Available online at www.sciencedirect.com
ScienceDirect
Theriogenology
Theriogenology 67 (2007) 1143–1149
www.theriojournal.com

Comparison of methods to stimulate ovarian follicular growth in cynomolgus and African green monkeys for collection of mature oocytes

N. Shimozawa [a,*], H. Okada [a,b], M. Hatori [a,c], T. Yoshida [a], T. Sankai [a]

[a] *Tsukuba Primate Research Center, National Institute of Biomedical Innovation, 1-1 Hachimandai, Tsukuba, Ibaraki 305-0843, Japan*
[b] *Japan Health Sciences Foundation, 13-4 Nihonnbashi-Kotenma-cho, Thuuo-ku, Tokyo 103-0001, Japan*
[c] *The Corporation for Production and Research of Laboratory Primates, 1-1 Hachimandai, Tsukuba, Ibaraki 305-0843, Japan*

Received 1 May 2006; accepted 14 October 2006

Abstract

The objective was to compare various gonadotropin-based methods to stimulate ovarian follicular growth in female cynomolgus ($n = 16$) and African green monkeys ($n = 8$) for collection of mature oocytes. On the 1st day of menstruation, the monkeys were treated with 3.75 mg leuprorelin acetate (a GnRH agonist). Starting 2–3 weeks later, ovarian follicular growth was stimulated as follows: (a) 25 IU/kg of human FSH (hFSH) in a glycerol solution given once daily for 9 d; (b) 200 IU of eCG given six times during a 9-d interval; (c) 75 IU/kg hFSH in a glycerol solution given three times (72 h intervals) during a 6-d interval. In addition, the monkeys were given 1200 or 4000 IU of hCG 36 h (Methods A and B) or 60 h (Method C) after the last gonadotropin treatment, and oocyte collection was attempted 36–38 h after hCG. Although there were no significant differences among methods in the number of oocytes collected, in cynomolgus monkeys, hFSH (Methods A and C) was better than eCG (Method B; 12 and 10 versus 7 mature oocytes, respectively), whereas in African green monkeys, eCG (Method B) was more effective than hFSH (Method A; 12 versus 7 mature oocytes). Furthermore, in cynomolgus monkeys, Method C was nearly as effective as Method A; using a glycerol solution as a solvent decreased the frequency of hFSH administration from nine to three times. In conclusion, in cynomolgus and African green monkeys, ovarian response depended on the species and on the individual, and in cynomolgus monkeys, hFSH in a glycerol solvent was effective.

Keywords: Monkey; Human FSH; eCG; Glycerol; Intracytoplasmic sperm injection

1. Introduction

Human embryonic stem (ES) cell lines have recently been established and regenerative medicine has rapidly advanced [1]. Applications to clinical medicine have been mainly based on research involving small laboratory rodents (e.g., mice or rats); before these research results can be directly applied to humans, preclinical studies involving non-human primates are needed, because they most closely resemble humans in numerous aspects. However, embryological studies in monkeys (e.g., in vitro culture and manipulation of oocytes and embryos) have lagged behind those in mice and humans. Methods to stimulate ovarian follicular growth in monkeys prior to oocyte collection need to be improved and simplified.

Collection of mature oocytes in monkeys typically includes administration of exogenous hormones to

* Corresponding author. Tel.: +81 29 837 2121; fax: +81 29 837 0218.
E-mail address: shimo@nibio.go.jp (N. Shimozawa).

0093-691X/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.theriogenology.2006.10.020

induce follicular growth. Ovarian stimulation protocols in monkeys differ from those used in other species, due to differences in reproductive physiology. Stimulation protocols vary widely [2–6], reflecting the difficulties inherent in working with primates. Therefore, there is considerable impetus to develop follicular growth stimulation protocols in non-human primates [7–11]. Cynomolgus and African green monkeys are commonly used in developmental studies (due to their widespread use in medical research and their origin as Vero and Cos7 cell lines, respectively). Furthermore, that both species are Old World monkeys, they most closely resemble humans in their evolution and thus are crucial experimental animals. The objective of the present study was to compare various gonadotropin-based methods to stimulate ovarian follicular growth in female cynomolgus and African green monkeys for collection of mature oocytes.

2. Materials and methods

2.1. Animals

We used 16 mature female and 2 mature male cynomolgus monkeys (*Macaca fascicularis*), and 8 mature female and 5 mature male African green monkeys (*Cercopithecus aethiops*). All animals were bred and maintained at the Tsukuba Primate Research Center, individually housed in an air-conditioned room with controlled illumination (12 h light:12 h dark), temperature (25 ± 2 °C), humidity (60 ± 5%) and ventilation (10 cycles/hr), given 75 g of commercial food twice daily (Type AS; Oriental Yeast Co., Ltd., Tokyo, Japan), 100 g of apples daily, and unlimited access to tap water. Every morning the general health (e.g., viability, appetite, fur-coat appearance) and menstruation status of each female monkey was assessed. The present study was conducted in accordance with the guidelines of the National Institutes of Biomedical Innovation for the care, use and biohazard countermeasures of laboratory animals.

2.2. Stimulation of ovarian follicular growth

On the 1st day of menstruation, female cynomolgus and African green monkeys were treated sc with 3.75 mg leuprorelin acetate, a GnRH agonist (GnRHa: Leuplin, Takeda Pharmaceutical Co., Ltd., Osaka, Japan). Two to three weeks later, ovarian follicular growth was stimulated by Methods A, B or C (Fig. 1) [12,13]. For Method A (Fig. 1A), 25 IU/kg of urinary human FSH (hFSH: Fertinorm, Serono Japan Co., Ltd., Tokyo, Japan) dissolved in a glycerol/physiological saline (1:1) solution was given sc once daily for 9 d (total of 225 IU/kg), and then 1200 IU of urinary human chorionic gonadotropin (hCG: Gonatropin, ASKA Pharmaceutical Co., Ltd., Tokyo, Japan) was given iv 36 h after the final hFSH treatment. For Method B (Fig. 1B), 200 IU of equine chorionic gonadotropin (eCG: Serotropin, ASKA Pharmaceutical Co.) dissolved in physiological saline was given im six times during a 10-d interval (total of 1200 IU), and 4000 IU of hCG was given im 36 h after the last eCG treatment. Finally, for Method C (Fig. 1C), female cynomolgus monkeys were given 75 IU/kg of hFSH sc every 3 d

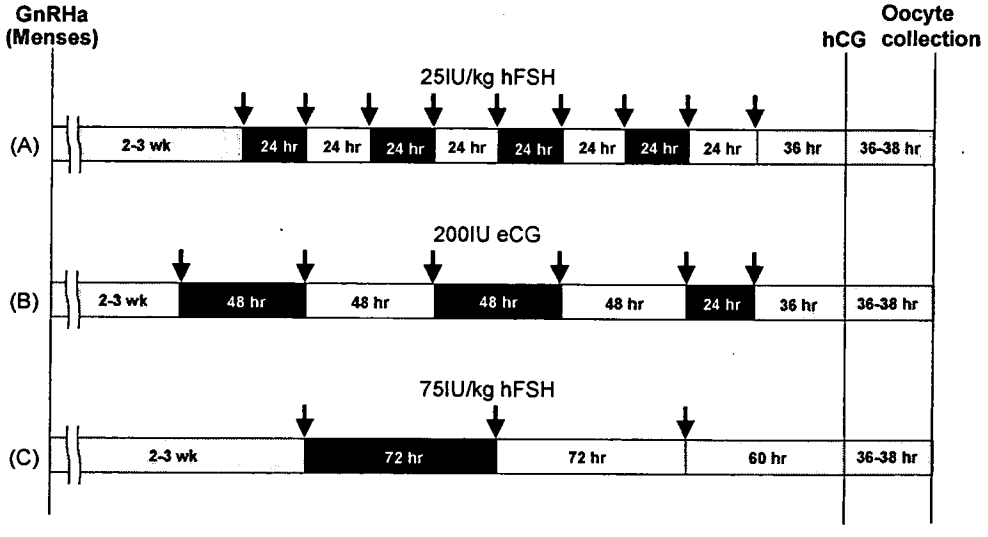


Fig. 1. Hormone administration methods applied to cynomolgus monkeys (A, B and C) and African green monkeys (A and B). A GnRH analogue was administrated on the 1st day of menses. Arrows indicate time of administration of hFSH and eCG (see Section 2 for detailed explanation).

during a 6-d interval (total of 225 IU/kg), followed by 1200 IU of hCG iv 60 h after the last treatment.

2.3. Oocyte collection

At 36–38 h after hCG administration, all females were anaesthetized with a combination of 10 mg/kg ketamine hydrochloride (Ketalar, Bayer Yakuhin Ltd., Osaka, Japan) and 1 mg/kg xylazine hydrochloride (Seractarl, Bayer Yakuhin Ltd.) given im. Ovaries were exposed through an abdominal incision, and the contents of the follicles were aspirated with a 25-gauge needle connected to a 2-mL syringe (Fig. 2A). The collected follicular contents were immediately diluted

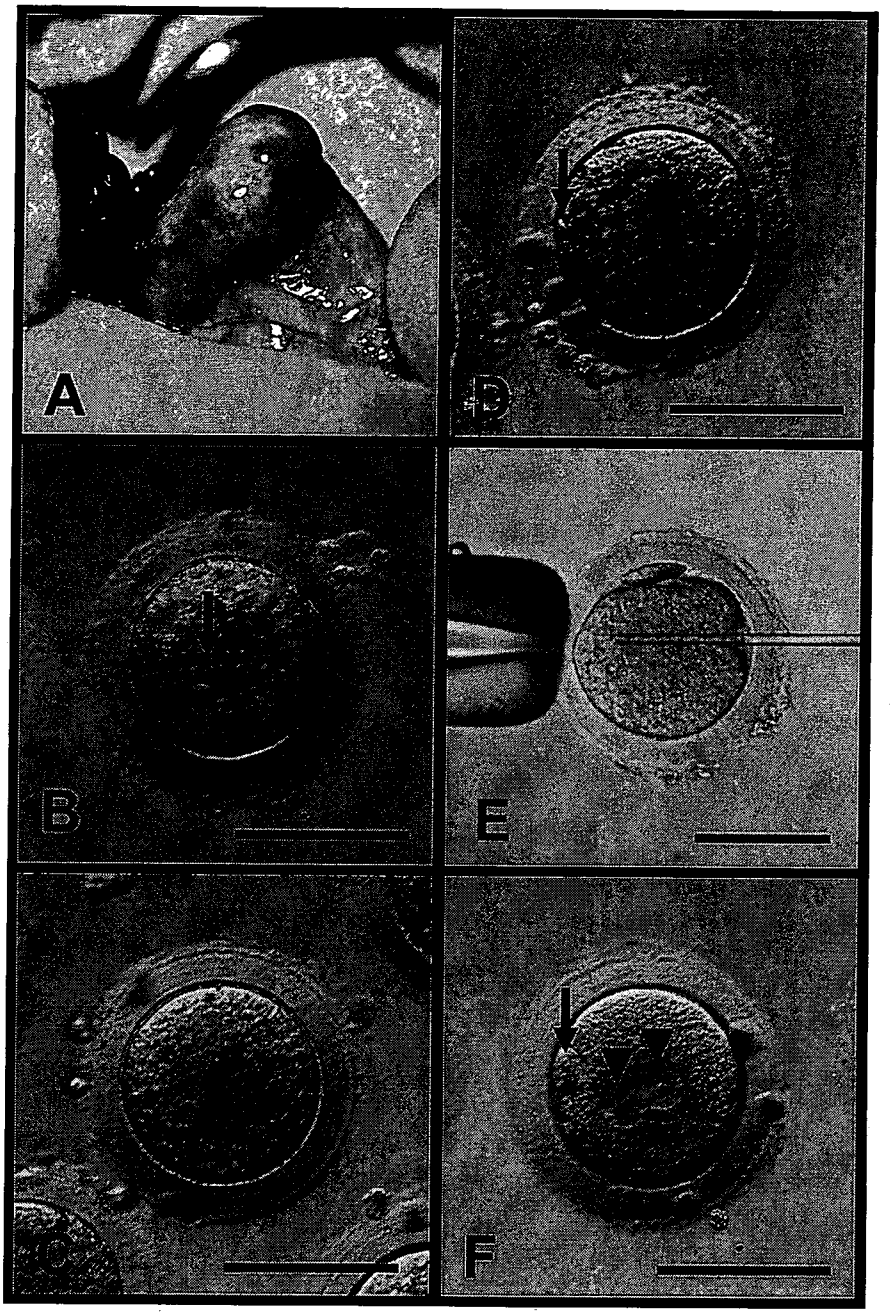


Fig. 2. (A) African green monkey ovary (at oocyte collection) with well developed follicles. (B–D) Oocytes collected from follicles. Germinal vesicle (GV) stage (B, arrow; GV), germinal vesicle breakdown (GVBD) stage (C) and mature oocytes (D, arrow; first polar body). (E) Intracytoplasmic sperm injection (ICSI). (F) Zygote with second polar body (arrow) and two pronuclei (arrowheads) after ICSI. Bar = 100 μm.

in a 50-mL centrifuge tube with TYH medium (119.37 mM NaCl, 4.78 mM KCl, 1.71 mM $CaCl_2 \cdot 2H_2O$, 1.19 mM $MgSO_4 \cdot 7H_2O$, 1.19 mM KH_2PO_4, 25.07 mM $NaHCO_3$, 5.56 mM glucose, 1.0 mM sodium pyruvate, 5 mg/mL BSA, penicillin-G (sodium salt) and streptomycin sulphate) [14] modified by adding Hepes (Hepes–TYH) containing 2.5 IU/mL of heparin (Novo Nordisk Pharma Co., Ltd., Tokyo, Japan). The preparation was treated with 0.1% hyaluronidase (Sigma, St. Louis, MO, USA) in Hepes–TYH medium and the oocytes were liberated from cumulus cells by pipetting. The oocytes were then washed with CMRL-1066 medium (Invitrogen, Carlsbad, CA, USA) containing 10% fetal bovine serum (FBS, Hyclone, Logan, UT, USA), Gluta-MAX (×100, Invitrogen), and penicillin–streptomycin solution (×100, Sigma), hereafter simply called CMRL. The washed oocytes were classified into the following three categories (Fig. 2B–D): (a) oocytes with a polar body were classified as metaphase II (M II: mature); (b) oocytes with a large nucleus were classified as the germinal vesicle (GV) stage; (c) oocytes without a polar body and large nucleus were classified as germinal vesicle breakdown (GVBD).

2.4. Fertilization ability

To determine the fertilization potential of the collected mature oocytes, microinjection was conducted by intracytoplasmic sperm injection (ICSI, Fig. 2E) using a micromanipulation system equipped with a piezo drive unit (Primetech, Ibaraki, Japan) under an inverted microscope (Nikon, Tokyo, Japan). According to the methods of Sankai et al. [15] and Okada et al. [16], male adult cynomolgus and African green monkeys were anaesthetized with ketamine hydrochloride and fresh semen was collected into TYH medium by electroejaculation. Each semen suspension was layered onto 90% Percoll (Amasham, Uppsala, Sweden) diluted with 9% NaCl solution and then centrifuged at 800 × *g* for 10 min. The precipitate (that contained sperm) was suspended in CZB (81.62 mM NaCl, 4.83 mM KCl, 1.7 mM $CaCl_2 \cdot 2H_2O$, 1.18 mM $MgSO_4 \cdot 7H_2O$, 1.18 mM KH_2PO_4, 15.0 mM $NaHCO_3$, 31.3 mM sodium lactate, 0.11 mM EDTA, 0.27 mM sodium pyruvate, 1.0 mM glutamine, 4 mg/mL BSA, penicillin-G (sodium salt) and streptomycin sulphate) [17], modified by adding Hepes (Hepes–CZB). Sperm suspended in Hepes–CZB with 10% polyvinylpyrrolidone (PVP, Sigma) was used in the microinjection. An immobilized spermatozoon drawn into the injection pipette (7–8 μm in diameter) was injected into the mature (M II) oocyte after breaking the plasma membrane with a few piezo pulses. Following injection of a sperm, oocytes were transferred into CMRL covered with mineral oil and cultured at 37.5 °C in a humidified atmosphere of 5% CO_2, 5% O_2 and 90% air. Fertilization potential of the oocytes was judged based on the formation of two pronuclei and the release of a second polar body (Fig. 2F), 15–16 h after ICSI.

2.5. Statistical analysis

One-way ANOVA analysis and Student's *t*-test were used to compare, among methods, the number of oocytes collected for cynomolgus monkeys and for African green monkeys, respectively. For all analyses, $P < 0.05$ was considered significant.

3. Results

3.1. Cynomolgus monkeys

For Method A, the mean and range of total oocytes collected were 31 and 10–92, of which 12 and 2–41 were mature oocytes (Table 1). For Method B, the corresponding values were 18 and 8–39 and 7 and 0–19, and for Method C, 21 and 11–32 and 10 and 0–17 (Table 1). More mature oocytes was collected following Methods A and C versus Method B although there were

Table 1
Mean (±S.E.M.) oocytes collected from cynomolgus and African green monkeys following stimulation of ovarian follicular growth

Monkey	Method	Number of monkeys	Age of monkeys (month)	Number of oocytes collected	Number of normal oocytes at each stage (%)		
					M II	GVBD	GV
Cynomolgus	A	7	96 ± 11	31 ± 12	12 ± 6 (38.7)	8 ± 4 (25.8)	9 ± 3 (29.0)
	B	8	98 ± 10	18 ± 4	7 ± 2 (38.9)	4 ± 3 (22.2)	7 ± 2 (38.9)
	C	5	88 ± 12	21 ± 4	10 ± 3 (47.6)	3 ± 2 (14.3)	7 ± 1 (33.3)
African green	A	6	138 ± 10	19 ± 6	7 ± 2 (36.8)	2 ± 1 (10.5)	3 ± 1 (15.8)
	B	5	121 ± 3	18 ± 4	12 ± 4 (66.7)	1 ± 1 (5.6)	1 ± 1 (5.6)

Table 2
Fertilization ability of *in vivo*-matured oocytes in cynomolgus and African green monkeys

Monkey	Method	Number of monkeys	Number of oocytes examined	Number of oocytes survived (%)	Number of fertilized oocytes with (%/examined)[a]			Number of oocytes unfertilized (%/examined)[b]
					2PN2PB	3PN1PB	mPN2PB	
Cynomolgus	A	4	62	53 (85.5)	28 (45.2)	2 (3.2)	1 (1.6)	22 (41.5)
	B	1	8	8 (100)	8 (100)			
African green	A	4	29	29 (100)	17 (58.6)			12 (41.4)
	B	4	36	36 (100)	28 (77.8)			8 (22.2)

[a] 2PN2PB, two pronuclei and two polar bodies; 3PN1PB, three pronuclei and one polar body; mPN2PB, multiple pronuclei and two polar bodies.
[b] Neither release of PB nor formation of PN were observed.

no significant differences among the three methods for total oocytes and mature oocytes ($P = 0.50$ and 0.57, respectively; Table 1).

3.2. African green monkeys

For Method A, the mean and range of total oocytes collected were 19 and 4–41, of which 7 and 0–13 were mature oocytes (Table 1). For Method B, the corresponding values were 18 and 5–27 and 12 and 4–24 (Table 1). More mature oocytes was collected following Method B versus Method A although there were no significant differences between the two methods for total oocytes and mature oocytes ($P = 0.90$ and 0.22, respectively; Table 1).

3.3. Fertilization ability

In the cynomolgus monkeys, ICSI was conducted using the mature oocytes collected from four and one monkeys after administration of Methods A and B, respectively. Of the 62 oocytes from Method A that were injected with sperm, 28 (45.2%) produced normal zygotes with female and male pronuclei and a second polar body, and all eight oocytes from Method B (100%) produced such zygotes (Table 2). Fertilization ability in oocytes collected after administration of Method C was not evaluated because all of those oocytes were used for other research. In African green monkeys, ICSI was conducted using the mature oocytes from four monkeys collected after Method A and from four monkeys after Method B. From these two sources, 17 of 29 oocytes (58.6%) and 28 of 36 oocytes (77.8%) developed into normal zygotes (Table 2).

4. Discussion

In this study, follicular growth stimulation methods were compared relative to the collection of a high number of mature oocytes in cynomolgus monkeys and African green monkeys. The most effective regimens depended on the type of monkey, although no significant differences were observed, and in cynomolgus monkeys a glycerol solution as the solvent decreased the hormone administration frequency.

Follicular growth stimulation methods to effectively collect a high number of mature oocytes from cynomolgus monkeys and African green monkeys were compared. A higher yield of mature oocytes was achieved with a combination of hFSH sc and hCG iv in cynomolgus monkeys, and a combination im administration of eCG and hCG in African green monkeys. Therefore, the hormone administration methods that effectively induced follicular growth differed according to the type of monkey, suggesting that sensitivity to a specific hormone varies among genera (*Macaca* and *Cercopithecus*). If the reason for this sensitivity can be identified, perhaps more effective follicular growth stimulation methods can be developed.

The proportion of oocytes that were mature oocytes in this study seemed lower than in other reports, including those in rhesus monkeys [5,18]. That maturation of oocytes was not high in our current study might be that the monkeys either had been excluded from the breeding colony at our research center due to long-term pregnancy failures or had not been mated for a long time. Sankai et al. [19,20] previously showed that the number of total oocytes from eCG-treated cynomolgus and African green monkeys was approximately 29 and 41, respectively, although the number of mature oocytes was not reported. The collection rate in our study seemed lower, namely, 18 and 18, respectively. One reason for this difference might be the poor fertility rate of the monkeys that we used, in contrast to the fertile monkeys that Sankai et al. used. Another reason might be that we induced follicular growth by suppressing endogenous gonadotropin [12,13] via GnRHa administration, whereas

Sankai et al. did not administer GnRHa. This difference in the number of collected oocytes also might reflect the difference in the number of total oocytes.

To examine the effect of glycerol in the solvent on the number of collected oocytes, we administered hFSH to cynomolgus monkeys every 3 d for a total period of 6 d (Method C). The number of mature oocytes collected after Method A was similar to that collected after Method C. The results clearly demonstrated the gradual, continuous effect of hFSH dissolved in 50% glycerol solution, and thus the possibility of decreasing the administration frequency by using this glycerol solution. A previous study reported that glycerol prevented loss of immunoreactive FSH from urine [21]. Stabilization of hFSH in a glycerol solution might allow a continual effect for follicular growth. Hormones dissolved in a solvent containing a high viscosity chemical (e.g., polyvinylpyrrolidone) as glycerol have been administered to cows, ewes and rabbits [22–24]. In our study, we demonstrated that collection of mature oocytes following this method of administration was also possible in cynomolgus monkeys, but not in African green monkeys due to a lower sensitivity to FSH at this dose. However, differences among Methods A, B and C in the amount of hCG given or the route of the administration may have affected maturity of oocytes. Further studies are needed to determine the optimal amount of hCG, the route of administration and the interval from treatment to aspiration.

Replicated tests by the same hormone administration to the same individual would be difficult, because an antibody is produced in monkeys given hormones derived from humans or horses [25–28]. Binding to antibodies formed after the first hCG administration might attenuate a reaction to human gonadotropins after the second follicle stimulation protocol [25,27]. VandeVoort and Tarantal [28] showed that no significant decrease in the number of collected oocytes was detected in replicates of follicular growth stimulation methods up to five times of the hFSH administration protocol. In this study, four cynomolgus monkeys and three African green monkeys were subjected to two follicular growth stimulation methods (hFSH and eCG), and were thus administered hCG twice. No significant difference in the rates of M II and GVBD stage oocytes to collected normal oocytes was detected between the first hCG administration and second hCG administration (data not shown), indicating that the anti-hCG antibody might not affect the maturation of oocytes induced by hCG administration.

Based on the ability of pronuclear formation and polar body released in the zygotes using the ICSI method, normal zygotes with two pronuclei and a second polar body were produced in the mature oocytes by either type of administration for both types of monkeys (Table 2). The in vivo mature oocytes collected by administration of either hormone used in this study had fertilization ability. However, some of the collected oocytes treated with ICSI were not activated, perhaps due to incomplete maturity or aging of the oocytes. However, many of the mature oocytes collected by the three hormone administration methods reported in this study could be applied to fertilization-related experiments.

In mice, after induction of follicular growth (i.e., oocyte growth induced by administration of eCG), maturity of oocytes and ovulation can be induced by hCG administration and the oocytes ovulated into the oviduct can be collected. In monkeys, however, follicular aspiration is typically performed using a needle because ovulation of most follicles does not occur even if the time from the hCG administration to oocyte collection is extended (unpublished data), or because all of the oocytes might not be captured by fimbria even if all of the follicles ovulated. The difference in fertilization ability among the oocytes might explain the difference in the degree of maturity during oocyte growth. The immature oocytes (e.g., GV and GVBD stage oocytes) collected in this study supported this hypothesis. The existence of aged mature oocytes might be the main cause for the decrease in fertilization ability because aged oocytes have lower fertilization ability due to chromosome aberration [29,30]. Collection of good quality oocytes can possibly be established by detailed examination of the time after hCG administration for collection of mature oocytes.

Mature oocytes are crucial experimental material in embryological studies that aim toward production of embryos or individuals. In particular, the low production rate of transgenic or embryonic/somatic cloned animals demands the development of a method to acquire high numbers of mature oocytes. In conclusion, in cynomolgus monkeys and African green monkeys, species and individual differences in ovarian reactivity were clearly evident. Additionally, in cynomolgus monkeys, use of a glycerol diluent apparently prolonged the duration of hormone activity. The three hormone administration methods reported here will be useful for further studies in collecting oocytes from monkeys.

Acknowledgments

This study was supported by grants from the Ministry of Health, Labour and Welfare of Japan. We

are grateful to the staff of the Corporation for Production and Research of Laboratory Primates for their kind cooperation in the collection of samples.

References

[1] Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1145–7.
[2] Hewitson L, Takahashi D, Dominko T, Simerly C, Schatten G. Fertilization and embryo development to blastocysts after intracytoplasmic sperm injection in the rhesus monkey. Hum Reprod 1998;13:3449–55.
[3] Nusser KD, Mitalipov S, Widmann A, Gerami-Naini B, Yeoman RR, Wolf DP. Developmental competence of oocytes after ICSI in the rhesus monkey. Hum Reprod 2001;16:130–7.
[4] Ogonuki N, Tsuchiya H, Hirose Y, Okada H, Ogura A, Sankai T. Pregnancy by the tubal transfer of embryos developed after injection of round spermatids into oocyte cytoplasm of the cynomolgus monkey (*Macaca fascicularis*). Hum Reprod 2003;18:1273–80.
[5] Hayes ES, Curnow EC, Trounson AO, Danielson LA, Unemori EN. Implantation and pregnancy following in vitro fertilization and the effect of recombinant human relaxin administration in *Macaca fascicularis*. Biol Reprod 2004;71:1591–7.
[6] Ng SC, Chen N, Yip WY, Liow SL, Tong GQ, Martelli B, et al. The first cell cycle after transfer of somatic cell nuclei in a non-human primate. Development 2004;131:2475–84.
[7] Wolf DP, Alexander M, Zelinski-Wooten M, Stouffer RL. Maturity and fertility of rhesus monkey oocytes collected at different intervals after an ovulatory stimulus (human chorionic gonadotropin) in in vitro fertilization cycles. Mol Reprod Dev 1996;43:76–81.
[8] Cseh S, Corselli J, Chan P, Bailey L. Superovulation using recombinant human FSH and ultrasound-guided transabdominal follicular aspiration in baboon (*Papio anubis*). Anim Reprod Sci 2002;70:287–93.
[9] VandeVoort CA, Leibo SP, Tarantal AF. Improved collection and developmental competence of immature macaque oocytes. Theriogenology 2003;59:699–707.
[10] Marshall VS, Browne MA, Knowles L, Golos TG, Thomson JA. Ovarian stimulation of marmoset monkeys (*Callithrix jacchus*) using recombinant human follicle stimulating hormone. J Med Primatol 2003;32:57–66.
[11] Yoshimoto N, Shimoda K, Mori Y, Honda R, Okamura H, Ide Y, et al. Ovarian follicular development stimulated by leuprorelin acetate plus human menopausal gonadotropin in chimpanzees. J Med Primatol 2005;34:73–85.
[12] Fleming R, Coutts JR. Induction of multiple follicular growth in normally menstruating women with endogenous gonadotropin suppression. Fertil Steril 1986;45:226–30.
[13] Torii R, Hosoi Y, Masuda Y, Iritani A, Nigi H. Birth of Japanese monkey (*Macaca fusucata*) infant following in vitro fertilization and embryo transfer. Primates 2000;39:399–406.
[14] Toyoda Y, Yokoyama M, Hoshi F. Studies on the fertilization of mouse eggs in vitro. I. In vitro fertilization of eggs by fresh epididymal sperm. Jpn J Anim Reprod 1971;16:147–51 (in Japanese).
[15] Sankai T, Terao K, Yanagimachi R, Cho F, Yoshikawa Y. Cryopreservation of spermatozoa from cynomolgus monkeys (*Macaca fascicularis*). J Reprod Fertil 1994;101:273–8.
[16] Okada A, Igarashi H, Kuroda M, Terao K, Yoshikawa Y, Sankai T. Cryopreservation-induced acrosomal vesiculation in live spermatozoa from cynomolgus monkeys (*Macaca fascicularis*). Hum Reprod 2001;16:2139–47.
[17] Chatot CL, Ziomek CA, Bavister BD, Lewis JL, Torres I. An improved culture medium supports development of random-bred 1-cell mouse embryos in vitro. J Reprod Fertil 1989;86:679–88.
[18] Borman SM, Chaffin CL, Schwinof KM, Stouffer RL, Zelinski-Wooten MB. Progesterone promotes oocyte maturation, but not ovulation, in nonhuman primate follicles without a gonadotropin surge. Biol Reprod 2004;71:366–73.
[19] Sankai T, Cho F, Yoshikawa Y. In vitro fertilization and pre-implantation embryo development of African green monkeys (*Cercopithecus aethiops*). Am J Primatol 1997;43:43–50.
[20] Sankai T, Ogonuki N, Tsuchiya H, Shimizu K, Cho F, Yoshikawa Y. Comparison of results from IVF-related studies for cynomolgus monkeys, Japanese monkeys, African green monkeys, and red-bellied tamarins. J Fertil Implant (Tokyo) 1998;15:177–9.
[21] Livesey JH, Roud HK, Metcalf MG, Donald RA. Glycerol prevents loss of immunoreactive follicle-stimulating hormone and luteinizing hormone from frozen urine. J Endocrinol 1983;98:381–4.
[22] Kanayama K, Sankai T, Nariai K, Endo T, Sakuma Y. Simplification of superovulation induction by using polyvinylpyrrolidone as a solvent for FSH in rabbits. J Vet Med Sci 1994;56:599–600.
[23] Sugano M, Shinogi T. Superovulation induction in Japanese Black cattle by a single intramuscular injection of hMG or FSH dissolved in polyvinylpyrrolidone. Anim Reprod Sci 1999;55:175–81.
[24] D'Alessandro AG, Martemucci G, Colonna MA, Borghese A, Terzano MG, Bellitti A. Superovulation in ewes by a single injection of pFSH dissolved in polyvinylpyrrolidone (PVP): effects of PVP molecular weight, concentration and schedule of treatment. Anim Reprod Sci 2001;5:255–64.
[25] Ottobre JS, Stouffer RL. Antibody production in rhesus monkeys following prolonged administration of human chorionic gonadotropin. Fertil Steril 1985;43:122–8.
[26] Bavister BD, Dees C, Schultz RD. Refractoriness of rhesus monkeys to repeated ovarian stimulation by exogenous gonadotropins is caused by nonprecipitating antibodies. Am J Reprod Immunol Microbiol 1986;11:11–6.
[27] Wolf DP, Vandevoort CA, Meyer-Haas GR, Zelinski-Wooten MB, Hess DL, Baughman WL, et al. In vitro fertilization and embryo transfer in the rhesus monkey. Biol Reprod 1989;41:335–46.
[28] VandeVoort CA, Tarantal AF. Recombinant human gonadotropins for macaque superovulation: repeated stimulations and post-treatment pregnancies. J Med Primatol 2001;30:304–7.
[29] Smith AL, Lodge JR. Interactions of aged gametes: in vitro fertilization using in vitro-aged sperm and in vivo-aged ova in the mouse. Gamete Res 1987;16:47–56.
[30] Mailhes JB, Young D, London SN. Postovulatory ageing of mouse oocytes in vivo and premature centromere separation and aneuploidy. Biol Reprod 1998;58:1206–10.

JOURNAL OF INTERFERON & CYTOKINE RESEARCH 27:32–37 (2007)

DOI: 10.1089/jir.2007.0099

Identification of a Novel CXCL1-Like Chemokine Gene in Macaques and Its Inactivation in Hominids

HISAYUKI NOMIYAMA,[1] KAORI OTSUKA-ONO,[1] RETSU MIURA,[1] NAOKI OSADA,[2] KEIJI TERAO,[3] OSAMU YOSHIE,[4] and JUN KUSUDA[2]

ABSTRACT

Chemokines are a rapidly evolving cytokine gene family. Because of various genome rearrangements after divergence of primates and rodents, humans and mice have different sets of chemokine genes, with humans having members outnumbering those of mice. Here, we report the occurrence of lineage-specific chemokine gene generation or inactivation events within primates. By using human chemokine sequences as queries, we isolated a novel cynomolgus macaque CXC chemokine cDNA. The encoded chemokine, termed CXCL1L (from CXCL1-like) showed the highest similarity to human CXCL1. A highly homologous gene was also found in the rhesus macaque genome. By comparing the genome organization of the major CXC chemokine clusters among the primates, we found that one copy of the duplicated CXCL1 genes turned into a pseudogene in the hominids, whereas the gene in macaques has been maintained as a functionally active CXCL1L. In addition, cynomolgus macaque was found to contain an additional CXC chemokine highly homologous to CXCL3, termed CXCL3L (from CXCL3-like). These results demonstrate the birth-and-death process of a new gene in association with gene duplication within the primates.

INTRODUCTION

CYNOMOLGUS MACAQUE (*Macaca fascicularis*) and rhesus macaque (*Macaca mulatta*) are closely related old world monkeys commonly used in experimental and toxicologic studies for drug and vaccine development.[1–4] Although both macaques are considered phylogenetically very close to humans, possible genetic differences between macaques and humans that may cause interspecies differences in drug responses and toxicity should be taken into account when the data obtained from macaques are extrapolated to humans. To unveil the genetic differences in primates and also to help identify genes in the human genome, expressed sequence tag (EST) and genome sequencing projects of these macaques are underway.

Chemokines are a large family of cytokines that regulate inflammation, leukocyte trafficking, and immune cell development.[5–7] There are at least 46 chemokine members in humans. Based on the arrangement of the conserved cysteine residues, chemokines are classified into four subfamilies. Two main subfamilies are CXC and CC chemokines, which have the first two conserved cysteines separated by one amino acid or juxtaposed, respectively. Chemokines can also be divided into two functional subgroups. Inflammatory chemokines attract mainly monocytes and neutrophils and mediate innate immunity, whereas homeostatic chemokines are constitutively expressed in organs, such as lymphoid tissues, and are involved in relocation of lymphocytes and dendritic cells (DCs).

The inflammatory CXC and CC chemokines are known to form a large gene cluster.[7] Human CXC and CC inflammatory chemokine gene clusters reside on chromosomes 4 and 17, respectively, and the respective gene clusters in mice are located on chromosomes 5 and 11. Comparison of these gene cluster organizations shows that the chemokine gene content in each cluster is greatly different between human and mouse due to lineage-specific gene duplication or deletion events or both after the divergence of primates and rodents.[8,9] In contrast, the

[1]Department of Molecular Enzymology, Kumamoto University Graduate School of Medical Sciences, Kumamoto 860-8556, Japan.
[2]Division of Biomedical Research Resources, National Institute of Biomedical Innovation, Ibaraki, Osaka 567-0085, Japan.
[3]Tsukuba Primate Research Center, National Institute of Biomedical Innovation, Tsukuba, Ibaraki 305-0843, Japan.
[4]Department of Microbiology, Kinki University School of Medicine, Osaka-Sayama, Osaka 589-8511, Japan.

Sequence data from this paper have been submitted to the GenBank/EBI/DDBJ databases with accession nos. AB262775 (CXCL1), AB262776 (CXCL2), AB262777 (CXCL3), AB262778 (CXCL1L), and AB262779 (CXCL3L).

emergence of noncluster chemokines, most of which are homeostatic chemokines, apparently predated the divergence of primates and rodents.[8,9]

To determine if gene rearrangements within the chemokine clusters occurred even after the diversification of humans and nonhuman primates, we searched the EST and genome databases of nonhuman primates for novel chemokines.

MATERIALS AND METHODS

cDNA cloning

Chemokine cDNAs were cloned by reverse transcriptase-polymerase chain reaction (RT-PCR). Total RNAs were prepared from various tissues of cynomolgus macaque and reverse transcribed. The cDNAs were synthesized using PrimeSTAR HS DNA polymerase (Takara Bio, Kyoto, Japan) and cloned with Mighty Cloning kit (Takara Bio). The primer sequences were designed based on the rhesus macaque genome sequences and were CXCL1, 5′-CTCCAGCTCCTCGCACAG and 5′-AGCCACCAATGAGCTTCTTC; CXCL1L, 5′-AGTTCCCCTGCTCCTCTCAC and 5′-GCCAGTATTTCTGACCAACG; CXCL2, 5′-CCGAAACGCCTGCTGAG and 5′-CTTCAGGAACAGCCACCAAT; CXCL3 and CXCL3L, 5′-TCCCATCCTGCTGAG and 5′-CCGCAGGAAGTGTCAATGT. The cDNAs used in the cDNA cloning as templates are those of liver (CXCL1 and CXCL2), spleen (CXCL1L), and stomach (CXCL3 and CXCL3L). The PCR conditions were 30 cycles at 98°C for 10 sec, 60°C for 5 sec, and 72°C for 1 min.

Tissue expression analysis of CXCL1 and CXCL1L

The prepared cDNAs were amplified with Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA). The primer sequences were CXCL1L, 5′-AGGGAATTCACCCCAAGAAC and 5′-GCAAACTCACCTGTTCAGCA; glyceraldehyde 3-phosphate dehydrogenase (GAPDH), 5′-GCCAAGGTCATCCATGACAACTTTGG and 5′-GCCTGCTTCACCACCTTCTTGATGTC. The primers used for CXCL1 were as described.

FIG. 1. Comparison of amino acid sequences of chemokines CXCL1, CXCL1L, CXCL2, CXCL3, and CXCL3L from human, chimpanzee, rhesus macaque, and cynomolgus macaque. Sequences other than those of cynomolgus macaque were taken from Ensembl (*www.ensembl.org/*) or UCSC Genome Browser (*genome.ucsc.edu/*). The amino-terminal sequence of the chimpanzee CXCL1 sequence is still unknown because of a gap in the genome sequence. Conserved four cysteine residues are boxed. ELR (Glu-Leu-Arg) motif is indicated by dots under the cynomolgus macaque CXCL3L sequence. Signal sequences are shown as lowercase letters. hs, human; pt, chimpanzee; mm, rhesus macaque; mf, cynomolgus macaque.

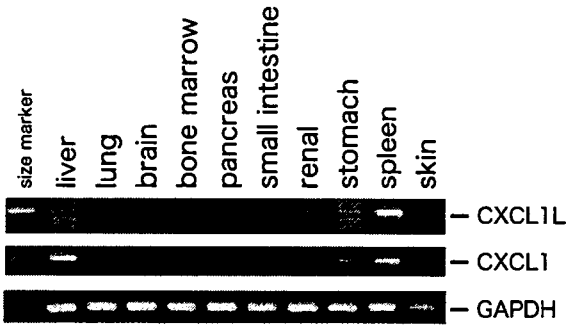


FIG. 2. RT-PCR analyses of CXCL1L and CXCL1 mRNAs in various cynomolgus macaque tissues. cDNAs were prepared from various tissues and amplified by PCR. GAPDH was used as an internal control.

The housekeeping gene GAPDH was used as an internal control. The PCR conditions were 35 cycles at 94°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min. The product sizes were 170 bp (CXCL1L), 398 bp (CXCL1), and 313 bp (GAPDH).

Computational methods

Signal sequences were predicted by SignalP (*www.cbs.dtu.dk/services/SignalP/*). Dot-plot analysis was performed using PipMaker (*pipmaker.bx.psu.edu/pipmaker/*) with the all matches option. For phylogenetic analysis, the chemokine amino acid sequences were aligned with ClustalX (*bips.u-strasbg.fr/fr/Documentation/ClustalX/*). The Neighbor-Joining tree was constructed with PAUP* (*paup.csit.fsu.edu/*) using the protein-Poisson distances, and only >50% bootstrap values are shown at each node (1000 replications). CXCL1P (human and chimpanzee CXCL1 pseudogene, exons 1 plus 2 sequences) and CXCL7P1 (PPBPL1, human, chimpanzee and rhesus macaque CXCL7 pseudogene exon sequences) were also used in the tree construction. The codon frames were inferred from the CXCL1 and CXCL7 genes.

RESULTS AND DISCUSSION

When human chemokine sequences were used as queries, one novel chemokine sequence (clone QnpA-12174) was found in the cynomolgus macaque EST database (Japan National Institute of Biomedical Innovation JCRB Gene Bank, *genebank.nibio.go.jp/*). Because the cDNA was a chimeric and 5′-truncated clone, the rhesus macaque genome database[10] (UCSC Genome Browser Database, *genome.ucsc.edu/*) was searched with the sequence, and the corresponding gene was found to be located in the inflammatory CXC chemokine gene cluster on chromosome 5 (position 55,707,566–55,708,094, January 2006 assembly). Based on this gene sequence, a PCR primer

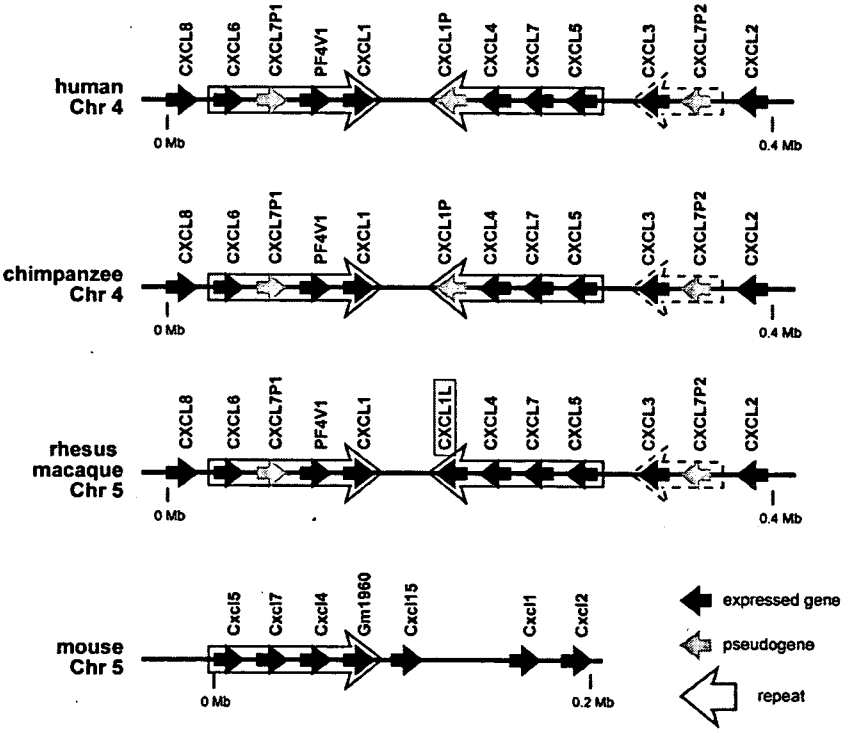


FIG. 3. Genome organization of CXC inflammatory chemokine gene clusters. Genome sequences taken from Ensembl or UCSC were analyzed and are shown schematically. Black and gray arrows denote functionally active gene and pseudogene, respectively, and arrowheads show the transcriptional orientation. Large arrows indicate duplicated regions.

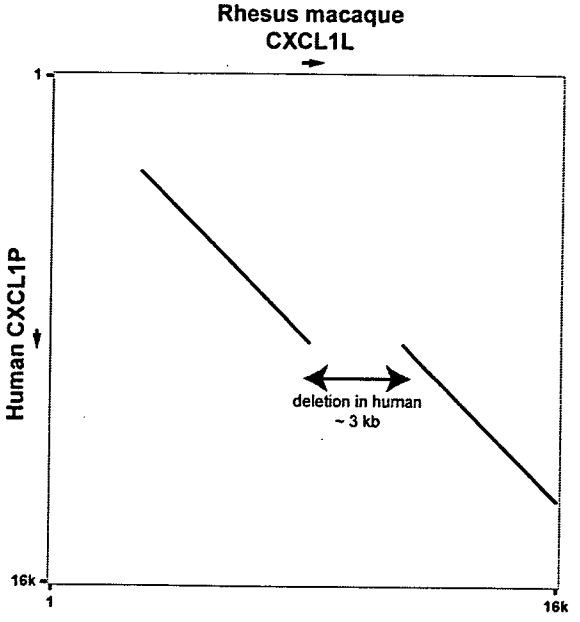


FIG. 4. Dot-plot analysis of the flanking sequences of the rhesus macaque CXCL1L gene and the human CXCL1 pseudogene (CXCL1P). The arrow with two heads shows a deletion of about 3 kb in length seen only in the human genome. The sizes of the genome sequences used were 16 kb for both species.

pair surrounding the coding sequence was prepared. A cDNA clone was isolated by PCR from a cDNA library prepared from cynomolgus macaque spleen. The cDNA encoded a polypeptide of 104 amino acid residues. As the encoded chemokine was highly similar to human CXCL1, CXCL2, and CXCL3, we also isolated cynomolgus macaque CXCL1, CXCL2, and CXCL3 cDNA clones from stomach or liver cDNA libraries for comparison (Fig. 1). The novel chemokine showed 80%, 76%, and 79% similarity to cynomolgus macaque CXCL1, CXCL2, and CXCL3, respectively, and, therefore, was called CXCL1L, from CXCL1-like. The corresponding rhesus macaque CXCL1L was identical to the cynomolgus macaque CXCL1L at the amino acid level but had one synonymous base substitution in the coding region. Interestingly, no orthologous gene for CXCL1L was found in other species, including hominids.

We then examined the expression of CXCL1L in various cynomolgus macaque tissues by semiquantitative RT-PCR. We also examined the expression of CXCL1 for comparison. CXCL1L mRNA was found to be expressed at high levels in spleen and to a lesser extent in brain, bone marrow, and pancreas (Fig. 2). In contrast, CXCL1 mRNA was expressed abundantly in liver, and the expression level in spleen was less than that in liver. Human CXCL1, CXCL2, and CXCL3 genes are likely to be expressed at an extremely low level in spleen according to the expression profiles (*www.ncbi.nlm.nih.gov/UniGene/*), and cynomolgus macaque CXCL1L and CXCL1 may have a unique function in the lymphoid organs.

When we prepared the cynomolgus macaque CXCL3 cDNA from stomach, a cDNA clone closely related to but slightly different from CXCL3 was also isolated. The cDNA encoded a chemokine of 107 amino acid residues that we named CXCL3L (from CXCL3-like) (Fig. 1). CXCL3L was just 5 amino acid residues shorter than CXCL3. Furthermore, although CXCL3L contained 13 base substitutions in the coding region, CXCL3L and CXCL3 differed by only 3 amino acid residues in the signal sequence. No counterpart of the CXCL3L gene was found in rhesus macaque or other primate species.

To obtain clues about the generation mechanism of the CXCL1L gene in macaques, we compared the maps of the inflammatory CXC chemokine gene clusters of human, chimpanzee, rhesus macaque, and mouse obtained from the websites (Ensembl and UCSC) (Fig. 3). The genome sequence of cynomolgus macaque is not available at present. In the primates, a large genome segment including four CXC chemokine genes was duplicated, forming an inverted repeat. Comparison of the maps shows that rhesus macaque CXCL1L gene locus corresponds to those of the CXCL1 pseudogene (CXCL1P) in the human and chimpanzee genomes. The pseudogene in both species contains exons 1 and 2 and the intron between them but lacks the downstream sequence.[11] To examine the genome organization at the nucleotide level, the sequences containing the rhesus macaque CXCL1L and human CXCL1P genes were compared by dot-plot analysis (Fig. 4). Figure 4 clearly shows that a deletion of approximately 3 kb in size is located 3′ downstream of the human CXCL1P gene. This result suggests that the human CXCL1P and rhesus macaque CXCL1L genes arose from a common ancestor gene and that a deletion event in the human genome made the gene inactive after the divergence of hominids and macaques.

Next, a phylogenetic tree was constructed to see the evolutionary relationship of the inflammatory CXC chemokines (Fig. 5). Protein coding sequences were used for the analysis, and the peptide sequences deduced from the human and chimpanzee CXCL1P sequences (exons 1 plus 2) were also included in the analysis for comparison. The tree shows that the macaque CXCL1L and human CXCL1P are indeed closely related, forming a unique branch, whereas the primate CXCL1, CXCL2, and CXCL3 form another relatively independent branch, suggesting the CXCL1L diverged from the common ancestor before CXCL1, CXCL2, and CXCL3 were generated.

Duplication of genes is recognized as the driving force of evolution, and multigene families evolve by a birth-and-death process of duplicated genes.[12] In the CXC chemokine gene cluster, CXCL1 and duplicated CXCL1L in the inverted repeat are both intact in macaques, whereas the latter gene became a pseudogene in hominids. Likewise, the CXCL7 gene in one copy of the repeat turned into a pseudogene CXCL7P1 in the primates (Fig. 3). In addition, another gene duplication appears to have occurred in cynomolgus macaque that generated the CXCL3L gene only in cynomolgus macaque. Given that inactivation of duplicated genes (CXCL1P and CXCL7P1) and emergence of a new gene (CXCL3L) are identified in different primate lineages, the birth-and-death process is still ongoing in the chemokine gene cluster of primates.

As the database search has not identified novel chemokine receptor genes in the macaque genome to date and macaque CXCL1L shows a high sequence similarity to human CXCL1, CXCL2, and CXCL3, the chemokine probably binds the chemokine receptor CXCR2 and has chemotactic activity for neutrophils. CXCL1L indeed contains the Glu-Leu-Arg (ELR)

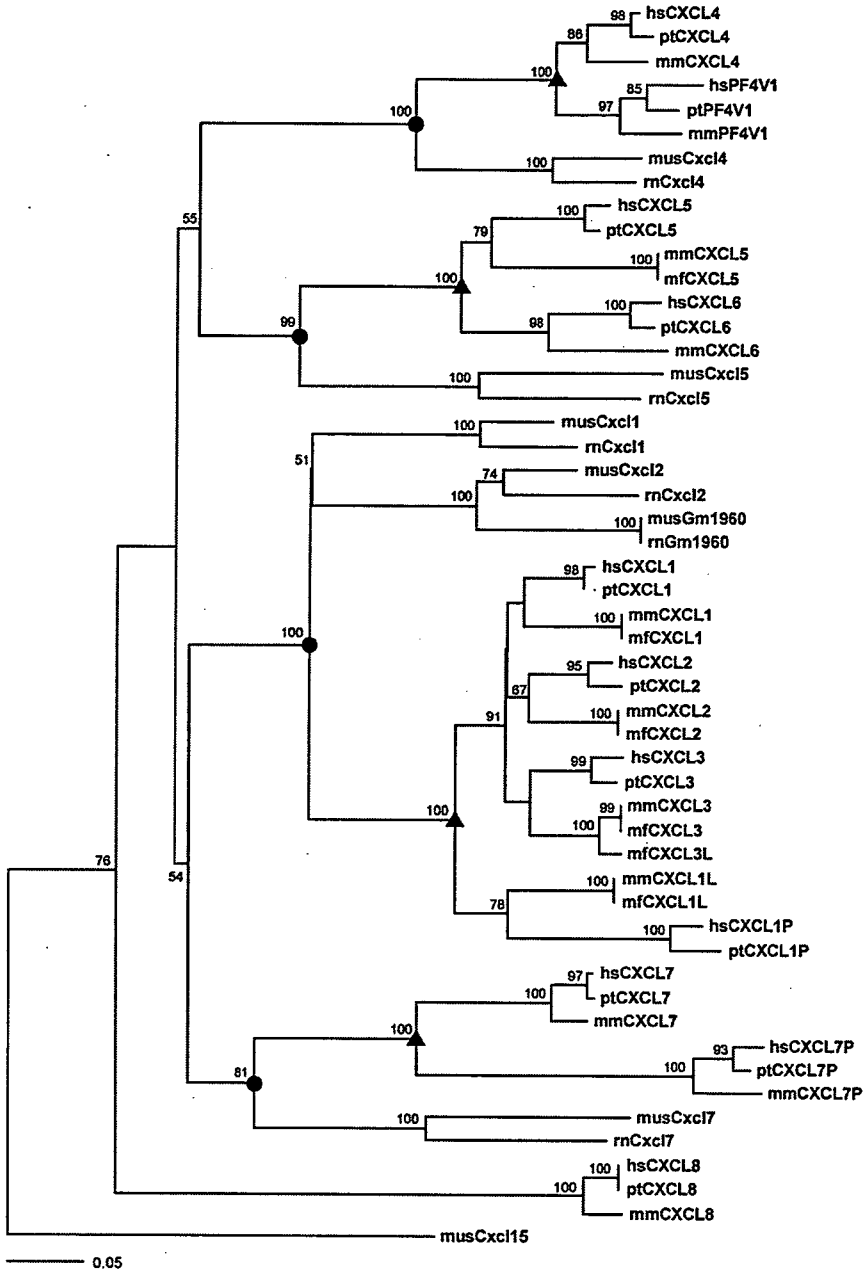


FIG. 5. Phylogenetic tree of CXC inflammatory chemokines. Circles at the nodes indicate divergence points between primates and rodents, and triangles show the points where segmental duplications have occurred in the primate lineage. Chemokine sequences were taken from Ensembl, UCSC, or Cytokine Family Database (*cytokine.medic.kumamoto-u.ac.jp/*). Mouse Cxcl15 was used as an outgroup. hs, human; pt, chimpanzee; mm, rhesus macaque; mf, cynomolgus macaque; mus, mouse; rn, rat.

motif preceding the first conserved cysteine residue that is essential for receptor binding, neutrophil activation, and angiogenic activity.[13,14] However, because gene duplication also allows one copy of the duplicated genes to acquire a new function,[12] we cannot exclude the possibility that CXCL1L might bind other receptors. Expression of CXCL1L at high levels in macaque spleen also suggests a new function in this lymphoid organ (Fig. 2). Recently, some chemokines were shown to be ligands for both chemokine receptors and those not categorized as chemokine receptors.[15,16] Thus, it will be interesting to see if CXCL1L has receptors and functions differently from CXCL1, CXCL2, and CXCL3.

ACKNOWLEDGMENTS

This work is supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by MEXT. HAITEKU (2002–2006).

REFERENCES

1. Jacqz E, Billante C, Moysan F, Mathieu H. The non-human primate: a possible model for human genetically determined polymorphisms in oxidative drug metabolism. *Mol. Pharmacol.* 1988;34:215–217.
2. Thorgeirsson UP, Dalgard DW, Reeves J, Adamson RH. Tumor incidence in a chemical carcinogenesis study of nonhuman primates. *Regul. Toxicol. Pharmacol.* 1994;19:130–151.
3. Daddario-DiCaprio KM, Geisbert TW, Stroher U, Geisbert JB, Grolla A, Fritz EA, Fernando L, Kagan E, Jahrling PB, Hensley LE, Jones SM, Feldmann H. Post-exposure protection against Marburg haemorrhagic fever with recombinant vesicular stomatitis virus vectors in non-human primates: an efficacy assessment. *Lancet* 2006;367:1399–1404.
4. Misumi S, Nakayama D, Kusaba M, Iiboshi T, Mukai R, Tachibana K, Nakasone T, Umeda M, Shibata H, Endo M, Takamune N, Shoji S. Effects of immunization with CCR5-based cycloimmunogen on simian/HIVSF162P3 challenge. *J. Immunol.* 2006;176:463–471.
5. Moser B, Loetscher P. Lymphocyte traffic control by chemokines. *Nat. Immunol.* 2001;2:123–128.
6. Zlotnik A, Yoshie O. Chemokines: a new classification system and their role in immunity. *Immunity* 2000;12:121–127.
7. Yoshie O, Imai T, Nomiyama H. Chemokines in immunity. *Adv. Immunol.* 2001;78:57–110.
8. Nomiyama H, Mera A, Ohneda O, Miura R, Suda T, Yoshie O. Organization of the chemokine genes in the human and mouse major clusters of CC and CXC chemokines: diversification between the two species. *Genes Immun.* 2001;2:110–113.
9. Nomiyama H, Egami K, Tanase S, Miura R, Hirakawa H, Kuhara S, Ogasawara J, Morishita S, Yoshie O, Kusuda J, Hashimoto K. Comparative DNA sequence analysis of mouse and human CC chemokine gene clusters. *J. Interferon Cytokine Res.* 2003; 23:37–45.
10. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ, Weber RJ, Haussler D, Kent WJ. The UCSC Genome Browser Database. *Nucleic Acids Res.* 2003;31:51–54.
11. Shattuck-Brandt RL, Wood LD, Richmond A. Identification and characterization of an MGSA/GRO pseudogene. *DNA Seq.* 1997;7: 379–386.
12. Nei M, Rooney AP. Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet.* 2005;39:121–152.
13. Clark-Lewis I, Dewald B, Geiser T, Moser B, Baggiolini M. Platelet factor 4 binds to interleukin 8 receptors and activates neutrophils when its N terminus is modified with Glu-Leu-Arg. *Proc. Natl. Acad. Sci. USA* 1993;90:3574–3577.
14. Strieter RM, Burdick MD, Gomperts BN, Belperio JA, Keane MP. CXC chemokines in angiogenesis. *Cytokine Growth Factor Rev.* 2005;16:593–609.
15. Elagoz A, Henderson D, Babu PS, Salter S, Grahames C, Bowers L, Roy MO, Laplante P, Grazzini E, Ahmad S, Lembo PM. A truncated form of CKbeta8-1 is a potent agonist for human formyl peptide-receptor-like 1 receptor. *Br. J. Pharmacol.* 2004;141:37–46.
16. Nakayama T, Kato Y, Hieshima K, Nagakubo D, Kunori Y, Fujisawa T, Yoshie O. Liver-expressed chemokine/CC chemokine ligand 16 attracts eosinophils by interacting with histamine H4 receptor. *J. Immunol.* 2004;173:2078–2083.

Address reprint requests or correspondence to:
Dr. Hisayuki Nomiyama
Department of Molecular Enzymology
Kumamoto University Graduate School of Medical Sciences
Honjo 1-1-1
Kumamoto 860-8556
Japan

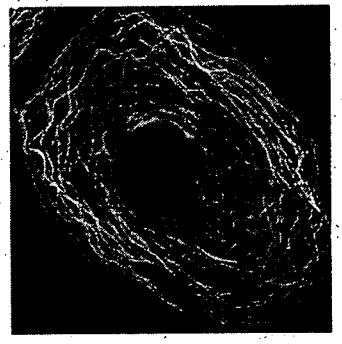
Tel: +81-96-373-5065
Fax: +81-96-373-5066
E-mail: nomiyama@gpo.kumamoto-u.ac.jp

Received 15 June 2006/Accepted 13 July 2006

FEBS Letters

The journal for rapid publication of short reports in molecular biosciences

- Spotlight on...Hans Eklund
- Non-canonical connexin Cx23 forms channels in zebrafish
- Nucleoporin Nup358 regulates interphase microtubules



Published by Elsevier on behalf of the Federation of European Biochemical Societies

Submit to: http://www.ees.elsevier.com/febsletters

ISSN 0014 5793
Volume 582 Number 2 23 January 2008

FEBS Letters 582 (2008) 351–358

Expressed sequence tags from cynomolgus monkey (*Macaca fascicularis*) liver: A systematic identification of drug-metabolizing enzymes

Yasuhiro Uno[a,*], Yutaka Suzuki[b,*], Hiroyuki Wakaguri[b], Yoshiko Sakamoto[a], Hitomi Sano[a], Naoki Osada[c], Katsuyuki Hashimoto[c], Sumio Sugano[b], Ituro Inoue[a,d]

[a] *Division of Genetic Diagnosis, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan*
[b] *Department of Medical Genome Sciences, Graduate School of Frontier Sciences, University of Tokyo, 4-6-1 Shirokanedai, Minatoku, Tokyo 108-8639, Japan*
[c] *Department of Biomedical Resources, National Institute of Biomedical Innovation, Ibaraki, Osaka, Japan*
[d] *Division of Molecular Life Science, School of Medicine, Tokai University, Shimokasuya 134, Isehara, Kanagawa 259-1193, Japan*

Received 1 October 2007; revised 14 December 2007; accepted 18 December 2007

Available online 31 December 2007

Edited by Takashi Gojobori

Abstract The liver, a major organ for drug metabolism, is physiologically similar between monkeys and humans. However, the paucity of identified genes has hampered a deep understanding of drug metabolism in monkeys. To provide such a genetic resource, 28655 expressed sequence tags (ESTs) were generated from a cynomolgus monkey liver full-length enriched cDNA library, which contained 23 unique ESTs homologous to human drug-metabolizing enzymes. Our comparative genomics approach identified nine lineage-specific candidate ESTs, including three drug-metabolizing enzymes, which could be important for understanding the physiological differences between monkeys and humans.

Keywords: Cynomolgus monkey; Drug metabolism; Drug-metabolizing enzyme; Expressed sequence tags; Lineage-specific gene; Liver

1. Introduction

Cynomolgus monkeys have been used as an animal model for the investigation of human physiology and disease because of their close genetic and physiological similarities to humans. Application of this animal model includes predicting metabolic fate of newly developed drugs due to pharmacokinetics similar to humans. However, we now know that differences in metabolic properties are occasionally seen for some drugs between monkeys and humans [1–7] possibly due to differences in genetic components essential for drug metabolism between the two lineages, such as lineage-specific genes and alternatively spliced transcripts. However, limited numbers of lineage-specific genes identified in monkeys have hampered complete knowledge of lineage differences in drug metabolism.

An expressed sequence tag (EST)-sequencing approach has been a rapid and efficient way to identify novel cDNAs that provide a basis to investigate genetic components essential to various physiological functions. In non-human primates, efforts have been made for the comprehensive identification of ESTs in chimpanzees [8], rhesus monkeys [9,10], and cynomolgus monkeys [11–13]. However, liver tissue has not been extensively sequenced for ESTs, thus only limited genetic information is available on liver physiological function such as drug metabolism. With the completion of a draft of the rhesus monkey genome sequence [14], EST analysis of macaques should be more feasible and accurate.

To provide a monkey genetic resource, 28655 ESTs from cynomolgus monkey liver were generated. These macaque ESTs analyzed against the rhesus genome identified 1064 unique ESTs, most of which (77.0%) matched the human RefSeq database. cDNAs highly homologous to human drug-metabolizing genes were identified, including those of cytochrome P450 (CYP), UDP-glucuronosyltransferase (UGT), glutathione *S*-transferase (GST), sulfotransferase (SULT), and flavin-containing monooxygenase (FMO). Moreover, our method to select lineage-specific ESTs successfully identified novel transcripts related to drug metabolism. This genetic information should help in discerning various physiological characteristics, including drug metabolism in monkeys.

2. Materials and methods

2.1. cDNA library construction and EST sequencing

Liver samples were collected from three adult cynomolgus monkeys (two males and one female) and used to generate a full-length enriched cDNA library using the pME18S-FL3 vector by the oligo-capping method as previously described [15]. Purified DNA was sequenced using the ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit, Version 2.0 (Applied Biosystems, Foster City, CA), followed by electrophoresis with ABI-3700 DNA Analyzer (Applied Biosystems) according to the manufacturer's instructions. Primers (5′-GGATGTTGCCTTTACTTCTA-3′ and 5′-TTTTTTTTTTTTTTTTV-3′) were used for single-pass sequencing of 5′ and 3′-ends for each cDNA, respectively.

*Corresponding authors.
E-mail addresses: unoxx001@pharm.hokudai.ac.jp (Y. Uno), ysuzuki@k.u-tokyo.ac.jp (Y. Suzuki).

Abbreviations: CYP, cytochrome P450; EST, expressed sequence tag; FMO, flavin-containing monooxygenase; GST, glutathione *S*-transferase; ORF, open reading frame; SULT, sulfotransferase; UGT, UDP-glucuronosyltransferase

0014-5793/$32.00
doi:10.1016/j.febslet.2007.12.031

2.2. Sequence data analysis

Vector sequence was trimmed and sequence quality was inspected using Phred (University of Washington). Only EST sequences longer than 200 bases were used. Generated EST sequences were first computationally mapped to the *Macaca mulatta* genomic sequence (rheMac2, UCSC Genome Browser). Computational mapping was carried out as previously described by consequential use of sequence alignment programs, BLAT and SIM4 [16]. Only ESTs over the entire sequence length that mapped perfectly at unique positions on the macaque genome were regarded as "mapped". Further information for each cDNA is presented in our database, DBTSS (http://dbtss.hgc.jp), and a user manual has been published [16].

The macaque genomic sequences to which our cynomolgus ESTs mapped were examined for any corresponding human genomic and RefSeq sequence. If any, the corresponding macaque EST was correlated with the human RefSeq gene. Based on information from the correlated human RefSeq gene, GO (Gene Ontology) classification was carried out for macaque ESTs using GO slim (http://www.geneontology.org/) for "Biological Process", Molecular Function", and "Cellular Component".

2.3. Identification of putative macaque-specific transcripts

To identify macaque ESTs that do not match to human genes, the ESTs were analyzed by either a genome- or cDNA-based approach. In the genome-based approach, we selected the EST sequence located outside human-macaque alignable regions according to the genome–genome alignment in the UCSC Genome Browser. In the cDNA-based approach, ESTs were first searched with the human RefSeq database using BLASTN (cut-off = 1.0e−100). Those ESTs with no hits were clustered with each other (cut-off = 0.0; >98% identity) and clusters containing more than 10 ESTs were selected. Those clustered cDNAs were searched against the human RefSeq database again (1.0e−50), and the generated sequence alignments were further manually inspected. For the macaque-specific transcript candidates, complete sequences were determined by primer-walking.

3. Results and discussion

3.1. Sequencing and clustering of macaque liver ESTs

A full-length cDNA library was constructed from cynomolgus monkey (*Macaca fascicularis*) livers using the oligo-capping method [15]. One-pass sequencing at 5′ and 3′-ends of the liver cDNA clones and sequence processing generated a total of 28 655 high quality ESTs (deposited in GenBank under Accession Nos. BB873801–BB902455). Only 3′ ESTs (27 959 entries) were further analyzed. Of these ESTs, 14 727 (53%) were successfully mapped to 1064 different regions in the rhesus macaque genome. Of the 1064 regions, 819 (77%) reside in genomic regions highly homologous to human RefSeq genes as revealed by a genome–genome comparison, and were annotated with human RefSeq genes (Table 1). Clustering of 27 959 ESTs was carried out by calculating the number of ESTs that mapped to the same region, which should represent the cluster size for the corresponding gene. This analysis for the 1064 mapped regions in the genome indicated that these 1064 unique ESTs consisted of 525 contigs (49.3%) and 539 singletons (50.7%). The number of members in each cluster ranged up to 4354, with a 26.9 average. The gene expression profile based on our EST data reflected liver functional characteristics because the most abundantly expressed genes were hepatocyte-specific markers, such as albumin, fibrinogen gamma and beta polypeptides, haptoglobin, and alcohol dehydrogenase, all of which comprised more than half of the identified ESTs (Table 1). Such high redundancy of hepatic ESTs from the non-normalized cDNA libraries has been also seen for human libraries [17–19].

3.2. Functional classification of ESTs

Provisional functional classification was carried out using GO slim terms based on the human RefSeq genes that correlated with our macaque ESTs (Fig. 1). Out of 819 unique ESTs that matched a human RefSeq entry, 786 (96.0%) were assigned to at least one main category; Biological Process, Molecular Function, and Cellular Component, to which 520, 458, and 373 sequences (48.9%, 43.0%, and 35.1%) were classified, respectively. Sequences from 133 ESTs (16.2%) were annotated into all three categories. The largest EST groups include metabolism, transcription, protein biosynthesis, electron transport, transport, signal transduction, and lipid metabolism (Fig. 1A) for Biological Process, and protein binding, transferase activity, and nucleotide binding for Molecular Function (Fig. 1B).

3.3. ESTs relevant to drug metabolism

Our major objective was identification of cDNAs important for drug metabolism, namely those encoding drug-metabolizing enzymes, which belong to CYP, UGT, GST, SULT, and FMO families. The 446 ESTs in 23 clusters were highly homologous to genes for such drug-metabolizing enzymes in humans (Table 2). For the CYP family, only ESTs for the CYP1 to CYP4 subfamilies are indicated in the list because of their importance in drug metabolism. The CYP family contained 231 entries (51.8%), the largest group among the ESTs for drug-metabolizing enzymes. CYP, a phase I drug-metabolizing enzyme, is involved in hydroxylation of a large number of

Table 1
Genes abundantly expressed in liver (>100 reads)

Contig number	ESTs	Human RefSeq ID	Annotation
15043	4354	NM_000477	Albumin
15223	3763	NM_000509	Fibrinogen gamma chain
15577	1097	NM_005141	Fibrinogen beta chain
10429	277	NM_005143	Haptoglobin
19271	253	NM_000668	Alcohol dehydrogenase IB (class I), beta polypeptide
17733	245	NM_001085	Serpin peptidase inhibitor, clade A, member 3
5070	227	NM_000035	Aldolase B, fructose-bisphosphate
6405	158	NM_016413	Carboxypeptidase B2 (plasma, carboxypeptidase U)
5438	154	NM_000638	Vitronectin
11141	130	NM_000354	Serpin peptidase inhibitor, clade A, member 7
12017	125	NM_001622	Alpha-2-HS-glycoprotein
15222	109	NM_000508	Fibrinogen alpha chain
17083	106	NM_001756	Serpin peptidase inhibitor, clade A, member 6

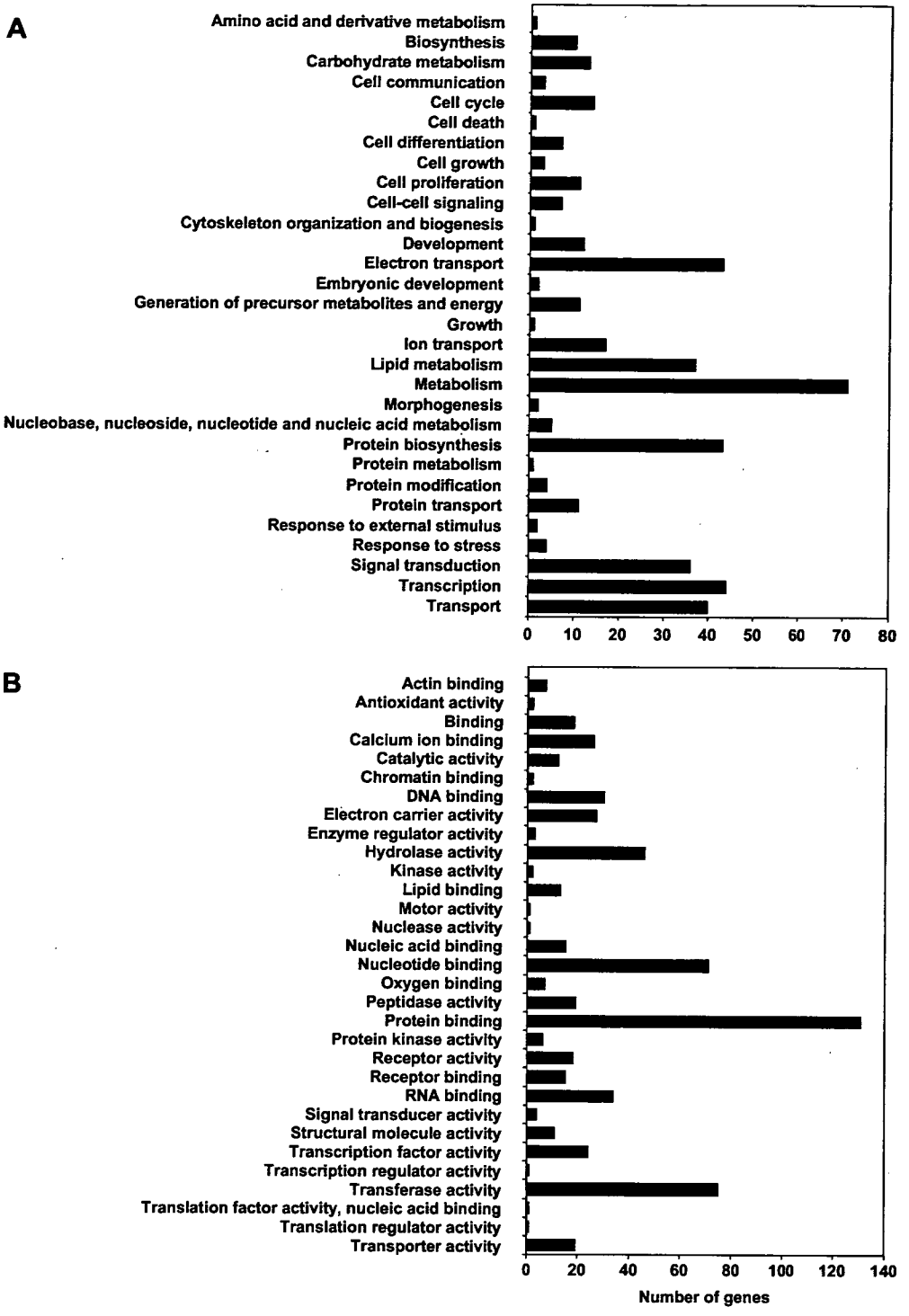


Fig. 1. Functional classification of cynomolgus liver ESTs. All non-redundant ESTs were assigned to each functional category as described in Section 2. Biological process (A) and Molecular function (B) are shown.

drugs [20]. Among the CYP ESTs identified, 124 (53.7%) belonged to the CYP2C subfamily that is important for metabolism of ~20% of all prescribed drugs such as tolbutamide, phenytoin, and warfarin [21]. Fifty-two ESTs belonged to the CYP3A subfamily. In humans, genes in the CYP3A subfamily (especially *CYP3A4*) are essential for drug metabolism, and are involved in the metabolism of more than half the currently prescribed drugs. Moreover, human CYP3A4 and CYP3A5 occupy more than half of the total CYP protein content in liver [20], contributing substantially overall drug

Table 2
Cynomolgus ESTs highly homologous to human drug-metabolizing enzyme families, CYP, UGT, GST, SULT, and FMO

Family	Contig number	Number of ESTs	Accession number	Matched human cDNA
CYP	18729	89	NM_0007700	CYP2C8
	12912	38	NM_017460	CYP3A4
	463	30	NM_000106	CYP2D6
	19109	21	NM_000769	CYP2C19
	19111	14	NM_000771	CYP2C9
	12910	14	NM_000777	CYP3A5
	19262	13	NM_000773	CYP2E1
	7671	8	NM_023944	CYP4F12
	8451	3	NM_000775	CYP2J2
	8410	1	NM_000778	CYP4A11
UGT	15423	75	NM_001074	UGT2B7
	2531	41	NM_019093	UGT1A3
	15424	30	NM_050394	UGT2B28
GST	19270	9	NM_145740	GSTA1
	14165	8	NM_000846	GSTA2
	9619	7	NM_000851	GSTM5
	1616	7	NM_145792	MGST1
	19167	3	NM_004832	GSTO1
	17697	2	NM_145870	GSTZ1
SULT	10631	10	NM_001055	SULT1A1
	19266	2	NM_001054	SULT1A2
	3203	1	NM_006588	SULT1C2
FMO	10051	20	NM_001002294	FMO3

metabolism in humans. Thirty ESTs were highly similar to human CYP2D6. In the human genome, three *CYP2D* genes are present including one functional *CYP2D* gene (*CYP2D6*) and two pseudogenes (*CYP2D7* and *CYP2D8*). CYP2D6 accounts for 5% of the total hepatic CYP content and is responsible for the metabolism of 25% of all drugs oxidized by CYPs [20]. In cynomolgus monkeys, CYP2D17, which is highly homologous to human CYP2D6, has been isolated [22]. Meanwhile, marmoset is known to have two functional CYP2Ds with different metabolic properties, CYP2D19 and CYP2D30 [23]. Further in-depth analysis of our EST clones could reveal whether *CYP2D17* is the only *CYP2D* gene expressed in cynomolgus monkey liver. Characterization of these CYP EST clones is currently in progress, such as full-length sequencing, tissue expression patterns, and metabolic assays, the outcome of which has been partly published [24,25].

Clusters for other drug-metabolizing enzymes of UGT, GST, SULT, and FMO families contained 146, 36, 13, and 20 ESTs, respectively (Table 2). UGT, a phase II drug-metabolizing enzyme, catalyzes the conjugation of various drugs to assist drug excretion and is composed of UGT1A, UGT2A, and UGT2B subfamilies in humans. The 146 ESTs for the UGT family were grouped into three clusters. Forty-one EST sequences were highly homologous to human UGT1A3. The *UGT1A* gene locus contains 13 distinct first exons (promoters) followed by exons 2–5 that are shared among all 13 transcripts, giving rise to nine different proteins (four pseudogenes) in humans [26]. Considering that these ESTs were 3′ cDNAs, the 41 EST clones possibly encode multiple *UGT1A* genes. Because only four *UGT1A* genes have been identified for macaques, further sequence analysis of these UGT1A ESTs could lead to the identification of novel *UGT1A* genes in this lineage. Seventy-five and 30 ESTs matched human UGT2B7 and UGT2B28, respectively. Initial full-length sequencing of the UGT2B EST clones revealed that the clones contained the UGT2B33 cDNA (GenBank Accession No. AB371703) newly identified in cynomolgus monkeys as well as the previously identified cDNAs for cynomolgus UGT2B9, UGT2B18, UGT2B19, UGT2B20, UGT2B23, and UGT2B30 (GenBank Accession Nos. U91582, AF016310, AF112112, AF072223, AF112113, and AF401657, respectively). In contrast to *UGT1A*, *UGT2B* genes have been frequently duplicated in many mammalian lineages [26]; therefore, some of these *UGT2Bs* are possibly lineage-specific genes as discussed below.

Other EST sequences were highly homologous to six human genes, two genes, and one gene in the GST, SULT, and FMO families, respectively (Table 2). GST is another phase II enzyme, catalyzing the conjugation of electrophilic substrates to glutathione and is composed of at least 16 genes for cytosolic, mitochondrial, and microsomal GSTs in humans [27]. SULT is also a gene family comprising at least 10 human genes, catalyzing sulfate conjugation of a wide variety of drugs [28]. FMO is a family of flavoproteins, catalyzing oxygenation of various drugs containing sulfur, nucleophilic nitrogen, and phosphorus heteroatoms [29]. Full-length sequencing and functional characterization of these EST clones are currently under investigation, by which novel genes could be identified because limited numbers of genes have been identified for these enzymes in monkeys. These results suggest that our EST-sequencing approach successfully identified a number of cDNA clones for various drug-metabolizing enzymes in monkeys.

3.4. Identification of lineage-specific genes

In order to better utilize monkeys as an animal model, it is essential to understand similarities or differences in genes expressed between monkeys and humans. The EST data should provide essential information on lineage-specific genes and transcripts. To identify macaque-specific transcripts, 27959 3′ ESTs were analyzed by either a genome- or cDNA-based approach. In the genome-based approach, we found 77 EST clusters, for which at least a part of the sequences were located outside human-macaque alignable regions. In the cDNA-based approach, we identified 12 clusters containing >10 ESTs that were unmatched to any human RefSeq genes according to BLASTN (cut-off = 1.0e−100). Clones available for the 10 remaining candidate clusters after subsequent manual inspection, along with clones for the 29 clusters randomly selected from 77 candidates in the genome-based approach, were subjected to full-length sequencing (excluding 1 overlapping clone). Sequence analysis of these 38 clones confirmed that nine clones contained lineage-specific candidate genes. Of these, six clones matched to human RefSeq sequences (Table 3): two clones lack a portion of human genome sequence and the other four matched to more than one member of a gene family. Thus, these four clones were potentially lineage-specific genes and were further characterized as described below.

One candidate clone (Qlv-U097A-G10) encoded CYP2C76 with a relatively low homology (~80%) to members of the human CYP2C subfamily, CYP2C8, CYP2C9, CYP2C18, and CYP2C19 (Table 3). The extent of homology was much lower than those for other ESTs (~95%). Our characterization

Table 3
Potential lineage-specific ESTs in cynomolgus monkey

Clone ID	GenBank Accession number	Nucleotide (bp)	ORF[a] (Number of amino acids)	Cynomolgus sequence	The most highly homologous human RefSeq cDNAs	Genome- or cDNA-based approach	Aligned location
Novel member of gene family							
Qlv-U042A-F11	AB362497	1637	454	None	CFH, CFHR3/4	Genome/cDNA	Intergenic
Qlv-U097A-G10	AB362507	1986	489	CYP2C76	CYP2C8/9/18/19	cDNA	Intergenic
Qlv-U346A-B11[b]	AB371605	1758	472	CYP2A23	CYP2A6/7/13	cDNA	Intergenic
Qlv-U405A-G11	AB362508	2225	528	UGT2B19	UGT2B4	cDNA	Intergenic
Partially unmatched to human genome							
Qlv-U244A-C6[b]	AB362499	1612	305	None	TSPAN12	Genome	Intergenic
Qlv-U258A-D7[b]	AB362500	1984	89	None	SS18L1	Genome	Intergenic
Unmatched to human genome							
Qlv-U050A-D10	AB362503	1700	34	None	None	Genome	Intron
Qlv-U295A-A3	AB362504	2278	118	None	None	Genome	Intergenic
Qlv-U389A-C1	AB362506	2043	90	None	None	Genome	Intergenic

[a]The longest ORF was selected.
[b]Transcript variants with different exon-intron structure from human homologs.

of CYP2C76 (GenBank Accession No. DQ074807) at the RNA, protein, and genomic level revealed that this CYP2C did not have any human ortholog because the corresponding genes were not found in the human genome [24]. Moreover, this CYP2C76 was at least partly responsible for lineage differences in drug metabolism [30]. These results confirmed that our comparative genomic approach succeeded in identifying macaque-specific transcripts that are absent in humans.

One clone (Qlv-U405A-G11) identified as a lineage-specific candidate contained the cDNA for UGT2B19 previously reported [31]. Cynomolgus UGT2B19 as well as UGT2B30 cDNAs were both highly homologous (92%) to human UGT2B4 cDNA [32]. A phylogenetic comparison (Fig. 2) indicated that the 1-to-1 orthologous relationship to the human UGT2Bs could not be determined for these cynomolgus UGT2Bs, raising the possibility that *UGT2B19* might be a lineage-specific gene. *UGT2B19* is expressed in cynomolgus monkey liver and prostate and has enzymatic activity to xenobiotics (1-naphthol) and steroids (testosterone) [31]. The UGT2B subfamily consisted of a number of member genes including a lineage-specific candidate [26], suggesting that UGT2B19 and other functional UGT2B enzymes in cynomolgus monkeys contribute not only to overall drug metabolism in monkey liver but also possibly to differences in drug metabolism.

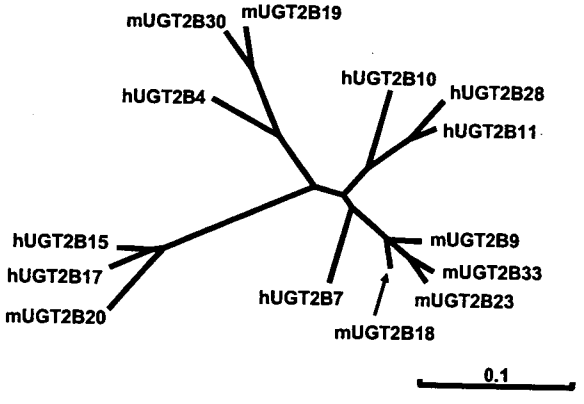


Fig. 2. A phylogenetic comparison of UGTs between macaque and human. The phylogenetic tree was based on amino acid sequence using the Clustal W program. Deduced amino acid sequences were used for cynomolgus monkeys (m) and human (h).

Another lineage-specific candidate clone (Qlv-U346A-B11) was cynomolgus CYP2A23 variant (tentatively named CYP2A23v), containing exons 1–8 with a partial intron 8 sequence and thus, lacking the entire exon 9 as compared to a complete CYP2A23 transcript. CYP2A23 and another cynomolgus CYP2A, CYP2A24, were both highly homologous (~95%) to the three human CYP2As, specifically CYP2A6, CYP2A7, and CYP2A13, indicating the difficulty in determining the orthologous relationship of CYP2A23 and CYP2A24 to human CYP2As [25]. This novel CYP2A23 transcript variant encodes a protein of 472 amino acids and lacks a part of a heme-binding region essential for CYP proteins (Fig. 3). The protein generated from this transcript, therefore, might not function as a drug-metabolizing enzyme. A similar transcript variant was also identified for CYP2C76 and UGT2B19 (data not shown). It remains to be determined whether the presence of these transcripts lacking a functional domain is limited to the animals that provided liver samples for the cDNA library construction and what roles these transcript variants play in drug metabolism.

Other than those for drug metabolism, one lineage-specific candidate (Qlv-U042A-F11) had high sequence homology to complement factor H (CFH) family genes in humans (Table 3). CFH (also called Factor H), an important complement regulator, forms a gene family along with CFH-related proteins (CFHL1-5) in humans [33]. This macaque transcript contained an open reading frame (ORF) of 454 amino acids. CFH and other genes important for immune response and T cell-mediated immunity such as *immunoglobulin-like* genes and MHC-related genes have been identified in macaques as the genes that went under positive selection [13,14,34], and thus, our finding of lineage-specific CFH-like sequence in macaques is not surprising. Further analysis of this CFH-like sequence indicated that the first 19 amino acids and the remaining amino acids were highly similar to CFH-related proteins (CFHR3 and CFHR4) and CFH in humans, respectively (data not shown), raising the possibility that this novel transcript might be a hybrid of CFH and CFH-related genes. In humans, a hybrid transcript of CFH and CFHR1 has been identified and implicated in atypical haemolytic uraemic syndrome [35].

```
mCYP2A23v   1:MLASGLLLVALLACLTVMVLMSVWQQRNSKGKLPPGPTPLPFIGNYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVLCGYDAVKEALVDQAEEFSG 100
mCYP2A23    1:MLASGLLLVALLACLTVMVLMSVWQQRNSKGKLPPGPTPLPFIGNYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVLCGYDAVKEALVDQAEEFSG 100
mCYP2A24    1:MLASGLLLVALLACLTVMVLMSVWQQRNSKGKLPPGPTPLPFIGNYLQLNTEQMCNSIMKISERYGPVFTIHLGPRRVVVLCGYDAVKEALVDQAEEFSG 100
hCYP2A6     1:MLASGMLLVALLVCLTVMVLMSVWQQRKSKGKLPPGPTPLPFIGNYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVLCGHDAVREALVDQAEEFSG 100
hCYP2A7     1:MLASGLLLVALLACLTVMVLMSVWQQRKSRGKLPPGPTPLPFIGNYLQLNTEHICDSIMKFSECYGPVFTIHLGPRRVVVLCGHDAVREALVDQAEEFSG 100
hCYP2A13    1:MLASGLLLVTLLACLTVMVLMSVWRQRKSRGKLPPGPTPLPFIGNYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVLCGHDAVKEALVDQAEEFSG 100
              ******.***.**.**********.** *.**********************....*.**.**.****************** ***.**********

mCYP2A23v 101:RGEQATFDWLFKGYGVVFSNGERAKQLRRFSIATLRDFGVGKRGIEERIQEEAGFLIEALRDTQGANIDPTFFLSRTVSNVISSIVFGDRFDYEDKEFLS 200
mCYP2A23  101:RGEQATFDWLFKGYGVVFSNGERAKQLRRFSIATLRDFGVGKRGIEERIQEEAGFLIEALRDTQGANIDPTFFLSRTVSNVISSIVFGDRFDYEDKEFLS 200
mCYP2A24  101:RGEQATFDWVFKGYGVVFSNGERAKQLRRFSIATLRDFGVGKRGIEERIQEEAGFLIDALRGTGGANIDPTFFLSRTVSNVISSIVFGDRFDYKDKEFLS 200
hCYP2A6   101:RGEQATFDWVFKGYGVVFSNGERAKQLRRFSIATLRDFGVGKRGIEERIQEEAGFLIEALRDTHGANIDPTFFLSRTVSNVISSIVFGDRFDYKDKEFLS 200
hCYP2A7   101:RGEQATFDWVFKGYGVAFSNGERAKQLLRFAIATLRDFGVGKRGIEERIQEESGFLIEAIRSTHGANIDPTFFLSRTVSNVISSIVFGDRFDYEDKEFLS 200
hCYP2A13  101:RGEQATFDWLFKGYGVAFSNGERAKQLRRFSIATLRGFGVGKRGIEERIQEEAGFLIDALRGTHGANIDPTFFLSRTVSNVISSIVFGDRFDYEDKEFLS 200
              ********* ******.**********.**.*****.***************.****.*.* * ******************************.******

mCYP2A23v 201:LLRMMLGSFQFTATSAGQLYEMFSSVMKHLPGPQQQAFKELQGLEDFIAKKVEHNRRTLDPNSPRDFIDSFLIRMQEEEKNPNTEFHLKNLVLTSLNLFF 300
mCYP2A23  201:LLRMMLGSFQFTATSAGQLYEMFSSVMKHLPGPQQQAFKELQGLEDFIAKKVEHNRRTLDPNSPRDFIDSFLIRMQEEEKNPNTEFHLKNLVLTSLNLFF 300
mCYP2A24  201:LLGMMLAIFQFTSTSTGQLYEMFSSVMKHLPGPQQQAFKELQGLEDFIAKKVEHNQRTLDPNSPRDFIDSFLIRMQEEEKNPNTEFYLKNLMMTTLNLFI 300
hCYP2A6   201:LLRMMLGIFQFTSTSTGQLYEMFSSVMKHLPGPQQQAFQLLQGLEDFIAKKVEHNQRTLDPNSPRDFIDSFLIRMQEEEKNPNTEFYLKNLVMTTLNLFI 300
hCYP2A7   201:LLSMMLGIFQFTSTSTGQLYEMFSSVMKHLPGPQQQAFKLLQGLEDFIAKKVEHNQRTLDPNSPQDFIDSFLIHMQEEEKNPNTEFYLKNLMMSTLNLFI 300
hCYP2A13  201:LLRMMLGSFQFTATSTGQLYEMFSSVMKHLPGPQQQAFKELQGLEDFIAKKVEHNQRTLDPNSPRDFIDSFLIRMQEEEKNPNTEFYLKNLVMTTLNLFF 300
              **.***. **** **.**********************..***************.********.********.************.****....****

mCYP2A23v 301:GGTETVSTTLRYGFLLLMKHPEVEAKVHEEIDRVIGKNRQPKFEDWAKMPYTEAVIHEIQRFGDMLPFGVAHRVIKDTKFRDFFLPKGTEVFPMLGSVLK 400
mCYP2A23  301:GGTETVSTTLRYGFLLLMKHPEVEAKVHEEIDRVIGKNRQPKFEDWAKMPYTEAVIHEIQRFGDMLPFGVAHRVIKDTKFRDFFLPKGTEVFPMLGSVLK 400
mCYP2A24  301:AGTETVSTTLRYGFLLLMKYPEVEAKVHEEIDRVIGKNRQPKFEDRVKMPYMEAVIHEIQRFGDVIPMSLARRVNKDTKFRDFFLPKGTEVFPMLGSVLR 400
hCYP2A6   301:GGTETVSTTLRYGFLLLMKHPEVEAKVHEEIDRVIGKNRQPKFEDRAKMPYMEAVIHEIQRFGDVIPMSLARRVKKDTKFRDFFLPKGTEVFPMLGSVLR 400
hCYP2A7   301:AGTETVSTTLRYGFLLLMKHPEVEAKVHEEIDRVIGKNRQPKFEDRTKMPYMEAVIHEIQRFGDVIPMSLARRVKKDTKFRDFFLPKGTEVFPMLGSVLR 400
hCYP2A13  301:AGTETVSTTLRYGFLLLMKHPEVEAKVHEEIDRVIGKNRQPKFEDRAKMPYTEAVIHEIQRFGDMLPMGLAHRVNKDTKFRDFFLPKGTEVFPMLGSVLR 400
              ******************.**************************..**** ************  *. .* ** ************************.

mCYP2A23v 401:DPKFFSNPQDFNPQHFLDEKGQFKKSDAFVPFSIGKRPLFAARPLLTPAGASLTQLPSVLCSLVFLQLGSSC                          472
mCYP2A23  401:DPKFFSNPQDFNPQHFLDEKGQFKKSDAFVPFSIGKRNCFGEGLARMELFLFLTTIMQNFRFKSPQSPKDIDVSPKHMGFATIPPNYTMSFLPR  494
mCYP2A24  401:DPRFFSNPQDFNPQHFLDEKGQFKKSDAFVPFSIGKRNCFGEGLARMELFLFFTTIMQNFRFKSPQSPKDIDVSPKHAGFATIPRNYTMSFLPR  494
hCYP2A6   401:DPSFFSNPQDFNPQHFLNEKGQFKKSDAFVPFSIGKRNCFGEGLARMELFLFFTTVMQNFRLKSSQSPKDIDVSPKHVGFATIPRNYTMSFLPR  494
hCYP2A7   401:DPSFFSNPQDFNPQHFLDDKGQFKKSDAFVPFSIGKRNCFGEGLARMELFLFFTTVMQNFRLKSSQSPKDIDVSPKHVVFATIPRNYTMSFLPR  494
hCYP2A13  401:DPRFFSNPRDFNPQHFLDKKGQFKKSDAFVPFSIGKRYCFGEGLARMELFLFFTTIMQNFRFKSPQSPKDIDVSPKHVGFATIPRNYTMSFLPR  494
              ** *****.********..**********.*.*.....*. ..... .. .......
```

Fig. 3. Alignment of the amino acid sequences deduced from cynomolgus monkey (m) and human (h) CYP2A cDNAs. The sequences were aligned using the Clustal W program. Asterisks and dots under the sequences indicate identical amino acids and conservatively unchanged amino acids, respectively. A black line under the amino acid sequences indicates the putative heme-binding region. The CYP2A23 variant (mCYP2A23v) newly identified lacks half of the putative heme-binding region.

These results suggest that our approach of lineage-specific gene identification successfully identified potential lineage-specific genes or transcripts, possibly relevant to the immune system. Further investigation of other ESTs should help make better use of the macaque for immunological studies.

Three candidate genes were unmatched to any human RefSeq sequence, and thus could be apparent lineage-specific genes (Table 3). The two candidate genes (Qlv-U295A-A3 and Qlv-U389A-C1) reside in intergenic regions of the macaque genome, which might be novel genes in monkeys. This was supported by the fact that these two sequences did not match any human ESTs by BLAST (data not shown). The two transcripts contained relatively small ORFs (<100 amino acids). Transcripts with small ORFs have been identified in mice and humans, some of which could be actually translated in vitro [36,37]. Alternatively, these mRNAs might be functioning as non-coding RNAs. A large proportion of transcripts are non-coding RNAs, including those having essential functions in transcriptional and translational control [38,39].

4. Conclusion

The data presented here provide an overview of genes expressed in cynomolgus liver to investigate liver physiology for macaques. ESTs for genes encoding a variety of drug-metabolizing enzymes hold great promise in deepening our understanding of drug metabolism in monkeys, which in turn helps to elucidate lineage differences between monkeys and humans. Indeed, our characterization of CYP2C ESTs has identified lineage-specific CYP2C76, which is responsible for lineage differences in drug metabolism [24,30]. Furthermore, the ESTs generated in this study can be a resource for the production of microarrays. Given that our cDNA library was generated with RNAs from only three animals, the EST sequencing using the library originated from the RNA samples of more animals would be useful for the identification of the allelic variants expressed in liver.

Many drug-metabolizing enzyme genes are confined to gene families, many of which have been subjected to gene duplication or gene loss during evolution, resulting in family size differences [40]. This indicates that lineage-specific genes could be identified for gene families even between evolutionarily close lineages such as monkeys and humans. Moreover, physiological differences should partly result from differences at the transcriptional level, for example, by alternative splicing and non-coding RNAs [41]. Further investigation of our EST data will lead to the identification of lineage-specific transcripts generated by alternative splicing and lineage-specific gene gain/loss, as the efforts for identifying such transcripts have succeeded partly in macaques [9,13]. The identified lineage-specific transcripts and genes will help lead to a better understanding