHLNKYFLENKPTKLISP :3	35 38		
pu3GalT mau3GalT	269: 274:	: Sanipří – ospřívlatifostpřovinitorčekgilo – řkení learwideshlnkyfilnkptki (959:33 : Sanipří – ospřívna ifostpiovinitorcekgili – říčno iraewideshlnkyfilnkpski (959:33	38 43		
A transf Forssman	252: 247:	SQAVIPKDBGJTYYLGGFFGGSVQEVQR5TRACHQAMMV-BQARGIRAVHHDESHLAKYLLRKPTKVHSP :3 : STAPVAENBGJTYYGGAVFGGRVAKVYEFTTGCHMAILA-DKANGIMAAWQEESHLNRRFISHKPSKVHSP :3	16		
bβ4GalT mβ4GalT	308: 305:	: PNNYNGNOGEDD I I'NRIAFRANSYSRPNAVIGKCRMIRHSRNKKNEPNPORFORIAHTKETMLSDGLNSITT : 3 : PNNYNGNOGEDD I ENRLYHKGNS I SRPNAVVGRCRMIRHSRNKKNEPNPORFORIAHTKETMRFOGLNSITT : 3	80 77		
β 4GalT β4GalT2	306: 297:	: PNNYWGYÖGDDD I FNRÖVFRÖMS ÖSRPNAVVGRCRHIRHSRÖKKŃEPNPORFDRIAHTKETMLSDGLNS II :3 : PNNWGYÖGDDD I FNRÖSLITMKI SRPDIRIGRYRMIKHDRÖKHÖEPNPORFTKIONTKLIMKRDGIGSVAY :3	78 69		
β4GalT3	252:	: PNEYWGWGGEDDDIATRYALACHKYSPPPTSVGHYKWYKHRGYKGNEENPHRFÛLLVRTONSWTODGANSUTY : 3 pnykgruffaren ywritwyr Mgi far pnau igkrem i risrnêknie nyrren i rahtretmssdginsiysy : 3	24 39		
CKII Lsβ4GnT	273: 317:	PNEYWGWGGEDDDIFNRISINGWKVSRPDIRMGRYRMIKHERWKHNEPNPORFTKIONTKMMKRDGISSUOY :3 SNTYFGGGEDDDLRWRÁVHM	45 337		
· · · · · · · · · · · · · · · · · · ·		- ···			
Family B			<u>II</u> ►	← IV →	
LgtC N.gon LgtC N.men	95: 96:	: (1)KVL51DT7VL7RDGFRFFMDTDLGGNWVCAC1D <20> VENASVD150NEKKWR <21> 0D0D1DN-GEFRGG : CDKVL51D1D7VL7NLSBTFFMDTDLGDNWLGAC1D <20> VENASVD157NEKKWR <21> 000D1DN-GEFRGG	VCYANSRENFMPTN VCYANSRENFMPTN	<28> VSHYCESAKEWHRDC : <28> VSHYCESAKEWHRDC :	255 256
REAL E.COLL	124:	: APRVIYITAD INCOGITE PUTINI SEPUDIKV MVVI <22> MENSEFILINITAQUA <23> POODUBI-MISLADK	TRUDIKATORS	<17> FURMICPTREAMHDWA :	211
Rfal S.typh	122.	: OTKVINTIDAD TACKGSTOED IDLNFALMED AVVA <22> MENA FETTINT XIPLAT <23> IDODVEN- IFLVNK I DEL MILA DANCKGDUSOU ULCIN-CAMBANAY <21> MENSOUNDED YMD <21> DODVEN- HI VOM	AREVOKKENTOFSL	<17> FÜHYIGPTKPWHSWG : <21> I ÜHYIGPTKPWHSWG :	276
Rfal S.typh RfaJ E.coli RfaJ S.typh BfbC k.pneu	123: 123: 120: 379:	: OTKVLYLDATIACKGSTOPPILLLIFALNEJAAVVA <22> YENASPITETXIPLAT <23> LOODVEN- LFLVAK : LARLAYLDA VYCKGUISOLILLGLA-GAVAVVK <21> YENSSYYTTÖLKKAA <21> POODVEN-VILLAR : VN LLAYLDA VYCKGISIADI IQLDL.T.FKLAAVVK <21> YENSSYYTTÖLKKAK <22> POODVEN- LELODK : YÜNYVETUSTIVVKADIKHULUVPLINNI VAAVC) <37> YÜNASPITISTIVVKAVK <22> LOODLAV- LEVYSK	ARAVDKKOŇTOFSL TLEUPREYNTIYTI VIFIPREYNTIYTI VTŘIPLETNYHCN	<17> FÜHYIGPTKENHSWG : <21> LÏHYIGATKENHKWA : <25> DIHYRGATKENHAWA : <26> MIHYAGENKENTEK :	276 276 274 552
Rfal S.typh RfaJ E.coli RfaJ S.typh RfbC k.pneu A. thaliana LpcA R.leg	123: 123: 120: 379: 165: 85:	: OTKVINIDATICKGSTORIULIIPALNEIJAVVA <22> YENASTITT <23> LOODYN- FFLUXA : LDRILNIDATYCKGDISOLIILGIN-GAVAAVVA <21> YENSSYYYDDIKKWA <21> PODYM-VILKGM : WILLATDATYCKGSIADITGIDIT. EKIJAVVK <21> YENSSYYTDIKKWA <21> PODYM-VILKGM : YÜKVYFIDSTWYKADIGELUVPLINNIYAVC) <37> YENSSYTTYNKKWA <21> PODYM-VILKGM : VINYFIDSTWYKADIGELUVPLINNIYAVC) <37> YENSSYTTYNKKWA <19> LOODIN-KVFYSR : VINYFIDSTWYKADIGELUVPLINNIYAVC) <37> YENSSYKTYNKKWA <19> LOODIN-KVFYSR : YENYFIDSTWYKADIGELUVPLINNIYAVC) <37> YENSSYKTYNKKWA <21> NOODIN-VFTCGV	ARŞVDKKENTOFSL TLEUPREYNTIYTI VIFUPREYNTIYTI VTFUPLENNVYHGN WLVIGPRZNT-OTG	<17> FVHY IGPT KPMHSWG : <21> LIHYT GAT KPMHSWG : <25> DIHYT GAT KPMHSWA : <25> DIHYT GAT KPMHSWA : <26> MIHYA GEM KPMTEK : <67> LIHWS GSC KWIRLD : <9> LIHPT GEM KWOATY :	276 276 274 552 325 228
Rfal S.typh RfaJ E.colı RfaJ S.typh RfbC k.pneu A. thalıana LpcA R.leg DUGT D.mel : Gpt1 S.pomb :	123: 120: 379: 165: 85: 1335: 1252:	: OTKVINIDATICKGSTOPIIDLIPAENEIJAVVA <22> MENAFTIGIKTPIAT <23> LOODURA-UFLKKK : LDRILNIDATYCKGDISOLIHLGLN-GAVIAVVK <21> TENSOUVIDLKKKA <21> PODYA-VILKKM : WILLTIDATYCKKGSTADIGIDLI-EKIAVVK <21> TENSOUVIDLKKKA <22> PODYA-VILKKM : YDKVFILSTVYKADIGEDLIVPIKINIÚNAVKS <37> TENGOVEDLKKKA <22> PODYA-VILKKM : VIKVFILSTVYKADIGEDLIVPIKINIÚNAVKS <37> TENGOVEDLKKKA <21> PODYA-VILKKM : TENGVIKUSSTVYKADIGEDLIVPIKINIÚNAVKS <37> TENGOVEDLKKKA <21> PODYA-VILKKM : TENGVIKUSSTVYKADIGEDLIVPIKINIÚNAVKS <37> TENGOVEDLKKKA : TENGVIKUSSTVYKADIGEDLIVPIKINIÚNAVKS <37> TENGOVEDLKKAR : TENGVIKUSSTVYKADIGEDLIVPIKINIÚNAVKS <37> TENGÓVEDLKKAR : TENGVIKUSSTVYKADIGEDLIVPIKINIÚNAVKS : TENGVIKUSSTVYKADIGEDLKGAPYYTES : VKKI INVOLALNÍTDIKSTVYDNDLGGAPYYTES : VKKI INVOLALNÍTDIKSTVYDNDLGGAPYYTES : TENGVIKADIGET : TENGVIKADIGEDUJUNELKGAPYSYTEM <26> MISATYVYDĽKKEK <22> LOODUSNAH	AREVOKKONTOFSL TLEIPERENTITTI VIEIPERENTITTI VIEIPERENTITTI VIEIPERENTOTI VIEIPERENT-QTG IKSIPPERENT-QTG IKSIPPERENT-QTG IKSIPPERENT	<pre><17> FÜHYIGPTKURHSWG : <21> LIHYTGTTKURHKWA : <25> DIHYRGTKURHAVA : <26> MIHYAGENKUNTEK : <67> LIHYSGEKKUNTEK : <67> LIHYSGEKKUNTEK : : : : : : : : : : : : : : : : : : :</pre>	276 276 274 552 325 228 462 L378
Rfal S.typh RfaJ E.coli RfaJ S.typh RfbC k.pneu A. thaliana LpcA R.leg DUGT D.mel : Gpt1 S.pomb :	123: 120: 379: 165: 85: 1335: 1252:	: OTKVINIDATICKGSTOREIDLITALNELLAVVA <22> TENSETTETKTPIKT <23> IDODTRA-TELIAV IDRUTTALAVYCKGSTOREISOLILLAUTALAVGAVK <21> TENSETTETKTELKOWA <21> PODDYS-TELIAGA VN RIATIDATVCKGSUADITQIDI.T-EKINAVVK <21> TENSETTETKTVKGW <21> PODDYS-TELIAGA VN RIATUDATVCKGSUADITQIDI.T-EKINAVVK <21> TENSETTETKTVKGW <21> PODDYS-TELIAGA VN RIATUDATVCKGSUADITQIDI.T-EKINAVVK <21> TENSETTETKTVKGW <21> PODDYS-TELIAGA VN RIATUDATVCKGSUADITCTNFOGKALAVDD <23> VENSETTETTETTETKTVKGW <21> NODDANT-VTEDGD VN RIATUDATVCKGSUADITCTNFOGKALAVDD <23> VENSETTETTETTETTETTETTETTETTETTETTETTETTETT	Arever and a second sec	 <17> FÜHY ICPT KINHSWG : <21> LIHYTCATKUCHKWAA : <25> EIHYRCATKUCHKWAA : <26> MIHYRCATKUCHKWAA : <36> LIHHSCENKUNTEK : <67> LIHHSCENKUNTEK : :11 	276 274 552 325 228 1462 1378
RfaJ E. coll RfaJ S. coll RfaJ S. typh RfbC k. pneu A. thallana LpCA R. leg DUGT D. mel Gptl S. pomb	123: 120: 379: 165: 85: 1335: 1252:	 OTKVINTLADICKGSIONEJILLIIPALNEIJAVVA <22> MENGETIELKENTY <23> LOODUNA- LELLANAVA IDRULATUATVCKGDUSOLIILGUA-GAVAVAVK <21> MENGETIELKENA <21> POODUNA- ULIKGM VINTUATVAVKKGUSOLIILGUA-GAVAVAVK <21> MENGEVUTVNIKLIKK <22> POODUNA- LILOK VINTUATVAVKKGUSOLIILIVULINITÜAVVK <21> MENGEVUTVNIKLIKK <22> POODUNA- LILOK VINTUATVAVKKGUSOLIIVULINITÜAVVK <21> MENGEVUTVNIKLIKK <22> POODUNA- LILOK VINTUATVAVKKGUSOLIIVULINITÜAVVC <37> MENGEVUTVNIKLIKK <22> POODUNA- LILOK VINTUATVAVKKGUSOLIIVULINITÜAVVC <37> MENGEVUTVNIKLIKK <22> DOODUNA- VIETGUSOLIIVUKONK VINTUATVAVKKGUSOLIIVULINITÜAVCO <37> MENGEVUTVNIKUKKKK VINTUATVAVKKGUSOLIIVULINITÜAVCO <37> MENGEVUTVNIKUKKKK VINTUATVAVKKGUSOLIIVULINITÜAVCO <37> MENGEVUTVNIKUKKKK VINTUATVAVKKGUSOLIIVULINITÜAVCO <37> MENGEVUTVNIKUKKKK VINTUATVAVKKGUSOLIIVULINITÜAVCO <37> MENGEVUTVINIKUKKKK VINTUATVAVKKGUSOLIIVULINITUKSIYVDULIGAPK VINKITVVAALINITUKSIYVDULGAPYYTTPE <26> MIISMEVVDULGAPK VINKITVVAALINITUKSIYVDULGAPYYTTPE <26> MIISMEVVDULGAPK VINKITVVAALINITUKSIYVDULGAPYYTTPE <26> MIISMEVVDULGAPK VINKITVVAALINITUKSIYDULGAPYYTTPE <26> MIISMEVVDULGAPKKKK VINKITVVAALINITUKSIYDULGAPYYTTPE <26> MIISMEVVDULGAPKKKK VINKITVVAALINITUKSIYDULGAPYYTTPE <26> MIISMEVVDULGAPKK VINKITVVAALINITUKSIYDULGAPYYTTPE <26> MIISMEVVDULGAPKK 	ARE DENG NUCESI TIL DERE NITITI VIEUPELE NITITI VIEUPELE NITITI VIEUPELE NITHGN MUNIGPRÄNT-OTG IKSIPPD LLWCOTW TYSIPPD LLWCETW	<pre><pre><pre>> FUH TOP TOP TOP HIM SWG :</pre></pre><pre><pre><pre>> ID HUTCAT TOP HIMA :</pre><pre><pre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre><lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></lpre></pre></pre></pre><pre></pre></pre><pre></pre></pre>	276 274 552 325 228 462 L378
Rfal E. coll RfaJ S. cyph RfbC k. pneu A. thallana LpcA R. leg DUGT D. mel Gpt1 S. pomb Family C rβ3GalT hβ3GalT	123: 120: 379: 165: 85: 1335: 1252: 77: 84:	: OTKVINIDATICKGS[OPE]IDLIPAENE]AVVA <22> ЙЕМАРТІЙХЙРІНТ <23> LOODINA- LFLUNK : LDRIJNLA VYCKGDISOLIILGIN-GAVAAVVK <21> ЙЕМАРТІЙХЙРІКТ <23> LOODINA- LFLUNK : VNILLATIDATYCKGDISOLIILGIN-GAVAAVVK <21> ЙЕМАРТІЙХЙРІКЦККА <21> РООDVA- VILKEM : VNILLATIDATYCKGDISOLIILGIN-GAVAAVVK <21> ЙЕМАРТІЙХЙРІКЦККА <22> РООDVA- VILKEM : YÜKVETIS TÜVYCADIGELLIVPLINILÜNK <2015 ЙЕМАРТІЙХЙКОК	ARE DENG NT OFFIL TLE DER NT IYTI VIE DER NT IYTI VIE DELE NVYHGN MUNIGER NT-OTG IKS DED NLWOTW IYS DED NLWOTW IYS DED NLWOETW III KV ASRAP KT SLHTR	<17> FÜH MEPRENKINSWE :: <21> LIHI MEATURANKA :: <25> JÜH MEATURANKA <26> HÜH MERNKANTEK : <67> LÜH MERNKANTEK :: <9> LÜH MERNKANTEK :: :1 	276 276 274 552 325 228 L462 L378
RfaJ E. coll RfaJ S. coll RfaJ S. typh RfbC k. pneu A. thallana LpCA R. leg DUGT D. mel : Gptl S. pomb : Family C rβ3GalT hβ3GalT Brainiac Fringe	123: 120: 379: 165: 85: 1335: 1252: 77: 84: 84: 155:	 OTKVINTLADY CKGSIGOEJ IDLIJPALNEJ JAVVA <22> ŽENA P I E IZKEPI T <23> LODDINA - IPLANA I DRIJATLADY CKGSIGOEJ IDLIJPALNEJ JAVVA <22> ŽENA P I E IZKEPI T <23> LODDINA - IPLANA V KIMATLADY CKGSIGOEJ IDLIJI - EKIJAVVK <21> JINS CVTVINIKLIKK <22> PODDINA - IRLONK 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I STUVNICH <22> PODDINA - IRLONK 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I STUVNICH <22> DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI GELIJVPLINIK JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI STUVKADI JAVC> 'YDVYTI STUVKADI JAVC> <37> ŽENA P I DODINA - VIPLGE 'YDVYTI STUVKADI JAVC> 'YDVYTI STUVKADI JAVC> 'YDVYTI	ARE DENG NT OFFIL TILD PERS NT IYTI VIE DER NT IYTI VIE DELE NVYHGN MUVI GERNNT-OTG IKS PEDDILWCOTW IYS PODDLWCETW III KVSVHLP VIACCEPT CICK	<17> FÜH MEPRENKINNE: <21> LIHM GATERHKNA: <25> JÜHMEATERHKNA: <26> ÜHMEATERHKANA: <67> LIHMSESCHWIRLD: <9> ILHTERKKNOATV: 11 	276 276 274 552 325 228 L462 L378
	123: 120: 379: 165: 85: 1335: 1252: 77: 84: 84: 84: 84: 5: 5:	International and the set of t	ARE DENG NT OFFIL TIL DENG NT IVTI VIEDERE NT IVTI VIEDERE NT IVTI VIEDERE NVYHEN WILVIGERINT-OTG IKSTPODILWCOTW IYSTPODILWCOTW IYSTPODILWCETW III KVSIHTR AASVHIR AASVHIR AASVHIR AASVHIR	<17> FÜH MEPRENKINKA : <21> LIHI MEATURAKAA <25> JIH MEATURAKAA <26> MIH MEATURAKAAA <26> MIH MEATURAKAAA <37> IIH MEATURAKAAA <39> IIH MERKANAATU : : : : : : : : : : : : : : : : : : :	276 276 274 552 325 228 L462 L378
RfaJ E. coll RfaJ S. coll RfaJ S. typh RfbC k. pneu A. thallana LpCA R. leg DUGT D. mel : Gptl S. pomb : Family C rβ3GalT bβ3GalT bβ3GalT Branniac Fringe LgtB N. men LgtB N. gon LgtB N. fon LgtB H. inf A. thallana	123: 120: 379: 165: 85: 1252: 77: 84: 84: 155: 5: 8: 435:	OTKUTULATI CKGSTOREJ DLJI JPALNEJ JAVVA <22> DENGET BERKEPINT <23> LODDURA- LFLUNK LIDRUJALDATVCKGSTADI SOLILICULA-GAVAVVK <21> TINS GVITUDLKORA <21> PODDUSA- ULIKOM VNIMITLATVCKGSTADI TQLDI.T. EKIJAVVK <21> TINS GVITUNKLIKK <22> PODDUSA- HILOD VNIMITLATVCKGSTADI TQLDI.T. EKIJAVVK <21> TINS GVITUNKLIKK <22> LODDI VI-KVTTSK TINS GVITUS SLVYDDI TAVKKTYDEDLGGATY YTEF <26> TINS GVITUS SACR <21> NDODAN-VTFDGD VNIKI TVD VALUT TUKSTYDEDLGGATY YTEF <26> TILKTODD VXVDELDKERK <22> LODDI VI-NNHHOVA LIKKYTYDAA TUKKTYDEDLGGATY YTEF <26> TILKTODD VXVDELDKERK <22> LODDI VI-NNHHOVA I HIS TYTYD VALUT TUKSTYDEDLGGATY YTEF <26> TILKTODD VXVDELVSED <71> SGTG VIIS ISVOLILL SUTTKEETABAAHEETYGEDENNFKSI <61> VYMKTOSDEDVNDNI TYTHI <44> SGTG TESADATE VTTE ASAAFERALIAATTS OLARDOVI <44> SGTG TESADATE VIIS TINSAAFERALIAATTS OLARDOVI <44> SGTG TESADATE VIIS TINSAAFERALIAATTS OLARDOVI <44> SGTG TESADATE VIIS TAVAATTS OLARDOVI <45> METVEDDTT SAAFT X55 VIIS TUKSTYDEDLATI SSGTG TISS VS5 TAVEDDTT VERDET SSGTG SSS VS5 VIIS TAVAATTS OLARDOVI VKSS <50> VIIS TUKSTODT VKVDAVIGEAE <41> VIIS VIIS VXVS VKSS <50> VIIS TAVEDT VKVDAVIGEAE <41> VIIS VIIS VXVS VKSS VKSS × 11 TV VIIS TAVEST VKSS VOOKLVKSS <50> VVIKCODT TVKVDAVIGEAE <41> VIIS VIIS VKSS VKSS × 11 TV VIIS VIIS TUKST VKSS VSSS × 500 VKVKSS <500 VVIKCODT TVKVDAVIGEAE <41> VIIS VKSS VKSS × 11 TVKSS × 11 TVKSSS × 11 TVKSS × 11 TVKSSS × 11 TVKSS × 11 TVKSSS × 11 TVKSSS	ARE DENG NT OFFIL TLE DERE NT IVTI VIEDERE NT IVTI VIEDERE NT IVTI VIEDERE NT IVTI VIEDERE NT OFFI INTERNATION INTERNATION III KV	<17> FÜH MEPRENKING: <21> LINN MEATERNIKA <25> JÜH REATERNIKA <26> JÜH REATERNIKA <26> JÜH REATERNIKA <26> JÜH REATERNIKA <3> ILHTERNIKA <3> ILHTERNIKA <3> ILHTERNIKA <3> ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <3 ILHTERNIKA <	276 276 274 552 325 228 462 L378
RfaJ E. coll RfaJ S. coll RfaJ S. coll RfaJ S. typh RfbC k. pneu A. thallana LpCA R. leg DUGT D. mel Gpt1 S. pomb Family C rβ3GalT bβ3GalT Brainiac Fringe LgtB N. men LgtE N. gon Lext H. inf A. thaliana	123: 120: 379: 165: 85: 1335: 1252: 77: 84: 84: 155: 5: 8: 435:	OTKINTLAUT CKGSTOPIIDLIFALNEI JAVVA <22> DENGET ISTRELIT <23> LODDINA- IFLANK LIDRUMUDATYCKGSTOPIISOLHILGUM-GAVAVVK <21> INSS GVTYDIKLAK <22> POODUS- IFLORK VN HANTDA VYCKGSTADITQUDI.T. FKIJAVVK <21> INSS GVTYDIKLAK <22> POODUS- IFLORK VN HANTDA VYCKGSTADITQUDI.T. FKIJAVVK <21> INSS GVTYDIKLAK <22> POODUS- IFLORK VN HANTDA VYCKGSTADITQUDI.T. FKIJAVVK <21> INSS GVTYDIKLAK <22> POODUS- IFLORK VN HANTDA VYCKGSTADITQUDI.T. FKIJAVVK <21> INSS GVTYDIKLAK <22> POODUS- IFLORK VN HANTDA VYCKGSTADITQUDI.T. FKIJAVVK <21> INSS GVTYDIKLAK <22> POODUS- IFLORK VN HANTDA VYCKGSTADITQUDI.KGAVY YTTPY <26> INST GVTIGTAVKGVY <19> LODDIN-VTFDGD VNKITTYD ALAUTTDIKGTYDIKDLGAAPY YTTPY <26> INSS GVTIGTAVKGVY <19> LODDIN-VTFDGD VNKITTYD ALAUTTDIKGTYDIKDLGAAPY YTTPY <26> INSS GVTIGTAVKGVY <19> LODDIN-VNFHQVA INSS GVTIGTAVST ISSUE STATISTICS (STATISTICS) VNKITTYD ALAUTTDIKGTYDIKDLGAAPY YTTPY <26> INSS GVTIGTAVKGVY <19> LODDIN-VNFHQVA INSS GVTIGTAVST ISSUE STATISTICS (STATISTICS) VNKITTYD ALAUTTDIKGTYDIKDLGAAPY YTTPY <26> INSS GVTIGTAVKTG <44> SGTG IFSADAVELIT INSS GVTIGTAVKTG <45 SGTG IFSADAVELTY VTGAANATKANINA STATISTICS (STATISTICS) VNKITTYD ALAUTTTDIKGTYGFFSTD <59> VTLYDDDYT SANAVLKTTG <44> SGTG IFSADAVELIT VIGTAANATKANING STATISTICS (STATISTICS) VIETADAVTTDIALUTTTDIKGT GSTG VTGO NUT VIGTAANATKANINGTICSRG STOF <52> VETVFEDDTUKGEZEZ- KFI <50> A SANATKANING STOKANING STOKANIN VIGTAANATKANINGTICSRG STOF <52> VETVFEDDTUKGEZEZ- KFI <50> A SANATKANING STOKANING STOKANING STOKANING STOKANING STOKANING STOKANING STOKANING STOKANING STOKANING VIGTAANATKANINGTICSRG STOF <52> VETVFEDDTUKGEZEZ- KFI <50> A SANATKANING STOKANING	ARE DENG NT OFFIL TILD DERES AT IVTI VIEDPELS AVTHAT VIEDPELS AVTHAN MUVIGPENAT-OTG IKS PODDLWCOTW III KV	<17> FÜH TEPRETARING: <21> LINY TEPRETARING <25> JÜH TREATSCHIKNA: <26> JÜH TREATSCHIKNA: <26> JÜH TREATSCHIKNA: <3> JÜH TREATSCHIKNA: <3> JÜH TREATSCHIKTER: 	276 276 274 352 325 228 228 237 8 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
RfaJ E. coll RfaJ S. coll RfaJ S. typh RfbC k. pneu A. thaliana LpcA R. leg DUGT D.mel : Gptl S. pomb : FB3GalT hβ3GalT Brainiac Fringe LgtB N.men RthF K.pn01	123: 123: 120: 379: 85: 1335: 1252: 77: 84: 84: 84: 84: 85: 5: 5: 5: 8: 8: 8: 125: 123: 125: 123: 125: 125: 125: 125: 125: 125: 125: 125	I OTKINTLA TI CKGSTOPI IDLITAENEI JAVVA (22> MENGET ISTKIPLIK (23> LOODUNA- LELINK IDRIJATIAA VCKGDI SOLIHIGIA-GAVAAVAK (21> MENSEVYTYJNIKUKK (23> DOODUNA- VILKEM VNIMTUAAVCKGDI SOLIHIGIA-GAVAAVAK (21> MINSEVYTYJNIKUKK (22> POODUNA- HILOD VNIMTUAAVCKGDI SOLIHIGIAAVAKISLISKII JAPRY (30> MINTENNIKOK (21> MOODUA-VTEDGI VNIMTUAAUVCKGDI SOLIHIGIAAVAKISLISKII JAPRY (30> MINTENNIKOK (21> MOODUA-VTEDGI VNIMTUAAUVCKGDI SOLIHIGIAAVAKISLISKII JAPRY (30> MINTENNIKOK (21> MOODUA-VTEDGI VNIKI HVDAALIATIKIN VIDAAUATIKISHIJAAPY (YTEM (26> MINSEVYJDIIDRER (21> MOODUA-VTEDGI VNIKI HVDAA-HVADOVIDIOHIJAHAAVA (21> MILKTODIVIVAVELVSEIJ (71> SGTEVIJE ISNOLILI ISITHKEETABOARBEINGEDENNEKSI (51> MYNKTOSSEEYNNINII IYUL KTIKNINITUALIINT FOLANDOVIDIOHIJAHAAVA (55> SYLEVODIVIVAVELVSEIJ (71> SGTEVIJE ISNOLILI KTIKNINITUALIINT FOLANDOVI (55> SYLEVODIVIVAVELVSEIJ (71> SGTEVIJE ISNOLILI KTIKNINTUALIINT FOLANDOVI (55> SYLEVODIVIVAVELVSEIJ (71> SGTEVIJE ISNOLILI VISLASAATERAIRAMTOSSGIPPOF (51> MENEBOVILEGEGAEOFL (50> (71+1)) VISLASAATERAIRAMTOSSGIPPOF (51> MENEBOVILEGEGAEOFL (50> (71+1)) (71+1) VISLASAATERAIRAMTOSSKIKISFSF (72> MINTEDIVILEGEGAE-OFL (50> (71+1)) (71+1) VISLASAATERAIRAMTOSSKIKISFSF (72> MINTEDIVILEGEGAE-OFL (50> (71+1)) (71+1) VISLASAATERAIRAMTOSSKIKISFSF (72> MINTEDIVILEGEGAE-OFL (50> (71+1)) (71+1)	ARE UDKIGNTOPSI TLE DPRENTITI TLE DPRENTITI VIE DPRENTITI VIE DPRENTITI VIE DPRENTITI VIE DPRENTITI VIE DPRENTITI VIE DPRENTITIE VIE DPRENTIE	<17> FÜH TEPRENKISWE : <21> LIHT TEATURANSWE : <25> JÜH TEATURANSWE : <25> JÜH TEATURANSWE : <26> LIHT ACTIVE HAVAA : <26> LIHT ACTIVE HAVAA : <26> LIHT ACTIVE : <3> LIHT CRASS (NITEK : : : : : : : : : : : : : : : : : : :	276 276 552 325 228 462 3378 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
RfaJ E. coll RfaJ S. coll RfaJ S. coll RfaJ S. typh RfbC K.pneu LpcA R.leg DUGT D.mel Gpt1 S.pomb Family C rβ3GalT hβ3GalT Branniac Fringe LgtB N.men LgtE N.gon Lext H.inf A.thaliana	123: 123: 120: 379: 85: 165: 85: 1335: 1252: 77: 84: 84: 84: 84: 155: 5: 5: 5: 8: 435: 203: 203: 203:	I OTKINTLA TI CKGSTOPI IDLITAENEI JAVVA (22> DENG FIETKEPLIK (23> LOODUNA- LEPLANK, UNKON (21> LORUTA), CKGSTOPI SOLUTIGLA-GAVAVK (21> LINS GVTYDELKOKA (21> POODUSA- HELONK VNIMTLAAVYCKGSTADITQLDI.T. EKITAVVK (21> INS GVTYDELKOKA (21> POODUSA- HELONK VNIMTLAAVYCKGSTADITQLDI.T. EKITAVVK (21> INS GVTYDELKOKA (21> POODUSA- HELONK VNIMTLAAVYCKGSTADITQLDI.T. EKITAVVK (21> INS GVTYDELKOKA (21> LOODIA- KVTYSK VNIMTLAAVYCKGSTADITQLDI.T. EKITAVVK (21> INS GVTYDELKOKA (21> LOODIA- KVTYSK VNIMTLAAVYCKGSTADITQLDI.T. EKITAVVK (21> INS GVTYDELKOKA (21> LOODIA- KVTYSK VNIMTLAAVYCKGSTADITQLDI.T. EKITAVVK (21> INS GVTYDELKOKA (21> NOODAN-VTEGED VNIMTLAAVYCKGSTADITQLDIAKJKKISISRIJ JAVAD (23> INS GVTIDEDWSACR (21> NOODAN-VTEGED VNIKTITVD AAVITTDIKGTYDEDLGAAPY YTTE* (25> HHISHTYDDEKKOR ILIKTIVD AAVITTDIKGTYDEDLGAAPY YTTE* (25> HHISHTYDDEKKOR IIISAVKASREEALRENTVGEDENNFKGI (61> VYNKTGSDEEVNANNIITYKLI (44> SETG IESAVKILOHLIP INS GVTATITYD AAVITTDIKGTYDEDLGAAPY (YTFM (26> HISADITYDIKKTG (46-SA- BATGAVKILOHLIP) VISTASAATTRAAHADITGRAATTSTOV (59> NTLEVDDDYTYSAKAVKKTG (46-SA- BATGAVKILOHLIP) VISTASAATTRAAHADITGRAATTSTOV (55> NTLEVDDDYTYSAKAVKKTG (46-SA- BATGAVKILOHLIP) VISTASAATTRAAHADITGRAATTSTOV (55> VILVEDDYTIEGGEZE-KEI (50> ()) VISTASAATTRAAHADITGRAATTSTOV (55> VILVEDDYTIEGGEZE-KEI (50> ()) VISTASAATTRAAHADITGSKGISTOS (55> VILVEDDYTIEGGEZE-KEI (50> ()) VISTASAATTRAAHADITGSKGISTOS (50> VINCODDITIVKVDAVIGEAE (41> ()) VISTASAATTRAAHADITGSKGISTSF (72> MEKITEDDYTIEGGEZE-KEI (50> ()) VISTASAATTRAA	ARE DENG NT OFFIL ARE DENG NT IVTI UIE DERENT IVTI VIE DELENVIHEN WLVIGPENNT-OTG IKS PODDLWCOTW III KVSVILD NA	<17> FÜH TEPRENASWE : <21> LINN TEATER HAWA : <25> JÜH REATER HAWA : <26> LINN SESSING HAWA : <67> LINN SESSING HAWA : <9> ILH TERRAK OATV : 	276 276 552 325 228 462 378
RfaJ E. coll RfaJ S. coll RfaJ S. coll RfbC k.pneu A. thaliana LpcA R.leg DUGT D.mel Gptl S.pomb FB3GalT hβ3GalT Brainiac Fringe LgtB N.men LgtB N.gon Lext H.inf A. thaliana	123: 120: 379: 379: 85: 85: 1252: 12	International Construction of the second	ARE DENG NT OFFIL ARE DENG NT IVTI VIEPPRENT IVTI VIEPPRENT IVTI VIEPPLENVYHEN MLVIGPRINT-QTG IKS PDDDILWCOFW IYS PODDLWCOFW IYS PODDLWCOFW IYS PODDLWCETW NA	<17> FÜH 10-PTGENASWG: <21> LINY 10-PTGENASWG: <25> JÜH PRCATCH HAWA: <26> KINY ACTACH HAWA: <26> KINY ACTACH HAWA: <26> KINY ACTACH HAWA: <26> KINY ACTACH HAWA: <27. INFO CS GIVER ID: <28. INFO CS GIVER ID: <29. ILHENGEN KIN OATV: YILP DUFY GVSARRYGE: <29. ILHENGEN KIN OATV: YILP DUFY GVSARRYGE: YILP DUFY GVSARRYGE: YILP DUFY GVSARRYGE: YILP DUFY GVSARRYGE: YILP DUFY GVSARRYGE: <	276 276 552 325 228 462 378 5 5 5 5 5 5 5 5 5 5
Rfal E. coll RfaJ S. typh RfaJ S. typh RfbC k. pneu LpCA R. leg DUGT D. mel Gptl S. pomb Family C rβ3GalT hβ3GalT Branniac Fringe LgtB N.men LgtB N.gon Lexl H.inf A. thaliana	123: 120: 379: 165: 85: 1252:	IDRALATION CREENESS LIDLID ALEMENT AVVA 22> DEMA SPECIFIC STATEMENT (23> IDDOUGA- IFLUMAN IDRALATION VCREENESS LIDLIGUN-GAVANVK 21> VENS SVYTPULKOMA 22> POODVSA- VILLOGA VINTATION VCREENESS LIDVELIDUT. END SCHTTPULKOMA 22> POODVSA- VILLOGA 22> POODVSA- VILLOGA VINTATION VCREENESS LIDVELIDUT. END SVYTPULKOMA 22> POODVSA- VILLOGA 22> POODVSA- VILLOGA VINTATION VCREENESS LIDVELAND STEVENESS VITVELESS LIDVELAND VILLON	ARE UNKNOWNESI ARE UNKNOWNESI TILD PRENTITYI VIEUPELE NVIHEN MUVIGPRANT-QTG IKS PODULWCOTW III KV	<17> FÜH TEPRENKISWE : <21> LIHI TEPRENKISWE : <25> JÜH REATERHIKVA : <25> JÜH REATERHIKVA : <26> JÜH REATERHIKVA : <3> FÜH REATERHIKENTER : <3> FÜH REATERHIKENTER : : : : : : : : : : : : : : : : : : :	276 276 552 325 325 322 228 462 1378
RfaJ E. coll RfaJ S. coll RfaJ S. coll RfaJ S. typh RfbC k. pneu A. thaliana LpcA R. leg DUGT D.mel Gptl S. pomb FB3GalT hβ3GalT Brainiac Fringe LgtB N.men LgtB N.men LgtB N.men LgtB N.men LgtB N.men LgtB N.men LgtB S.go RfbF S.mar RfbF S.mar RfbF S.mar RfbF S.dys CaplE S.pne Rfak N.men AnsD B.any Tage B.sub Rfab E.coll	123: 120: 379: 379: 855: 1252:	IDENTIFY TO CREATE OF THE TANK AND A 22> THE STRUCT CASE TO COMPANY CASE TO CASE AND THE TO CREATE THE TANK AND A 22> THE STRUCTURE AND THE STRUCTURE STRUCTURE STRUCTURE AND THE STRUCTURE STRUC	ARE DENG NT OFFIL ARE DENG NT IVTI UIEDPRENT IVTI VIEDPRENT IVTI VIEDPLENVIHEN WILVIGPRINT-QTG IKSIPPODILWCOFW IVSIPPODILWCOFW IVSIPPODILWCOFW IVSIPPODILWCOFW NT SIPPODILWCOFW SLIIDNDN IIVKSN SLIIDNDN IIVKSN SLIIDNDN IIVKSN SLIIDNDN IIVKSN SLIIDNDN IIVKSN SLIIDNDN IIVKSN SLIIDNDN IIVKSN SLIIDNDN IIVKSN SLIIDNDN IVKSN SLIIDNDN SLIX	<pre></pre> <pre></pre> <pre></pre> <pre></pre> <pre></pre> <pre></pre> <pre></pre> <pre><td>276 2774 552 325 325 228 462 378 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5</td></pre>	276 2774 552 325 325 228 462 378 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
RfaJ E. coll RfaJ E. coll RfaJ S. typh RfbC k. pneu A. thallana LpCA R. leg DUGT D. mel Gpt1 S. pomb Family C rß3GalT bß3GalT bß3GalT Brainiac Fringe LgtB N. men LgtB N. men LgtB N. men LgtB N. gon Lext H. 1nf A. thallana Family D RfbF K. pn01 RfbF S. mar RfbF C. hyo RfbB S. dys CaplE S. pne RfaK N. men Amap E. awy Tage B. sub RfaB E. coll SiaD N. men	123: 120: 379: 165: 85: 1252:	IDRIATIONALIZACIONES OLI IDLIPALMENTANA 422> ENACTERISTICTULATIONES OLI IDLIPALMENTANA 422> ENACTERISTICTULATIONES OLIVICATIONES OLIVICATIONES AVECTORIA 422> ENACTERISTICATION 422>	ARE UNKNOWNESI TILD DER NTIYTI VIEDERANT-OTS ILD DER NTIYTI VIEDERANT-OTS IKS DEDUILWCOTW IVS DEDUILWCOTW IVS DEDUILWCOTW IVS DEDUILWCETW III KV	<17) FÜH TEPRERAMSWE : <21) LINITGATERHISWE : <25) JÜH REATERHISWA : <26) JÜH REATERHISWARA : <26) JÜH REATERHISWARA : <30) JÜH REATERHISMENTER : : : : : : : : : : : : : : : : : : :	276 276 552 325 325 322 228 462 1378
RfaJ E. coll RfaJ S. coll RfaJ S. coll RfaJ S. typh RfbC K. pneu A. thallana LpcA R. leg DUGT D. mel Gpt1 S. pomb FB3GalT hβ3GalT Braniac Fringe LgtB N. men LgtB S. dys CaplE S. pne RfbF K. pn01 RfbF S. mar RfbF S. mar RfbF S. mar RfbF S. sub RfaB S. coll SiaD N. men Family E Sub RfaB E. coll SiaD N. men Family E Cps14J S.pne	123: 123: 120: 379: 85: 85: 1252: 12	IDRIANIZACIÓ CRESTO EL IDLIVPALNEJ AVVA <22> EPARCETETICTETICT (23) LOOVAG- IELVANC IDRIANIZACIÓN (2000) ESQUELICALINGUACIÓN (21) DISSUYTE DIVERSACIA <22> PODVAG- IELADA VNN HATLA AVYCKUSI DOLLADIA (10) I.T. ERILANVK <21> DISSUYTE VIKURK <22> PODVAG- IELADA VNN HATLA AVYCKUSI DOLLADIA (10) I.T. ERILANVK <21> DISSUYTE VIKURK <22> PODVAG- IELADA VNN HATLA AVYCKUSI DOLLADIA (10) I.T. ERILANVK <21> DISSUYTE VIKURK <22> PODVAG- IELADA VNN HATLA AVYCKUSI DOLLADIA (10) I.T. ERILANVK <21> DISSUYTE VIKURK <22> PODVAG- IELADA VNN HATLA AVYCKUSI DOLLADIA (10) I.T. ERILANVK <21> DISSUYTE VIKURK <22> PODVAG- IELADA VNN HATLA AVYCKUSI DOLLADIA (10) I.T. ERILANVK <21> DISSUYTE VIKURK <22> DISSUETING (10) I.T. ERILANVK <21> DISSUETING (10) I.T. ERILANVK /110 I.T. E	ARE DENG NT OPEL ARE DENG NT OPEL ILL DENE NT IVTI VIEDPLE NVIHEN WLVIGPRINT-QTG IKSTPDDJLWCOFW ILL KVSUHTR AASUHTR AASUHTR AASUHTR AASUHTR AASUHTR AASUHTR SLIIDEDEGIIVKSN SLIIDEDEGIIVKSN SLIIDEDEGIIVKSN SLIIDEDEGIIVKSN SLIIDEDEGIIVKSN SLIIDEDEGIIVKSN SLIIDEDEGIIVKSN SLIIDENEGIIVKSN SLIIDENEGIIVKSN SLIIDENEGIIVKSN SLIIDENEGIIVKSN SLIIDENEGIIVKSN SLUTDGANGVVIEVI SKOYVKKUHS :10	<17> FÜH TEPREN SWG : <21> LINY TEPREN SWG : <25> JÜH REATER SWAA : <25> JÜH REATER SWG : <26> ÄÜH REATER SWG : <26> ÄÜH REATER SWG : <27> ILHISES SUPERID : <28< AUTOR SUPERION SUPERION : YILPEDUFYEVSARVG: 296 YILPEDUFYEVSARVG: 296 YILPEDUFYEVSARVG: 296 YILPEDUFYEVSARVG: 296 YILPEDUFYEVSARVG: 338 <	276 2774 552 325 322 228 462 378
Rfal S. coll RfaJ S. coll RfaJ S. coll RfaJ S. typh RfbC k. pneu A. thaliana LpcA R. leg DUGT D.mel Gptl S. pomb Family C r\$3GalT hβ3GalT Brainiac Fringe LgtB N.men LgtB N.men LgtB N.men LgtB N.men LgtB N.men RfbF S.mar RfbF S.mar	123: 123: 120: 379: 165: 85: 1252: 1	IDENSITY IN THE ACTIVE AS A DESCRIPTION OF A DESCRIPTI	ARE UNKNOWNESI ARE UNKNOWNESI TILD DREAMINTOPISI TILD DREAMINTOPISI WINTGERANTOPIS WINTGERANTOPIS WINTGERANTOPIS WINTGERANTOPIS WINTGERANTOPIS WINTGERANTOPIS WINTGERANTOPIS WINTGERANTOPIS WINTGERANTOPIS SLIDNDN IIVKSN SLIDNDN IIVKSN SLIDNDN IIVKSN SLIDNDN IIVKSN SLIDNT SLUAMIR EIVKDRETYIPOGE SLYDHORETYIPOGE SLYDHORE ILLDRI ALISAND VVVKN EIVKDRETYIPOGE SLYDHORE ILLDRI DI COPUNCHIVOFE DYSNYNEN CUVEKNE SKOYUMKIYHS :10 DEFAMISKNYOF :10 SPANERKIVGE :110	(17) FÜH TEPRENKISKE (21) LIHITGATEKHIKKA (25) JÜHRGATEKHIKKA (25) JÜHRGATEKHIKKA (26) JÜHRGATEKHIKKA (26) JÜHRGATEKHIKKA (27) IÜHRGATEKHIGATEK (27) JÜHRGATEK (28) JÜHRGATEK (28) JÜHRGATEK (28) JÜHRGATEK (28) JÜHRGATEK (29) JÜHRGATEK (29) JÜHRGATEK (20) JÜHR	276 276 552 325 325 228 462 1378
RfaJ E. coll RfaJ E. coll RfaJ S. typh RfbC K. pneu A. thallana LpcA R. leg DUGT D. mel Gpt1 S. pomb RfbGalT hβ3GalT Brainiac Fringe LgtB N. men LgtB N. men LgtB N. men LgtB N. men LgtB S. gon RfbF K. pn01 RfbF K. pn01 RfbF S. mar RfbF S. mar RfbF S. snar Rfab E. coll SiaD N.men Family E Cps14J S.pne LgtD N.gon Exco R. leg	123: 123: 123: 120: 379: 85: 1335: 1252: 12	I OTKINTPACTACKGS[DOP]IDURABLE[AVVA (22) EPACPTEXTELIT (23) LODUAL FILVAKG URLAND AV CKGSADTOLDIT - KRUAVKA (22) EPACPTEXTERT (23) PODVAL FILVAKG WILLAND VCKGSADTOLDIT - KRUAVKA (22) EPACPTEXTERT (23) PODVAL - KVTYSR WILLAND VCKGSADTOLDIT - KRUAVKA (22) EPACPTEXTERT (23) PODVAL - KVTYSR WILLAND VLAT (24) FILVAR (25) HILLAND VCG (25) EPACPTEXTERT (25) HILLANS (22) LODIEST - KVTYSR WILLAND VLAT (24) FILVAR (25) FILVAR (26) EPACPTEXTER (25) LODIEST - KVTYSR HILLAND VLAT (24) FILVAR (25) FILVAR (26) EPACPTEXTER (25) LODIEST - KVTYSR HILLAND VLAT (24) FILVAR (25) FILVAR (26) EPACPTEXTER (25) FILVAR (24) EPACPTEXTER (25) FILVAR (26) EPACPTEXTER (25) FILVAR (26) EPACPTEXTER (25) FILVAR (26) EPACPTEXTER (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) FILVAR (26) EPACPTEXTER (26) FILVAR (26)	ARE UNKNYNESI ARE UNKNYNESI ILL UNKNYNESI ILL UNKNYNESI ILL UNKNYNESI SKUYWKIYKS 10 SKUYWKIYKS 10 SKUYWKIYKS 11 SKUYWKIYKS 11 SKUYWKIYKI 11 SKUYYKI 11 SKUYWKIYKI 11 SKUYWKIYKI 11 SKUYWKIYKI 11 SKUYYKI 11 SKUYYYKI 11 SKUYYKI 11 SKUYYKI	(17) FÜH TEPRENISWE : (21) LININGATERINSWE : (25) JÜHRCATERINSWE : (25) JÜHRCATERINSWE : (267) LÜHRCATERINER (267) LÜHRCATERINER (27) LÜHRCATERINE (27) LÜHRCA	276 2774 552 325 228 462 378

binding domain (23-27). One of these residues, Y309, in human β 4-GalT was found to be conserved in all protein sequences (25). Interestingly, the DKKN motif is absent in the homologous snail β 4-GlcNAc-transferase, thus supporting the role of this region in discrimination of the Gal (or GlcNAc) moiety of the nucleotide sugar, as previously suggested by Bakker et al. (28). Several attempts were made to delineate the regions and the residues in the catalytic domain of β 4-GalTs that are responsible for UDP-Gal binding, and for α -LA and acceptor binding (23, 25, 26, 29). From these studies it appears that (i) the binding sites for α -LA and an acceptor are mostly located in the N-terminus of the catalytic domain in a region which encompasses residues 120-260 in human β 4-GalT, thus including the DVD motif; (ii) the UDP-Gal binding site seems to be mainly restricted to the C-terminal half region, *i.e.* residues 261-402 in human β 4-GalT; and (iii) residues in both regions are probably needed for correct accommodation of the donor and acceptor, and for catalysis. For the highly homologous blood A and B transferases, the modification of four critical amino acids (Arg176→Gly, Gly235 \rightarrow Ser, Leu266 \rightarrow Met, and Gly268 \rightarrow Ala) can affect the binding of both the sugar donor (UDP-GalNAc/UDP-Gal) and the acceptor. Recently, Seto et al. (30) suggested a role in enzyme turnover for residue 176, and a possible role in acceptor binding for residue 235, and also proposed that segments around residues 266 and 268 could be critical for binding of the nucleotide sugar. Thus, for both the eukaryotic α 3-Gal(NAc)Ts and β 4-GalTs, the nucleotide binding domain seems to be located in the carboxy-terminal portion of the catalytic domain, suggesting that these enzymes have a similar domain organization.

Family B. The *rfaI* gene from *Salmonella typhimurium* (group III) encodes an α 3-GalT involved in the biosynthesis of the hexose region of LPS. This protein exhibits 53% sequence identity with a protein encoded by the rfaI gene from Escherichia coli, which has a different function since it catalyzes the addition of an α 1,3-glucose to a glucose acceptor (31). As a result of the HCA analysis, a region spanning nearly 200 residues in RfaI proteins was found to exhibit significant homology with other classes of prokaryotic and eukaryotic glycosyltransferases. Among the homologous sequences (Table II), we found the bacterial α 4-GalTs (group V) from *Neisseriae* species, two bacterial α 2-glucosyltransferases (RfaJ), and several proteins of unknown function. Homology was also found with a sequence (LpcA) from Rhizobium leguminosarum, which was recently described as an α 6-galactosyltransferase (32). This sequence belonging to group VI exhibits no homology with the other known α 6-GalTs of group VII. The most striking finding is that the same homologous region was observed in a putative plant glycosyltransferase from

Arabidopsis thaliana, and in the two glycoprotein: α 3glucosyltransferases from Schizosaccharomyces pombe and Drosophila melanogaster. The two latter proteins are large soluble proteins which participate in the quality control mechanism for glycoprotein folding in the endoplasmic reticulum, where they catalyze the addition of a glucose to the protein-linked Man₇₋₉GlcNAc₂ (33, 34). These enzvmes exhibit a great degree of homology (60% identity) in the C-terminal domain (\sim 300 residues), the region which was found to be homologous to bacterial sequences. It is likely that this region corresponds to the catalytic domain responsible for the α -glucosylation of misfolded glycoproteins. All the family B enzymes in Table II have different acceptor and donor specificities, but they have in common the use of a UDP-sugar and the retaining character of the transfer reaction. The HCA plots shown in Fig. 2 illustrate the effectiveness of this method for detecting structural similarities in protein sequences with low levels of identity. The alignment is straightforward as the shapes of the hydrophobic clusters are relatively well conserved, and several regions of similarity could be identified. Multiple sequence alignment of the four most conserved regions found in proteins of family B is shown in Fig. 1. The DxD (DAQ in S. pombe Gpt1) motif present in region I just after a vertical hydrophobic cluster, indicative of a β -strand, is the most conserved feature observed in this family. It can be compared to the DVD motif described for family A, as the two motifs exhibit similar HCA patterns. Another motif, DQDxxN, in region III, where x is a non-aromatic hydrophobic residue or proline, is present in all but one sequence (A. thaliana). The sequence, HyxGxxKPW, in region IV, where y is an aromatic residue and x any type of amino acid, was also found in all sequences except the glycoprotein: α 3-glucosyltransferases.

Family C. A third family was defined which includes the β -GalTs of groups VIII and X. Like family B, it includes eukaryotic and prokaryotic sequences such as the rat β 3-GalT involved in the biosynthesis of gangliosides (35), and the β 4 GalTs (LgtB and LgtE) from *Neisseriae* species which participate to the biosynthesis of the lacto-N-neotetraose terminal LPS structure (36, 37). These galactosyltransferases exhibit homology with other protein sequences of unknown function, a sequence from A. thaliana and several sequences from C. elegans (Table II). Recently, Yuan et al. (38) reported that a class of signaling molecules involved in developmental processes, such as Fringe and Brainiac, may be glycosyltransferases, as they exhibit local homology with bacterial enzymes LgtB and LgtE. The similarities between the β 3-GalTs and Brainiac extend over the entire catalytic domain, with an overall identity score of 29%. Multialignment of the three most conserved regions found in proteins of family C is shown in Fig. 1. Region II contains an acidic motif, DDD (or EDD or DSD), surrounded by two stretches of hydrophobic residues. which appears to be characteristic of this family, and which also exhibits HCA similarity with the DVD and DxD motifs of families A and B, respectively.

Family D. This family comprises the protein sequences from groups IV and VII corresponding to $\alpha 3$ - and $\alpha 6$ -GalTs, respectively. The RfbF proteins (group IV) are involved in the biosynthesis of D-galactan I in LPS. In fact, RfbF from *Klebsiella pneumoniae* 0:1 was proposed to be a bifunctional enzyme adding disaccharide Gal*f* β 1,3Gal $p\alpha$ 1,

Fig. 1. Multiple sequence alignment of protein sequences belonging to families A to E. Protein names are as defined in Table II. The sequence alignments were performed using ClustalW (49) and refined manually from the HCA results. The invariant or highly similar residues are given in white letters on a black background and other conserved amino acids are in black on a grey background. Invariant Asp or Glu residues are indicated by asterisks. Dashes indicate gaps. Dots indicate that the consensus sequence is absent in the protein. Numbers in brackets indicate the numbers of amino acids between two conserved regions.



Fig. 2. HCA alignment of selected proteins of family B. Protein names are as indicated in Table II. The HCA plots were obtained from the drawhca server on the Internet (http://www.lmcp.jussieu.fr/ \sim soyer/www-hca/hca-form.html). The protein sequences are represented on a duplicated α -helical net, and the clusters of contiguous hydrophobic residues (V, I, L, F, M, W, and Y) are shown. The

3- to a lipid-linked GlcNAc (39). The Galf precursor was recently identified as UDP- α -D-Galf (40). These sequences exhibit significant similarities in their C-terminal ends with many other bacterial sequences. Most of them are considered to be putative glycosyltransferases but their functions have not yet been determined. Among the homologous proteins of which the functions have been determined, we found enzymes belonging to different classes of glycosyltransferases: an α 2-GlcNAc-transferase (RfaK) and a polysiallytransferase (SiaD), both from N. *meningitidis*, and an α -galacturonosyltransferase (Cap1E) from Streptococcus pneumoniae (Table II). Figure 1 shows linear alignment of the two most conserved regions identified in all these sequences. Two invariant glutamic acid residues in region II constitute a signature for this family. Interestingly, the same signature was recently reported for bacterial α -mannosyltransferases, and is present also in sucrose and sucrose phosphate synthase (41). The most striking feature is that the conserved motif characteristic of

analysis involved the visual comparison of hydrophobic cluster shapes and their distribution in order to find similarities between the sequences. The one-letter code is used for amino acids except for Gly, Pro, Ser, and Thr, which are represented by \blacklozenge , \bigstar , Ξ , and \Box , respectively. Vertical lines delineate the best conserved regions (I-IV). The invariant residues are indicated on a black background.

this family is also found in glycosyltransferases, which utilize very different nucleotide sugars (UDP- α -D-Gal, UDP- α -D-GalA, UDP- α -D-GlcNAc, GDP- α -D-Man, and CMP- β -NeuAc) and different mechanisms of action (retention and inversion of the anomeric configuration).

Family E. The group XI β 4-GalT from S. pneumoniae serotype 14 (Cps14J), involved in the biosynthesis of capsular polysaccharide (42), does not exhibit homology with other galactosyltransferase sequences, but exhibits extensive homology, in its N-terminal end, to a large number of glycosyltransferases of prokaryotic and eukaryotic origin. The most representative protein members of family E are listed in Table II, and multiple sequence alignment is shown in Fig. 1. It includes bacterial β 3-GlcNAc- and β 3-GalNAc-transferases. Cps14J also exhibits homology with Kfic from E. coli, which is a bi-functional enzyme catalyzing the biosynthesis of the repeated disaccharide unit, GlcA β 1,4GlcNAc α 1,4-, of the K5 antigen (43), and with the Exo proteins of Rhizobium meliloti

involved in the biosynthesis of succinoglycan (4). Significant homology was also observed with a family of mammalian α -GalNAc-transferases which catalyze the first step of mucin-type O-linked glycosylation (for a review see Ref. 44). Conserved regions that span more than one-hundred residues are found at the N-terminal end of all proteins but one, Kfic, in which it is located in the C-terminal portion of the protein. Three aspartate (or glutamate) residues as well as several other polar amino acids (Ser, Thr, and Asn) were found to be invariant in this family. In region II, the DxD motif is present in all sequences except in the polypeptide: α -GalNAc-transferases. The homology can be extended to other oligo and polysaccharide synthases since the same motif has been identified not only in NodC, a bacterial chitin oligosaccharide synthase, but also in cellulose, hyaluronan and other polymer synthases (43, 45, 46). Considering all these sequences, only two Asp residues (D40 and D90 in cps14J) were found to be invariant.

Other families. The ceramide β -GalTs (CGTs) of group XII exhibit no significant homology with other groups of GalTs. However, it is well known that these enzymes are highly homologous to the UDP-glucuronosyltransferases, suggesting an evolutionary link between CGTs and these detoxifying enzymes (47). The yeast α 2-GalTs (Group I) constitute a homogenous group with no evident link with other glycosyltransferases. Interestingly, the DxD motif was also observed in the HCA plots of this group of galactosyltransferases (data not shown).

Concluding remarks

The galactosyltransferases constitute a large and heterogenous class of enzymes. On the basis of their protein sequence identity they can be classified into twelve groups. This classification also reflects the nature of the glycosidic linkage formed. Within one group, the proteins are expected to be evolutionary related and to share a similar overall

3D structure. With the combination of HCA and BLAST analyses, we searched for local homologies which could be present in the different groups of galactosyltransferases as well as in other glycosyltransferases. This resulted in the definition of five families, each including protein sequences from various sources and classes of glycosyltransferases. Therefore, it should be noted that the prediction of a specific function for a newly identified gene solely on the basis of sequence similarity can be erroneous. These families present some interesting points. First, some families include enzymes from very diverse origins, ranging from bacteria to mammals and even plants (*i.e.* families B and C). Second, some families include enzymes catalyzing very different reactions, *i.e.* the transfer of a carbohydrate residue with inversion or retention of the anomeric configuration of the glycosidic linkage (families A, D, and E). All these observations suggest that these enzymes may have evolved from an ancestral gene before the divergence of prokaryotes and eukaryotes. Interestingly, in families A, B, and C the anomeric configuration of the acceptor is strictly conserved (β for families A and C, and α for family B). This may reflect the fact that the recognition of the acceptor was a driving force in the evolution of these enzymes.

The lengths and locations of the conserved regions in galactosyltransferases vary from one family to another, as illustrated in Fig. 3. In family A, the regions of similarity are mostly found in the C-terminal half of the catalytic domain, while in family E the two most conserved regions are localized in the N-terminal half of the proteins. The present analysis allowed the identification of the conserved amino acids in each family and therefore can serve as a template for further structure-function studies. Invariant glutamate or aspartate residues were indicated for each family, as it has been shown for fucosyltransferases that these amino acids exhibit side chain chemical properties





checkered areas indicate the transmembrane domain. The arrows indicate the minimal size of the catalytic domain as determined for mouse α 3-GalT (50) and bovine β 4-GalT (51).

appropriate for catalysis of the transfer reaction (48). Of greater interest is the presence of a possible signature for a wide range of glycosyltransferases. An acidic motif (DxD or DDD) present just after a vertical hydrophobic cluster, indicative of a β -strand, was found in almost all the families (Fig. 3), the exception being family D, which instead has the consensus sequence, ExxxxxxE. As this motif is widespread in different classes of glycosyltransferases, it is likely to be involved in the enzyme function. The common feature of all enzymes belonging to families A, B, C, and E is the use of a UDP-sugar. Therefore, the acidic motif could be involved in either UDP binding and/or the catalytic process.

Note Added in Proof: During the reviewing of the paper, one human (Y15014) and 3 mouse (AF029790, 91, and 92) β 3-GalT sequences were published. These enzymes, which are involved in the synthesis of the type 1 antigen (Gal β 1, 3GlcNAc β -R), are homologous to proteins of family C.

REFERENCES

- 1. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97-130
- Paulson, J.C. and Colley, K.J. (1989) Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. J. Biol. Chem. 264, 17615-17618
- Joziasse, D.H. (1992) Mammalian glycosyltransferases: genomic organization and protein structure. *Glycobiology* 2, 271-277
- Glucksmann, M.A., Reuber, T.L., and Walker, G.C. (1993) Family of glycosyltransferases needed for the synthesis of succinoglycan by *Rhizobium meliloti*. J. Bacteriol. 175, 7033-7044
- Liu, D., Haase, A.M., Lindqvist, L., Lindberg, A.A., and Reeves, P.R. (1993) Glycosyltransferases of O-antigen biosynthesis in Salmonella enterica: identification and characterization of transferase genes of groups B, C2, and E1. J. Bacteriol. 175, 3408-3413
- 6. Shaper, J.H., Joziasse, D.H., Meurer, J.A., Chou, T.-D.D., Schnaar, R.A., and Shaper, N.L. (1995) The chicken genome contains two functional non-allelic β 1,4-galactosyltransferase genes. *Glycoconjugate J.* 12, 477-478
- 7. Shaper, N.L., Meurer, J.A., Joziasse, D.H., Chou, T.-D.D., Smith, E.J., Schnaar, R.L., and Shaper, J.H. (1997) The chicken genome contains two functional non-allelic β 1,4-galactosyltransferase genes. Chromosomal assignment to syntenic regions tracks fate of the two gene lineages in the human genome. J. Biol. Chem. 272, 31389-31399
- 8. Almeida, R., Amado, M., David, L., Levery, S.B., Holmes, E.H., Merkx, G., van Kessel, A.G., Rygaard, E., Hassan, H., Bennett, E., and Clausen, H. (1997) A family of human β 4-galactosyltransferases: cloning and expression of two novel UDP-Galactose: β -N-acetylglucosamine β 1,4Galactosyltransferase, β 4Gal-T2 and β 4Gal-T3. J. Biol. Chem. 272, 31979-31991
- 9. Shaper, J.H. and Shaper, N.L. (1992) Enzymes associated with glycosylation. Curr. Opin. Struct. Biol. 2, 701-709
- Sandrin, M.S., Vaughan, H.A., Dabkowski, P.L., and McKenzie, I.F.C. (1993) Anti-pig IgM antibodies in human serum react predominantly with Gal(α1-3)Gal epitopes. Proc. Natl. Acad. Sci. USA 90, 11391-11395
- Oriol, R., Ye, Y., Koren, E., and Cooper, D.K.C. (1993) Carbohydrate antigens of pig tissues reacting with human natural antibodies as potential targets for hyperacute vascular rejection in pig-to-man organ xenotransplantation. *Transplantation* 56, 1433-1442
- Joziasse, D.H., Shaper, J.H., Van den Eijnden, D.H., Van Tunen, A.J., and Shaper, N.L. (1989) Bovine α1-3-galactosyltransferase: isolation and characterization of a cDNA clone. Identification of homologous sequences in human genomic DNA. J. Biol. Chem. 264, 14290-14297

- Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) Molecular genetic basis of the human histo-blood group ABO system. *Nature* 345, 229-233
- Bakker, H., Agterberg, M., Van Tetering, A., Koeleman, C.A.M., Van den Eijnden, D.H., and Van Die, I. (1994) A Lymnaea stagnalis gene, with sequence similarity to that of mammalian β1-4-galactosyltransferase, encodes a novel UDP-GlcNAc:Glc-NAcβ-Rβ1-4-N-acetylglucosaminyltransferase. J. Biol. Chem. 269, 30326-30333
- 15. Van Die, I., Bakker, H., and Van den Eijden, D.H. (1997) Identification of conserved amino acids in members of the $\beta 1 \rightarrow$ 4-galactosyltransferase gene family. *Glycobiology* 8, v-ix
- Breton, C., Oriol, R., and Imberty, A. (1998) Conserved structural features in eukaryotic and prokaryotic fucosyltransferases. *Glycobiology* 8, 87-94
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403-410
- 18. Uehara, K. and Muramatsu, T. (1997) Molecular cloning and characterization of β -1,4-galactosyltransferase expressed in mouse testis. *Eur. J. Biochem.* **244**, 706–712
- Gaboriaud, C., Bissery, V.L., Benchetrit, T., and Mornon, J.P. (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Lett.* 224, 149– 155
- Woodcock, S., Mornon, J.P., and Henrissat, B. (1992) Detection of secondary structure elements in proteins by hydrophobic cluster analysis. *Protein Eng.* 5, 629-635
- Yamamoto, F.-I. and Hakomori, S.-I. (1990) Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. J. Biol. Chem. 265, 19257-19262
- Haslam, D.B. and Baenziger, J.U. (1996) Expression cloning of Forssman glycolipid synthetase: a novel member of the histoblood group ABO gene family. Proc. Natl. Acad. Sci. USA 93, 10697-10702
- Wang, W., Wong, S.S., Fukuda, M.N., Zu, H., Liu, Z., Tang, Q., and Appert, H.E. (1994) Identification of functional cysteine residues in human galactosyltransferases. *Biochem. Biophys. Res. Commun.* 204, 701-709
- Boeggeman, E.E., Balaji, P.V., and Qasba, P.K. (1995) Functional domains of bovine β-1,4-galactosyltransferase. Glycoconj. J. 12, 865-878
- Aoki, D., Appert, H.E., Johnson, D., Wong, S.S., and Fukuda, M.N. (1990) Analysis of the substrate binding sites of human galactosyltransferase by protein engineering. *EMBO J.* 9, 3171-3178
- 26. Yadav, S. and Brew, K. (1990) Identification of a region of UDP-galactose: N-acetylglucosamine β 4-galactosyltransferase involved in UDP-galactose binding by differential labeling. J. Biol. Chem. 265, 14163-14269
- Zu, H., Fukuda, M.N., Wong, S.S., Wang, Y., Liu, W.Z., Tang, Q., and Appert, H.E. (1995) Use of site-directed mutagenesis to identify the galactosyltransferase binding sites for UDP-galactose. *Biochem. Biophys. Res. Commun.* 206, 362-369
- 28. Bakker, H., Van Tetering, A., Agterberg, M., Smit, A.B., Van den Eijnden, D.H., and Van Die, I. (1997) Deletion of two exons from the Lymnaea stagnalis β 1-4-N-acetylglucosaminyltransferase gene elevates the kinetic efficiency of the encoded enzyme for both UDP-sugar donor and acceptor substrates. J. Biol. Chem. 272, 18580-18585
- 29. Yadav, S. and Brew, K. (1991) Structure and function in galactosyltransferase. Sequence locations of α -lactalbumin binding site, thiol groups, and disulfide bond. J. Biol. Chem. 266, 698-703
- 30. Seto, N.O.L., Palcic, M.M., Compston, C.A., Li, H., Bundle, D.R., and Narang, S.A. (1997) Sequential interchange of four amino acids from blood group B glycosyltransferase boots catalytic activity and progressively modifies substrate recognition in human recombinant enzymes. J. Biol. Chem. 272, 14133-14138
- 31. Pradel, E., Parker, C.T., and Schnaitman, C.A. (1992) Structures of the rfaB, rfaI, rfaJ, and rfaS genes of *Escherichia coli* K-12 and

their roles in the assembly of the lipopolysaccharide core. J. Bacteriol. 174, 4736-4745

- 32. Allaway, D., Jeyaretnam, B., Carlson, R.W., and Poole, P.S. (1996) Genetic and characterization of a mutant that disrupts synthesis of the lipopolysaccharide core tetrasaccharide in *Rhizobium leguminosarum. J. Bacteriol.* 178, 6403-6406
- 33. Parker, C.G., Fessler, L.I., Nelson, R.E., and Fessler, J.H. (1995) Drosophila UDP-glucose:glycoprotein glucosyltransferase: sequence and characterization of an enzyme that distinguishes between denatured and native proteins. EMBO J. 14, 1294-1303
- 34. Fernandez, F., Jannatipour, M., Hellman, U., Rokeach, L.A., and Parodi, A.J. (1996) A new stress protein: synthesis of *Schizosaccharomyces pombe* UDP-Glc:glycoprotein glucosyltransferase mRNA is induced by stress conditions but the enzyme is not essential for cell viability. *EMBO J.* 15, 705-713
- Miyazaki, H., Fukumoto, S., Okada, M., Hasegawa, T., Furukawa, K., and Furukawa, K. (1997) Expression cloning of rat cDNA encoding UDP-galactose:GD₂ β1,3-galactosyltransferase that determines the expression of G_{D1b}, G_{M1}, G_{A1}. J. Biol. Chem. 40, 24794-24799
- Gotschlich, E.C. (1994) Genetic locus for the biosynthesis of the variable portion of Neisseria gonorrhoeae lipooligosaccharide. J. Exp. Med. 180, 2181-2190
- Jennings, M.P., Hood, D.W., Peak, I.R.A., Virji, M., and Moxon, E.R. (1995) Molecular analysis of a locus for the biosynthesis and phase variable expression of the lacto-N-neotetraose terminal lipopolysaccharide structure in Neisseriae meningitidis. Mol. Microbiol. 18, 729-740
- Yuan, Y.P., Schultz, J., Mlodzik, M., and Bork, P. (1997) Secreted fringe-like signaling molecules may be glycosyltransferases. *Cell* 88, 9–11
- Clarke, B.R., Bronner, D., Keenleyside, W.J., Severn, W.B., Richards, J.C., and Whitfield, C. (1995) Role of Rfe and RfbF in the initiation of biosynthesis of D-galactan I, the lipopolysaccharide O antigen from *Klebsiella pneumoniae* serotype O1. J. Bacteriol. 177, 5411-5418
- 40. Koplin, R., Brisson, J.R., and Whitfield, C. (1997) UDP-galactofuranose precursor required for formation of the lipopolysaccharide O antigen of *Klebsiella pneumoniae* serotype O1 is synthesized by the product of the rfbDKPO1 gene. J. Biol. Chem. 272, 4121-4128
- 41. Geremia, R.A., Petroni, E.A., Ielpi, L., and Henrissat, B. (1996) Towards a classification of glycosyltransferases based on amino

acid sequence similarities: prokaryotic α -mannosyltransferases. Biochem. J. 318, 133-138

- 42. Kolkman, M.A.B., Wakarchuk, W., Nuijten, P.J.M., and van der Zeijst, B.A.M. (1997) Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete cps locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the tetrasaccharide subunit. *Mol. Microbiol.* 26, 197-208
- Petit, C., Rigg, G.P., Pazzani, C., Smith, A., Sieberth, V., Stevens, M., Boulnois, G., Jann, K., and Roberts, I.S. (1995) Region 2 of the *Escherichia coli* K5 capsule gene cluster encoding proteins for the biosynthesis of the K5 polysaccharide. *Mol. Microbiol.* 17, 611-620
- Clausen, H. and Bennett, E.P. (1996) A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases control the initiation of mucin-type O-linked glycosylation. Glycobiology 6, 635-646
- 45. Saxena, I.M., Brown Jr., R.M., Fevre, M., Geremia, R.A., and Henrissat, B. (1995) Multidomain architecture of β-glycosyl transferases:implications for mechanism of action. J. Bacteriol. 177, 1419-1424
- Keenleyside, W.J. and Whitfield, C. (1996) A novel pathway for O-polysaccharide biosynthesis in Salmonella enterica serovar Borreze. J. Biol. Chem. 271, 28581-28592
- Stahl, N., Jurevics, H., Morell, P., Suzuki, K., and Popko, B. (1994) Isolation, characterization and expression of cDNA clones that encode rat UDP-galactose:ceramide galactosyltransferase. J. Neurosci. Res. 38, 234-242
- Murray, B.W., Takayama, S., Schultz, J., and Wong, C.-H. (1996) Mechanism and specificity of human α-1,3-fucosyltransferase. *Biochemistry* 35, 11183-11195
- 49. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUS-TAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680
- 50. Henion, T.R., Macher, B.A., Anaraki, F., and Galili, U. (1994) Defining the minimal size of catalytically active primate $\alpha 1,3$ -galactosyltransferase: structure-function studies on the recombinant truncated enzyme. *Glycobiology* 4, 193-201
- 51. Boeggeman, E.E., Balaji, P.V., Sethi, N., Masibay, A.S., and Qasba, P.K. (1993) Expression of deletion constructs of bovine β -1,4-galactosyltransferase in *Escherichia coli*: importance of Cys134 for its activity. *Protein Eng.* **6**, 779-785