11. Flint S, Gibb DM. Recurrent second trimester miscarriage. *Curr Opin Obstet Gynecol* 1996; **8**: 449–453.
12. Lotgering FK. Clinical aspects of cervical insufficiency. *BMC Pregnancy Childbirth* 2007; **7** (Suppl 1): S17.
13. Abortion. In: Cunningham FG, Leveno KJ, Bloom SL, Hauth JC, Gilstrap LC, Wenstrom KD (eds). *Williams Obstetrics*, 23rd edn. New York, NY: McGraw-Hill Professional, 2005; 215–237.
14. Morency AM, Bujold E. The effect of second-trimester antibiotic therapy on the rate of preterm birth. *J Obstet Gynaecol Can* 2007; **29**: 35–44.
15. Shiffman RL. Continuous low-dose antibiotics and cerclage for recurrent second-trimester pregnancy loss. *J Reprod Med* 2000; **45**: 323–326.
16. Treiber G. Decrease of plasma fibrinogen after eradication of *Helicobacter pylori* infection in patients with ischaemic heart disease. *Heart* 1999; **82**: 646.
17. Raghupathy R, Al Mutawa E, Makhseed M, Al-Azemi M, Azizieh F. Redirection of cytokine production by lymphocytes from women with pre-term delivery by dydrogesterone. *Am J Reprod Immunol* 2007; **58**: 31–38.

The Balance of the Immune System between T Cells and NK Cells in Miscarriage

Akitoshi Nakashima, Tomoko Shima, Kumiko Inada, Mika Ito, Shigeru Saito

Department of Obstetrics and Gynecology, Faculty of Medicine, University of Toyama, Toyama, Japan

Keywords
Decidual NK cell, Granulysin, regulatory T cell, miscarriage, Th17 cell

Correspondence
Shigeru Saito, MD, Ph. D., Department of Obstetrics and Gynecology, Faculty of Medicine, University of Toyama, 2630 Sugitani, Toyama, 930-0194, Japan.
E-mail: s30saito@med.u-toyama.ac.jp

Submission December 28, 2011;
accepted February 2, 2012.

Citation
Nakashima A, Shima T, Inada K, Ito M, Saito S. The balance of the immune system between T cells and NK Cells in miscarriage. Am J Reprod Immunol 2012; 67: 304–310

doi:10.1111/j.1600-0897.2012.01115.x

Immunological dysfunction has been proposed to explain the etiology of recurrent pregnancy loss (RPL). The immunological environment differs between the decidua basalis and decidua parietalis, and also between RPL cases with normal fetal chromosomes and those with abnormal fetal chromosomes. The problem with analyzing decidual tissues from spontaneous abortions is that cause versus effect phenomena are difficult to distinguish. Recent data show that the immune system in a late-stage miscarriage is completely different from that in an early-stage miscarriage. If immunocompetent cells can cause RPL, the immunological environment may be a causative factor, especially in an early-stage miscarriage, at the decidua basalis, and/or in cases of RPL with a normal embryo. Careful examination of the immune system at the decidua basalis in an early-stage miscarriage in RPL cases with normal fetal chromosomes may reveal useful information. This paper aimed at finding a cause of RPL by analyzing the balance of the immune system between T cells and NK cells in an early-stage miscarriage.

Introduction

Recurrent pregnancy loss (RPL) has many known etiologies such as parental cytogenetic anomalies, uterine malformation, endocrine dysfunction, infections, antiphospholipid syndrome, and inherited thrombophilia, yet these identified causes account for only 20–50% of cases.[1–3] A significant proportion of RPL cases are associated with immune etiologies; however, 50% of unexplained RPL is attributed to chromosome abnormalities in the embryo. If immune dysfunction causes RPL, the immunological environment may be important, especially in cases with a normal embryo. The fetus is a semi-allograft for the maternal host, and the maternal immune system is activated by fetal antigens at the feto-maternal interface. In the early stage of pregnancy, fetal tolerance by maternal immune cells is vital to the maintenance of a pregnancy. Maternal tolerance toward fetal alloantigens has been explained by a predominantly Th2-type immunity during pregnancy.[4] Other reports support the idea that Th2-dominant immunity protects the fetus from maternal Th1 cell attack.[5,6] We previously reported decreases in Th2 and Tc2 cells at the decidua basalis but not decidua parietalis in unexplained RPL with normal chromosomal content.[7] Our data also suggested a shift in the Th1/Th2 balance toward a Th1-predominant stage at the implantation site in unexplained RPL with normal chromosomal content. However, this simplistic characterization did not fully explain the immunopathogenesis of abortion/miscarriage.[8,9] Indeed, Th2-dominant immunity was reported in recurrent miscarriage cases, and even in the combined absence of IL-5, IL-9, and IL-13, mice showed a normal pregnancy, even in cases of allogeneic pregnancy.[10] However, this work did not include IL-10 – a Th2-type cytokine that has been reported to be of importance in pregnancy.[11–14]

Recently, the classic Th1/Th2 paradigm of effector T cell differentiation has been expanded by the discovery of another lineage of T helper cells that includes Th17 and Th9 cells,[15] and immunoregulatory T cells, termed Treg cells. IL-17 has an important role in induction of the protective immune response against extracellular bacteria or fungal pathogens.[16,17] Conversely, regulatory T (Treg) cells are potent suppressors of inflammatory immune responses and play pivotal roles in the maintenance of allografted pregnancies in mice and humans.[18–20] Recent findings show the reciprocal development of pathways between the Th17/Treg subsets,[21–23] and an imbalance of Th17/Treg cell development has been reported in recurrent pregnancy loss (RPL).[19,24,25] One should therefore discuss again reproductive events from the view point of the new Th1/Th2/Th17 and Treg paradigms. Additionally, we have noticed that immunological changes at the implantation site (decidua basalis) are usually more emphasized compared with those at the decidua parietalis. The distribution of immune cells in the decidua could therefore provide a clue for verifying the immunological dysfunction in cases of miscarriage. Multiple factors are known to correlate with RPL, and it is quite difficult to unravel them. However, NK cells and T cells are clearly at least a part of the cause of abortion/miscarriage in mice and humans. This paper aimed at finding a cause of RPL by analyzing the balance of the immune system between T cells and NK cells in miscarriage.

Immunopathology

Decidual Natural Killer Cells are Regulated by Hormones

Decidualization of the human endometrium following embryo implantation is normally associated with the massive recruitment of $CD16^{-}CD56^{bright}$ natural killer (NK) cells. In the early pregnant decidua, $CD16^{-}CD56^{bright}$ NK cells constitute the major immune cell population accounting for more than 70% of lymphocytes, whereas $CD4^{+}$ T cells and $CD8^{+}$ T cells are a minor population (<5%).[26,27] NK cells interact with target cells via a series of inhibitory and activating NK cell receptors constitutively expressed by decidual NK cells. Furthermore, $CD16^{-}CD56^{bright}$ NK cells recognize the major histocompatibility complex class I molecules HLA-C, HLA-E, and HLA-G via these inhibitory and activating receptors.[28] To clarify the characteristics of decidual NK cells necessary for maintaining pregnancy, we compared the cytokine expression patterns in decidual $CD16^{-}CD56^{bright}$ NK cells between normal pregnancy and miscarriage. Prior to this analysis, decidual NK cells were divided into four groups: NK1: Th1 cytokine-producing cells, NK2: Th2 cytokine-producing cells, NK3: TGF-β-producing cells, and NKr1: IL-10-producing cells. Normal pregnant women had a significantly increased number of NK3 cells among decidual lymphocytes, compared with women who miscarried.[29] Additionally, a significant increase in NKr1 cells was also observed among peripheral blood lymphocytes in normal pregnant women.[29] There were no significant differences in the proportion of NK1 or NK2 cells between the normal pregnant groups and miscarriage groups. NK3 cells were induced to form from peripheral blood mononuclear cells by prolactin and soluble HLA-G1 *in vitro* (data not shown). Conversely, NKr1 cells were generated from peripheral blood mononuclear cells by human chorionic gonadotropin and progesterone *in vitro*. Progesterone is known to play a crucial role in peripheral blood NK cell accumulation to the uterus.[30] Thus, progesterone, prolactin, human chorionic gonadotropin, and soluble HLA-G1 could contribute to fetal tolerance by inducing the production of immunosuppressive NK subsets (Fig. 1).

NK Cell-Derived Granulysin Induces Apoptosis of Extravillous Trophoblasts

As mentioned before, the proportions of immunosuppressive NK subsets, NK3 and NKr1, were suppressed in miscarriage cases, compared with those of normal pregnancy. In other words, once the immune balance of NK cells is lost, excessive activation of these cells may induce miscarriage.[31,32] Immune changes are well known to occur in RPL, but it was unclear whether maternal NK cells directly attack fetus-derived trophoblasts in humans. To clarify the immunological causes of miscarriage, we examined the relationship between cytotoxic granule proteins, such as granulysin, granzyme B, and perforin, in decidual lymphocytes and the induction of apoptosis in extravillous trophoblasts (EVTs). In the decidua basalis, the number of granulysin-positive $CD56^{bright}$ NK cells significantly increased in miscarriage compared with normal pregnancy; however, granzyme B- and per-

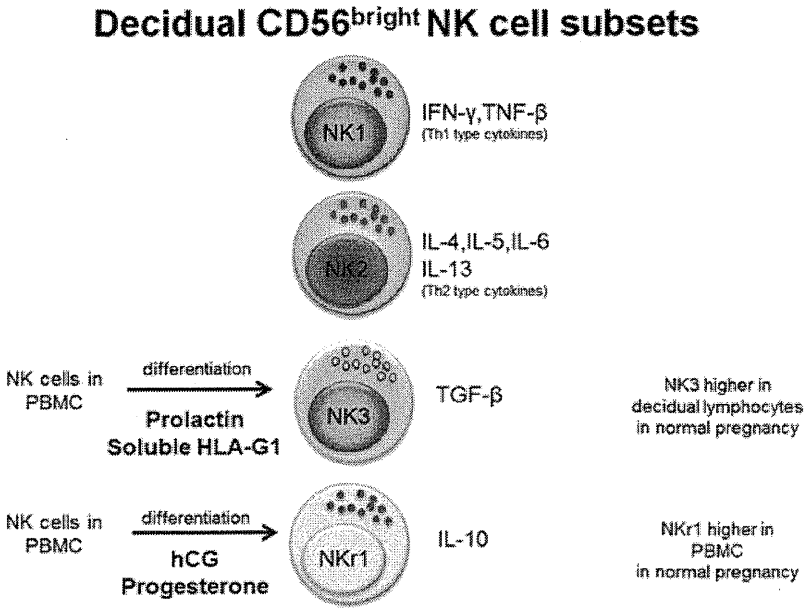
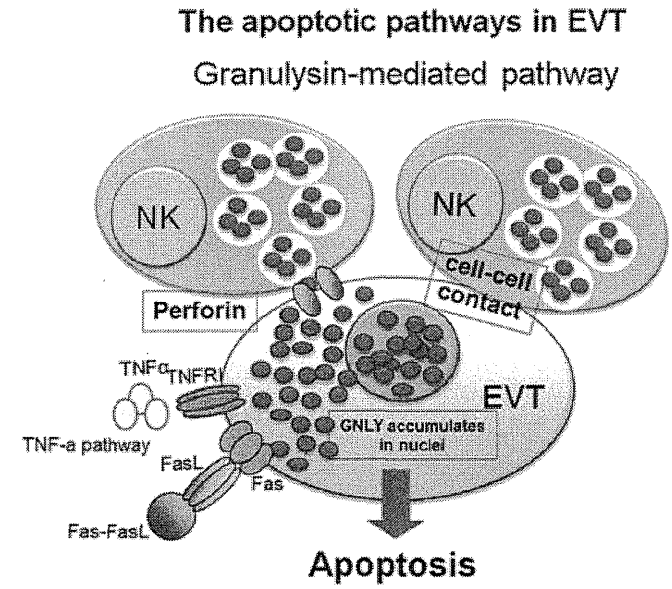


Fig. 1 Decidual CD56^{bright} NK cell subsets.

forin-positive cell numbers did not change.[33] Additionally, granulysin was mainly detected in CD56^{bright} NK cells, but not CD3^{+} T cells in decidual lymphocytes, and the proportion of granulysin in CD56^{bright} NK cells was significantly higher in cases of miscarriage than normal pregnancy. *In vitro* study showed that IL-2-stimulated CD56^{bright} NK cells induced apoptosis in EVT cells (HTR-8/SV40neo) dependent on cell to cell contact and the expression of perforin in decidual NK cells. Furthermore, transfected cells expressing a GFP-granulysin fusion protein induced apoptosis in HTR-8/SV40neo cells independently of caspases. Our results suggest that granulysin-positive decidual NK cells attack EVTs, causing the death of EVTs because of apoptosis (Fig. 2). The data support a new apoptotic path-



Decidual CD56^{bright}NK cell directly induces apoptosis in EVT by Granulysin-transfer.

Fig. 2 The apoptotic pathways in EVT.

way for trophoblasts via decidual NK cell-derived granulysin, suggesting that granulysin is involved in miscarriage. These NK cell-related studies clearly showed the bilateral character for inducing abortion. Conversely, the results suggest the suppression of NK cell activity to be a potential therapeutic target for RPL.

IL-17 in the Pathogenesis of Miscarriage

The identification of a novel helper T (Th)-cell subset, the IL-17-producing Th (Th17) cells, has provided new insight into our understanding of the molecular mechanisms of reproduction. Most of the IL-17-producing cells are $CD4^+$ T cells in the peripheral blood and deciduas.[34–36] Our results showed that Th17 levels in peripheral blood lymphocytes did not change during normal pregnancy,[36] and the proportion of $IL\text{-}17^+$ decidual lymphocytes was significantly higher than that of peripheral blood lymphocytes early in normal pregnancy.[36] But Santner-Nanan et al.[37] reported that the population of Th17 cells in the late stages of pregnancy was significantly smaller than that in non-pregnant subjects. These findings suggest IL-17 to play some role in the establishment of pregnancy. We therefore explored the correlation between IL-17 and miscarriage. IL-17 expression was observed in the decidua parietalis as well as decidua basalis in miscarriage cases.[35] IL-17-stained cells were restricted to lymphocytes, and staining for IL-17A was not observed in villous trophoblasts, EVTs, endometrial gland cells, or endometrial stromal cells.[34] Immunohistochemistry showed that the number of $IL\text{-}17^+$ cells in inevitable abortion cases, which were accompanied with genital bleeding or severe abdominal pain, was significantly higher than that in normal pregnancy cases. $IL\text{-}17^+$ cells were distributed over the entire region of the decidua, the cell column in the decidua basalis, as well as the decidua parietalis.[35] Importantly, there was no significant difference in the number of $IL\text{-}17^+$ cells between missed abortion cases, which were accompanied with no symptoms, and normal pregnancy cases. Additionally, a significant positive correlation was observed between the number of $IL\text{-}17^+$ cells and the number of neutrophils in the abortion cases. These findings suggest IL-17 to be involved in the appearance of symptoms in cases of miscarriage and that IL-17 expression might not be the cause, but rather an effect, of miscarriage.

The Role of Treg Cells in Pregnant Mice

We established pregnant mice with lower numbers of regulatory T cells by an administration of anti-CD25 monoclonal antibodies at days 4.5 and 7.5 of gestation. The ratio of regulatory T cells in $CD4^+$ T cells was decreased proportionally to the dosage of anti-CD25 monoclonal antibody. The miscarriage ratio was significantly increased in proportion to the decrease in regulatory T cells in allogeneic pregnancies, but not in syngeneic pregnancies. Thus, regulatory T cells prevent fetal rejection at the early phase in an allogeneic pregnancy. Our preliminary data showed that paternal antigen-reactive regulatory T (PA-Treg) cells increased in regional lymph nodes of the uterus at day 3.5 and increased in the uterus from day 5.5 in an allogeneic pregnancy. Implantation occurs at day 4.5 in mice, and the increase in PA-Treg cells occurred before implantation in regional lymph nodes. This result suggests that implantation might enhance the migration of PA-Treg cells to the uterus for maintaining pregnancy. Additionally, seminal fluid seems to contribute to the priming by the paternal antigen of maternal immune cells.[38] Taken together, maternal immune cells recognize the paternal antigen in seminal fluid. Subsequently, PA-Treg cells increase in regional lymph nodes before implantation and help to maintain the pregnancy after implantation. Adversely, the decrease in Treg cells solely induces miscarriage in an allogeneic pregnancy. In other words, a decrease in maternal regulatory T cells could cause the fetus to be rejected, suggesting a cause of miscarriage in humans.

Current clinical approach

The Regulation of Peripheral NK Cell Activity Helps to Prevent RPL

We performed a clinical retrospective study to analyze the relationship between peripheral NK cell activity and RPL. Outpatients with unexplained RPL, which all had high NK cell activity in peripheral blood (>40%, $n = 32$), were treated with prednisolone p.o. These patients were then divided into two groups after peripheral NK cell activity was re-estimated: a normalized group (<40%, $n = 21$) and a high cell activity group (>40%, $n = 11$). The number of recurrent pregnancy losses was significantly decreased in the normalized group, compared with

the high activity group ($P = 0.001$). All abortive fetuses were analyzed as to karyotype, and miscarriage cases with an abnormal chromosomal content were excluded. These results suggest peripheral NK cell activity to be one of the markers for disruption of fetal tolerance. However, a randomized control study is needed to estimate the relationship.

The Distribution of Treg Cells in the Decidua of Fetuses with Normal and Abnormal Chromosomal Content

We first reported that $CD4^+$ $CD25^{high}$ regulatory T cells increased in human peripheral blood and further increased in the early pregnant decidua to three times the level found in peripheral blood.[19] Somerset et al.[39] reported that circulating $CD4^+$ $CD25^+$ T cells increase during early pregnancy, peak during the second trimester, and then decline postpartum. Interestingly, these increased peripheral and decidual $CD4^+$ $CD25^{high}$ Treg cell ratios decreased to non-pregnancy levels in miscarriage or RPL cases,[19] suggesting that increased $CD4^+$ $CD25^{high}$ Treg cell numbers are needed to maintain a pregnancy in humans. To further clarify whether immunological dysfunction closely relates to miscarriage in humans, a comparison between fetuses with normal and abnormal chromosomal content would be required. Our preliminary data showed the proportion of Treg cells among decidual lymphocytes to be significantly lower in cases of miscarriage with normal fetal chromosomes than in normal pregnancies. Immunohistochemistry also showed that the number of Treg cells at the decidua basalis was significantly smaller in cases of miscarriage with normal chromosomal content than those with abnormal chromosomal content. These results suggest the decrease in Treg cells at the implantation site to be the cause of miscarriage in humans and are consistent with the data obtained from the pregnant mouse model. However, further studies are needed to test this hypothesis.

Immunological Environment at the Decidua Basalis and Decidua Parietalis

Immune changes are known to occur in RPL, although it is difficult to obtain proof. From our data, numbers of decidual Th17 cells were higher in inevitable abortion cases, but not in missed abortion cases.[35] We presume that missed abortions, which have no symptoms, are an early stage of abortion, whereas inevitable abortions, which do have clinical symptoms such as genital bleeding and severe abdominal pain, are a late stage of abortion. An increase in Th17 cells was observed in inevitable abortion cases but not missed abortion cases and was observed at the decidua parietalis as well as decidua basalis in inevitable abortion cases. Th17 cells therefore might not be the cause, but rather an effect, of miscarriage. Granulysin expression in decidual $CD56^{bright}$ NK cells was significantly higher in miscarriage, compared with that in normal pregnancy. Immunohistochemistry showed that granulysin expression was quite significantly higher in the deicuda basalis than decidua parietalis in missed abortion cases.[33] Our preliminary findings show that there are no significant differences in the number of granulysin-positive cells in the deicuda basalis between miscarriage cases with normal chromosome content and those with abnormal chromosomal content. These results may indicate granulysin to be in a higher cascade than Th17 cells, but this might just be the effect of miscarriage. Conversely, the decrease in Treg cells was observed only in cases of miscarriage with normal chromosomal content.

Potential clinical approach

We propose that (i) A decrease of Treg cells disturbs fetal tolerance by activating maternal immune cells, which might be a cause of miscarriage in fetuses with a normal chromosomal content. (ii) Granulysin-positive $CD56^{bright}$ NK cells directly attack fetal-derived EVT cells, resulting in excessive inflammation at the feto-maternal interface. Reducing peripheral NK cell activity could decrease the

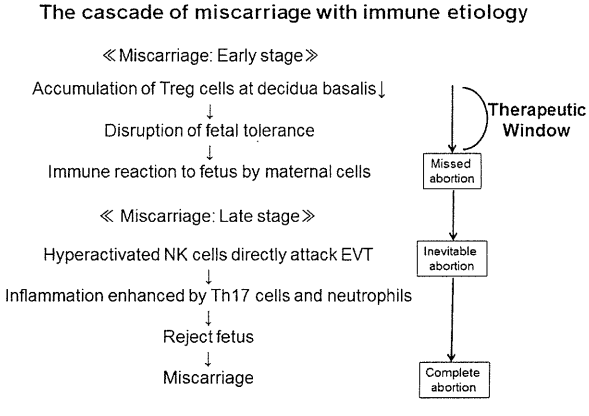


Fig. 3 The cascade of miscarriage with immune etiology.

rate of miscarriage for RPL patients. Further studies are needed to clarify the correlation between Granulysin-positive $CD56^{bright}$ NK cells and higher peripheral NK cell activity in RPL. (iii) Th17 cells directly abrogate the conceptus from the uterus at the final stage of miscarriage (Fig. 3).

To clarify the pathophysiology of unexplained RPL associated with immune etiologies, cases of miscarriage should be classified according to chromosomal content. With regard to immune status, a comparison between cases involving fetuses with a normal versus abnormal chromosomal content would provide useful information. To verify the role of immune cells in pregnancy, not only the proportion but also the distribution of cells should be analyzed. Immunohistochemistry would be useful to clarify the distribution of maternal immune cells in between the decidua basalis and decidua parietalis. Cause versus effect phenomena are difficult to distinguish, but the finding that immunological changes occurred only in missed abortion cases (early stage) with a normal chromosomal content is important to establishing a cause of miscarriage. The onset of miscarriage continues to be a problem. More research will explain the pathophysiology of unexplained RPL associated with immune etiology and may elucidate useful treatments.

References

1 Branch DW, Gibson M, Silver RM: Clinical practice. Recurrent miscarriage. *N Engl J Med* 2010; 363:1740–1747.

2 Christiansen OB, Nielsen HS, Kolte AM: Future directions of failed implantation and recurrent miscarriage research. *Reprod Biomed Online* 2006; 13:71–83.

3 Stephenson MD: Frequency of factors associated with habitual abortion in 197 couples. *Fertil Steril* 1996; 66:24–29.

4 Wegmann TG, Lin H, Guilbert L, Mosmann TR: Bidirectional cytokine interactions in the maternal-fetal relationship: is successful pregnancy a TH2 phenomenon? *Immunol Today* 1993; 14:353–356.

5 Piccinni MP, Beloni L, Livi C, Maggi E, Scarselli G, Romagnani S: Defective production of both leukemia inhibitory factor and type 2 T-helper cytokines by decidual T cells in unexplained recurrent abortions. *Nat Med* 1998; 4:1020–1024.

6 Raghupathy R: Th1-type immunity is incompatible with successful pregnancy. *Immunol Today* 1997; 18:478–482.

7 Michimata T, Sakai M, Miyazaki S, Ogasawara MS, Suzumori K, Aoki K, Nagata K, Saito S: Decrease of T-helper 2 and T-cytotoxic 2 cells at implantation sites occurs in unexplained recurrent spontaneous abortion with normal chromosomal content. *Hum Reprod* 2003; 18:1523–1528.

8 Bates MD, Quenby S, Takakuwa K, Johnson PM, Vince GS: Aberrant cytokine production by peripheral blood mononuclear cells in recurrent pregnancy loss? *Hum Reprod* 2002; 17:2439–2444.

9 Chaouat G, Ledee-Bataille N, Zourbas S, Ostojic S, Dubanchet S, Martal J, Frydman R: Cytokines, implantation and early abortion: re-examining the Th1/Th2 paradigm leads to question the single pathway, single therapy concept. *Am J Reprod Immunol* 2003; 50:177–186.

10 Fallon PG, Jolin HE, Smith P, Emson CL, Townsend MJ, Fallon R, McKenzie AN: IL-4 induces characteristic Th2 responses even in the combined absence of IL-5, IL-9, and IL-13. *Immunity* 2002; 17:7–17.

11 Chaouat G, Assal Meliani A, Martal J, Raghupathy R, Elliott JF, Mosmann T, Wegmann TG: IL-10 prevents naturally occurring fetal loss in the CBA x DBA/2 mating combination, and local defect in IL-10 production in this abortion-prone combination is corrected by in vivo injection of IFN-tau. *J Immunol* 1995; 154:4261–4268.

12 Nagamatsu T, Schust DJ: The immunomodulatory roles of macrophages at the maternal-fetal interface. *Reprod Sci* 2010; 17:209–218.

13 Thaxton JE, Sharma S: Interleukin-10: a multi-faceted agent of pregnancy. *Am J Reprod Immunol* 2010; 63:482–491.

14 Wilczynski JR, Radwan M, Kalinka J: The characterization and role of regulatory T cells in immune reactions. *Front Biosci* 2008; 13:2266–2274.

15 Xing J, Wu Y, Ni B: Th9: a new player in asthma pathogenesis? *J Asthma* 2011; 48:115–125.

16 Kolls JK, Khader SA: The role of Th17 cytokines in primary mucosal immunity. *Cytokine Growth Factor Rev* 2010; 21:443–448.

17 Matsuzaki G, Umemura M: Interleukin-17 as an effector molecule of innate and acquired immunity against infections. *Microbiol Immunol* 2007; 51:1139–1147.

18 Aluvihare VR, Kallikourdis M, Betz AG: Regulatory T cells mediate maternal tolerance to the fetus. *Nat Immunol* 2004; 5:266–271.

19 Sasaki Y, Sakai M, Miyazaki S, Higuma S, Shiozaki A, Saito S: Decidual and peripheral blood CD4+CD25+ regulatory T cells in early pregnancy subjects and spontaneous abortion cases. *Mol Hum Reprod* 2004; 10:347–353.

20 Zenclussen AC, Gerlof K, Zenclussen ML, Sollwedel A, Bertoja AZ, Ritter T, Kotsch K, Leber J, Volk HD: Abnormal T-cell reactivity against paternal antigens in spontaneous abortion: adoptive transfer of pregnancy-induced CD4+CD25+ T regulatory cells prevents fetal rejection in a murine abortion model. *Am J Pathol* 2005; 166:811–822.

21 Crome SQ, Wang AY, Levings MK: Translational mini-review series on Th17 cells: function and regulation of human T helper 17 cells in health and disease. *Clin Exp Immunol* 2010; 159:109–119.

22 Iwakura Y, Ishigame H, Saijo S, Nakae S: Functional specialization of interleukin-17 family members. *Immunity* 2011; 34:149–162.

23 Matsushita S, Higashi T: Human Th17 cell clones and natural immune responses. *Allergol Int* 2008; 57:135–140.

24 Arruvito L, Sanz M, Banham AH, Fainboim L: Expansion of CD4+CD25+and FOXP3+ regulatory T cells during the follicular phase of the menstrual cycle: implications for human reproduction. *J Immunol* 2007; 178:2572–2578.

25 Yang H, Qiu L, Chen G, Ye Z, Lu C, Lin Q: Proportional change of CD4+CD25+ regulatory T cells in decidua and peripheral blood in unexplained recurrent spontaneous abortion patients. *Fertil Steril* 2008; 89:656–661.

26 Moffett-King A: Natural killer cells and pregnancy. *Nat Rev Immunol* 2002; 2:656–663.

27 Nishikawa K, Saito S, Morii T, Hamada K, Ako H, Narita N, Ichijo M, Kurahayashi M, Sugamura K: Accumulation of CD16-CD56+ natural killer cells with high affinity interleukin 2

receptors in human early pregnancy decidua. *Int Immunol* 1991; 3:743–750.

28 Dosiou C, Giudice LC: Natural killer cells in pregnancy and recurrent pregnancy loss: endocrine and immunologic perspectives. *Endocr Rev* 2005; 26:44–62.

29 Higuma-Myojo S, Sasaki Y, Miyazaki S, Sakai M, Siozaki A, Miwa N, Saito S: Cytokine profile of natural killer cells in early human pregnancy. *Am J Reprod Immunol* 2005; 54:21–29.

30 Carlino C, Stabile H, Morrone S, Bulla R, Soriani A, Agostinis C, Bossi F, Mocci C, Sarazani F, Tedesco F, Santoni A, Gismondi A: Recruitment of circulating NK cells through decidual tissues: a possible mechanism controlling NK cell accumulation in the uterus during early pregnancy. *Blood* 2008; 111:3108–3115.

31 Lee J, Choi BC, Cho C, Hill JA, Baek KH, Kim JW: Trophoblast apoptosis is increased in women with evidence of TH1 immunity. *Fertil Steril* 2005; 83:1047–1049.

32 Olivares EG, Munoz R, Tejerizo G, Montes MJ, Gomez-Molina F, Abadia-Molina AC: Decidual lymphocytes of human spontaneous abortions induce apoptosis but not necrosis in JEG-3 extravillous trophoblast cells. *Biol Reprod* 2002; 67:1211–1217.

33 Nakashima A, Shiozaki A, Myojo S, Ito M, Tatematsu M, Sakai M, Takamori Y, Ogawa K, Nagata K, Saito S: Granulysin produced by uterine natural killer cells induces apoptosis of extravillous trophoblasts in spontaneous abortion. *Am J Pathol* 2008; 173:653–664.

34 Ito M, Nakashima A, Hidaka T, Okabe M, Bac ND, Ina S, Yoneda S, Shiozaki A, Sumi S, Tsuneyama K, Nikaido T, Saito S: A role for IL-17 in induction of an inflammation at the fetomaternal interface in preterm labour. *J Reprod Immunol* 2010; 84:75–85.

35 Nakashima A, Ito M, Shima T, Bac ND, Hidaka T, Saito S: Accumulation of IL-17-positive cells in decidua of inevitable abortion cases. *Am J Reprod Immunol* 2010; 64:4–11.

36 Nakashima A, Ito M, Yoneda S, Shiozaki A, Hidaka T, Saito S: Circulating and decidual Th17 cell levels in healthy pregnancy. *Am J Reprod Immunol* 2010; 63:104–109.

37 Santner-Nanan B, Peek MJ, Khanam R, Richarts L, Zhu E, Fazekas de St Groth B, Nanan R: Systemic increase in the ratio between Foxp3+ and IL-17-producing CD4+ T cells in healthy pregnancy but not in preeclampsia. *J Immunol* 2009; 183:7023–7030.

38 Guerin LR, Moldenhauer LM, Prins JR, Bromfield JJ, Hayball JD, Robertson SA: Seminal fluid regulates accumulation of FOXP3+ regulatory T Cells in the preimplantation mouse uterus through expanding the FOXP3+ cell pool and CCL19-mediated recruitment. *Biol Reprod* 2011; 85:397–408.

39 Somerset DA, Zheng Y, Kilby MD, Sansom DM, Drayson MT: Normal human pregnancy is associated with an elevation in the immune suppressive CD25+ CD4+ regulatory T-cell subset. *Immunology* 2004; 112:38–43.

The Frequency of Peripheral Blood CD4+ CD25high FoxP3+ and CD4+ CD25− FoxP3+ Regulatory T Cells in Normal Pregnancy and Pre-Eclampsia

Gergely Toldi[1], Shigeru Saito[2], Tomoko Shima[2], Amrita Halmos[3], Zoltán Veresh[3], Barna Vásárhelyi[4], János Rigó Jr[3], Attila Molvarec[3]

[1]First Department of Pediatrics, Semmelweis University, Budapest, Hungary;
[2]Department of Obstetrics and Gynecology, University of Toyama, Toyama, Japan;
[3]First Department of Obstetrics and Gynecology, Semmelweis University, Budapest, Hungary;
[4]Department of Laboratory Medicine, Semmelweis University, Budapest, Hungary

Keywords
CD25, FoxP3, pre-eclampsia, pregnancy, Treg

Correspondence
Attila Molvarec, Hangya lépcső 9, Budapest H-1121, Hungary.
E-mail: molvarec@freemail.hu

Submission February 6, 2012;
accepted March 23, 2012.

Citation
Toldi G, Saito S, Shima T, Halmos A, Veresh Z, Vásárhelyi B, Rigó J Jr, Molvarec A. The frequency of peripheral blood CD4+ CD25high FoxP3+ and CD4+ CD25− FoxP3+ regulatory T cells in normal pregnancy and pre-eclampsia. Am J Reprod Immunol 2012; 68: 175–180

doi:10.1111/j.1600-0897.2012.01145.x

Problem
Regulatory T cells (Tregs) play an important role in the development of pregnancy-specific immune tolerance. We aimed to determine the peripheral frequency of a recently described Treg subpopulation, the CD4+ CD25− FoxP3+ Treg subset, and its correlation with the conventional CD4+ CD25high FoxP3+ Tregs in normal pregnancy (NP) and pre-eclampsia (PE) compared to non-pregnant (non-P) women. We also examined the proportion of the activated CD4+ CD25high FoxP3high Treg subset within conventional Treg cells.

Method
We took peripheral blood samples from 20 PE, 20 NP, and 12 non-P women and determined the frequency of the above Treg subsets using flow cytometry.

Results
The frequency of conventional CD4+ CD25high FoxP3+ Tregs and activated CD4+ CD25high FoxP3high Tregs, but also that of non-conventional CD4+ CD25− FoxP3+ Tregs was higher in NP compared to non-P women, but lower again in PE, reaching comparable levels to the non-P group. The ratios of CD4+ CD25high FoxP3+ and CD4+ CD25− FoxP3+ Treg subsets were constant in all three investigated groups.

Conclusion
Our results indicate that the frequency of conventional and non-conventional Tregs alters simultaneously, and the presence in circulation of both of these Treg subsets is similarly important in the adequate development of pregnancy-specific immune tolerance.

Introduction

Pre-eclampsia (PE) is an immune-mediated, pregnancy-specific syndrome that affects at least 5% of all pregnancies. The main clinical symptoms of this disorder are hypertension and proteinuria that develop in the second half of pregnancy in a previously normotensive woman. In most of the cases, these symptoms disappear rapidly after delivery.

Pre-eclampsia is characterized by an excessive maternal systemic inflammatory response to pregnancy with activation of both the innate and adaptive

arms of the immune system.[1,2] Activated neutrophils, monocytes, and natural killer cells initiate inflammation, which induces endothelial dysfunction, and activated T cells may support inadequate tolerance during pregnancy. An important feature of systemic inflammation in PE is the absence of Th2 skewness characteristic for normal pregnancy (NP), and thus the predominance of a Th1-type immunity.[3]

Recently, the discovery of two novel T helper subsets, referred to as Th17 and regulatory T cells (Tregs), led to the transformation of the Th1/Th2 concept of immunity into a four-component paradigm of a complex and mutually interacting network of T helper cells.[4] Indeed, besides the imbalance of Th1 and Th2 cells, alterations of the prevalence of Th17 and Treg cells have also been suggested to be of importance in the development of maternal systemic inflammation in PE.[5] Th17 cells have an important role in the pathogenesis of autoimmune disorders and in the induction and maintenance of chronic inflammation.[6] Indeed, the frequency of Th17 cells was found to be elevated in the peripheral blood of PE patients compared to NP women in the third trimester of pregnancy.[7,8]

The effect of Th17 cells on the inflammatory balance is opposed by Tregs. Lower than normal Treg frequency may contribute to the exaggerated systemic inflammation present in PE. Indeed, a number of groups including ours demonstrated that the prevalence of peripheral Tregs is lower in PE compared to healthy pregnancy.[9–13] Of note, contradictory data suggesting no alteration with regard to the prevalence of Tregs in PE are also available.[14,15]

Emerging evidence indicates that Tregs display molecular and functional heterogeneity, also supporting the idea that Treg cells retain plasticity.[16] Therefore, the population of Tregs has been further grouped based on intracellular and cell surface markers. These Treg subgroups were shown to have distinct functional properties that play different roles in the regulation of the inflammatory response. Very recently, Steinborn et al. demonstrated that the composition of the total Treg cell pool is characterized by distinct changes in pregnancy-associated disorders compared to NP. For example, its percentage of naive DR− CD45RA+ Tregs was reduced, while its percentage of DRhigh+ CD45RA− and DRlow+ CD45RA− Tregs was increased in the presence of PE.[17]

Recent studies showed that FoxP3, required for the development of Tregs, is specifically expressed in most CD4+ CD25+ T cells but also in a small part of CD4+ CD25− T cells. Furthermore, the ectopic expression of FoxP3 induced suppressive function in peripheral CD4+ CD25− T cells. These results demonstrate that FoxP3 is a key player in Tregs and that CD25 is a useful cell surface marker for many, but not all Treg cells.[18–21]

Based on the findings of Miyara et al.,[22] the conventional CD4+ CD25high Treg subset might be further divided into activated (CD4+ CD25high FoxP3high) and resting (CD4+ CD25high FoxP3low) regulatory T cells. Both of these subtypes are suppressive; however, they differ in proliferation dynamics and responsiveness. Terminally differentiated activated Treg cells are more responsive, but die rapidly, whereas resting Tregs have limited responsiveness, but are able to proliferate and convert into activated Treg cells with elevated FoxP3 expression.

In this study, we aimed to determine the peripheral frequency of a recently described Treg subpopulation, the non-conventional CD4+ CD25− FoxP3+ Treg subset and its correlation with the conventional CD4+ CD25high FoxP3+ Tregs in NP and PE compared to non-pregnant (non-P), healthy age-matched women. We also examined the proportion of the activated CD4+ CD25high FoxP3high Treg subset within conventional Treg cells.

Materials and methods

We took peripheral blood samples from 20 PE, 20 NP, and 12 non-P women. Clinical characteristics of the study participants are shown in Table I. PE was defined by increased blood pressure ($\geq$ 140 mmHg systolic and/or $\geq$ 90 mmHg diastolic on $\geq$ 2 occasions at least 6 hr apart) that occurred after 20 weeks of gestation in a woman with previously normal blood pressure, accompanied by proteinuria ($\geq$ 0.3 g/24 hr or $\geq$ 1+ on dipstick in the absence of urinary tract infection). PE was regarded as severe if any of the following criteria was present: blood pressure $\geq$ 160 mmHg systolic or $\geq$ 110 mmHg diastolic, or proteinuria $\geq$ 5 g/24 hr (or $\geq$ 3+ on dipstick). Early onset of PE was defined as onset of the disease before 34 weeks of gestation. Fetal growth restriction (IUGR) was diagnosed if the fetal birth weight was below the 10th percentile for gestational age and gender, based on Hungarian birth weight percentiles. Exclusion criteria were multifetal gestation, chronic hypertension, diabetes mellitus, autoimmune disease, angiopathy, renal disorder, maternal or fetal infection, and fetal congenital anomaly.

American Journal of Reproductive Immunology 68 (2012) 175–180

Table I Clinical Characteristics of Healthy Non-Pregnant and Normal Pregnant Women and Pre-Eclamptic Patients

	Healthy non-pregnant women (n = 12)	Normal pregnant women (n = 20)	Pre-eclamptic patients (n = 20)
Age (years)	34.5 (30.5–35.5)	32 (27–34)	32.5 (30–34.5)
Pre-pregnancy BMI (kg/m^2)	23.9 (21.4–27.0)	22.4 (20.4–27.5)	25.5 (21.5–32.6)
Smokers (%)	3 (25.0)	0 (0)[a]	2 (10.0)
Primiparas (%)	NA	13 (65.0)	12 (60.0)
Systolic blood pressure (mmHg)	110 (100–110)	110 (110–120)	150 (140–160)[b,c]
Diastolic blood pressure (mmHg)	65 (60–70)	70 (60–77)	100 (93–100)[b,c]
Gestational age at blood draw (weeks)	NA	37 (36–38)	33.5 (31–37)
Gestational age at delivery (weeks)	NA	38 (38–39)	34.5 (32–38)[d]
Fetal birth weight (g)	NA	3335 (3290–3570)	1710 (1445–2885)[c]
Fetal growth restriction (%)	NA	0 (0)	7 (35.0)[d]

NA, not applicable; BMI, body mass index.
Data are presented as median (interquartile range) for continuous variables and as number (%) for categorical variables.
[a]$P < 0.05$ versus healthy non-pregnant women.
[b]$P < 0.001$ versus healthy non-pregnant women.
[c]$P < 0.001$ pre-eclamptic patients versus healthy pregnant women.
[d]$P < 0.05$ pre-eclamptic patients versus healthy pregnant women.

None of the pregnant women were in active labor, and none had rupture of membranes at the time of blood sampling. Informed consent was obtained from all subjects, and our study was reviewed and approved by an independent ethical committee of the institution. The study was adhered to the tenets of the most recent revision of the Declaration of Helsinki.

Peripheral blood mononuclear cells (PBMCs) were separated by a standard density gradient centrifugation (27 min, 400 *g*, 22°C, Ficoll Paque; Amersham Biosciences AB, Uppsala, Sweden) from freshly drawn blood collected in lithium heparin-treated tubes (BD Vacutainer; BD Biosciences, San Jose, CA, USA). This cell suspension was washed twice in phosphate-buffered saline. Cells were suspended in RPMI 1640 medium (Sigma-Aldrich, St. Louis, MO, USA).

Peripheral blood mononuclear cells were stained for 30 min at 4°C with PE Cy7-conjugated CD4 and APC-conjugated CD25 mAbs (PharMingen, San Diego, CA, USA). After washing, cells were fixed with fixation/permeabilization solution and treated with permeabilization buffer according to the manufacturer's instructions (eBioscience, San Diego, CA, USA). They were then stained with PE-conjugated FoxP3 mAb (eBioscience) for 30 min at 4°C. After washing, cells were analyzed on a BD FACSAria flow cytometer (BD Biosciences). About 200,000 cells were recorded. The population of lymphocytes was gated from PBMCs according to forward scatter characteristics and side scatter characteristics. Isotype-matched PE-conjugated mouse IgG1 antibody was used as a control (eBioscience). The intra-assay coefficient of variation of our flow cytometric measurements was below 25%.

Data are expressed as median and interquartile range. Comparisons between sample populations were made with the Kruskal–Wallis analysis of variance by ranks test. Multiple comparisons of mean ranks for all groups were carried out as post hoc tests. Adjustment for gestational age at blood sampling was performed with analysis of covariance (ANCOVA). Two-tailed P-values <0.05 were considered significant. Statistics were calculated using the STATISTICA software (version 8.0; StatSoft, Inc., Tulsa, Oklahoma, USA).

Results

Our results are summarized in Fig. 1. The frequency of CD4+ CD25high FoxP3+ cells was lower in non-P women than in NP and higher in NP than in PE [2.99 (2.41–3.48)% versus 5.02 (4.20–5.46)% versus 2.97 (2.23–3.40)%, $P < 0.001$, Fig. 1a].

The proportion of activated CD4+ CD25high FoxP3high Tregs among CD4+ CD25high cells was also lower in non-P women than in NP and higher in NP

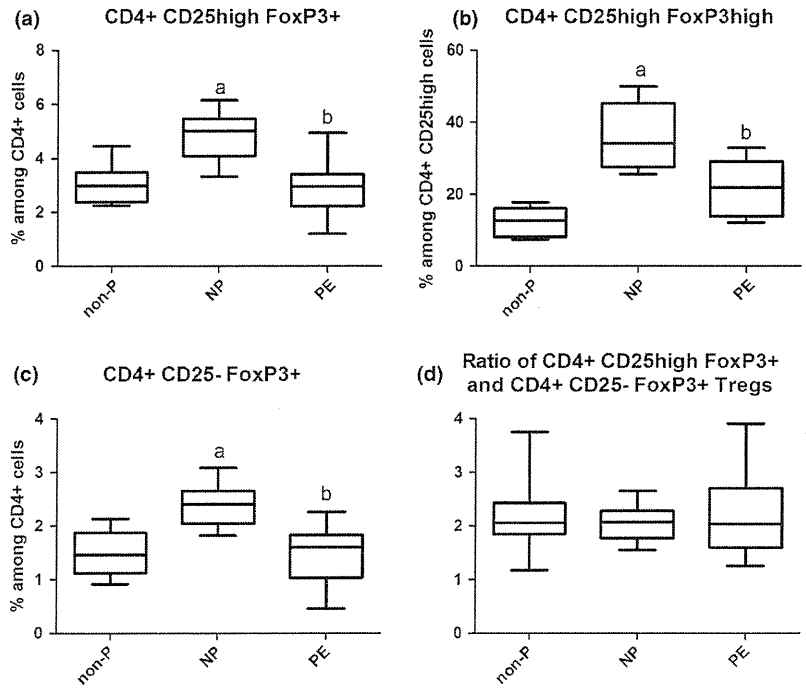


Fig. 1 Box-plots representing the frequency of peripheral blood CD4+ CD25high FoxP3+ (a), CD4+ CD25high FoxP3high (b) and CD4+ CD25− FoxP3+ (c) regulatory T cells (Tregs) as well as the ratio of CD25high FoxP3+ and CD25− FoxP3+ subsets (d) in healthy non-pregnant (non-P) and normal pregnant (NP) women and pre-eclamptic (PE) patients. Horizontal line: median; Box: interquartile range (25–75 percentile); Whisker: range. $^{a}P < 0.001$ versus healthy non-pregnant women; $^{b}P < 0.001$ versus healthy pregnant women.

than in PE [12.60 (8.15–16.10)% versus 34.20 (27.60–45.35)% versus 21.90 (13.85–29.20)%, $P < 0.001$, Fig. 1b].

Similarly, the frequency of CD4+ CD25− FoxP3+ cells was higher in NP than in non-P women, and lower again in PE [1.46 (1.17–1.76)% versus 2.41 (2.06–2.65)% versus 1.60 (1.05–1.80)%, $P < 0.001$, Fig. 1c].

The differences in the proportions of CD4+ CD25high FoxP3+ and CD4+ CD25− FoxP3+ cells among CD4+ cells, as well as in those of activated CD4+ CD25high FoxP3high Tregs among CD4+ CD25high cells between NP and PE remained significant even after adjustment for gestational age at blood draw in ANCOVA.

Within PE patients, no difference was detected in the frequency of the above subsets when patients were compared based on early or late onset of PE (n of early PE = 11), severity of PE (n of severe PE = 9), or the presence of fetal growth restriction (n = 7) (data not shown).

Additionally, we examined the relationship between CD4+ CD25high FoxP3+ and CD4+ CD25− FoxP3+ cells in all three study groups. No difference was detected between the ratios of the above subsets in non-P, NP, and PE women (Fig. 1d).

Discussion

In this study, we demonstrated that the frequency of the conventional CD4+ CD25high FoxP3+ Treg subset, but also that of non-conventional CD4+ CD25− FoxP3+ Tregs is higher in NP compared to non-P women, but lower again in PE, reaching comparable levels to the non-P group.

Several studies have investigated the frequency of Treg cells in NP and PE. Heikkinen et al. and Somerset et al. observed an increase in the population of peripheral Tregs in early NP. They described a peak of this population during the second trimester and a subsequent gradual decrease to levels slightly higher than non-P levels during the post-partum period.[23,24]

Data concerning the alteration of circulating Treg frequency in PE compared to NP are not fully in line. Paeschke et al.[14] and Hu et al.[15] reported comparable, while Sasaki et al.,[9] Darmochwal-Kolarz et al.,[10] Prins et al.,[11] Steinborn et al.,[12] and our group[13] demonstrated lower Treg frequency in peripheral blood of PE patients compared to NP women. Of note, Miko et al.[25] also found lower Treg percentages in PE in comparison with NP and also to non-P women. Furthermore, Sasaki et al.[9] reported

that the prevalence of Tregs is lower not only in peripheral blood but also in deciduas of PE patients compared with NP women.

In the above investigations, Tregs were identified based on the expression of a cell surface activation marker, CD25 (CD4+ CD25high cells), or based on intracellular FoxP3 positivity (CD4+ FoxP3+ or CD4+ CD25high FoxP3+ cells). Miyara et al.[22] further divided Treg cells into subgroups using these and other markers. They described that terminally differentiated activated CD4+ CD25high FoxP3high Treg cells are more responsive, but less capable to proliferate than resting CD4+ CD25high FoxP3low Tregs. Based on our results, not only the prevalence of Tregs in general, but also the frequency of this more responsive subset is lower in PE compared to NP, which might contribute to the activation of the innate and adaptive immune system. This observation is in agreement with earlier findings: in PE, the significantly diminished percentage of CD4+ CD25+ FoxP3high+ Tregs was associated with lower suppressive capacity.[12]

Previously, Nishioka et al.[20] identified FoxP3+ suppressive cells among hyporesponsive CD4+ CD25− T cells in aged mice. *In vitro* analyses showed that this population was negligible in the young controls. Their results suggest that the age-related decline in T cell-mediated immune responses is ascribable to changes in the CD4+ CD25− T-cell population and not to a functional augmentation of suppressive CD4+ CD25+ T cells. Furthermore, their results demonstrate that FoxP3, but not CD25, is the best marker of suppressive T cells in aged hosts.[20] The regulatory function of CD4+ CD25− FoxP3+ T cells was reinforced in human samples.[21] This is of interest because the two earlier studies reporting comparable number of Tregs in PE[14,15] used CD25, but not FoxP3 as a marker to identify Tregs.

In our study, the ratios of CD4+ CD25high FoxP3+ and CD4+ CD25− FoxP3+ Treg subsets were constant in all three investigated groups, showing that the frequency of conventional and non-conventional Tregs alters simultaneously, and the presence in circulation of both of these Treg subsets is similarly important in the adequate development of pregnancy-specific immune tolerance.

Taken together, these data indicate that the expression of FoxP3 is a more important factor in Treg function, and a better marker of Tregs in pregnancy than the cell surface IL-2 receptor alpha chain, CD25.

Conclusion

In conclusion, our data indicate that not only the frequency of conventional CD4+ CD25high FoxP3+ Tregs, but also that of the non-conventional CD4+ CD25− FoxP3+ Treg subset is lower in PE than in NP. As these Treg subsets alter simultaneously in peripheral blood, we conclude that not only conventional Tregs but also the CD25− non-conventional Treg subset is of importance in the development of pregnancy-specific immune tolerance.

Acknowledgements

This study was supported by grants OTKA 101661 and TÁMOP-4.2.2.-08/1/KMR-2008-0004, as well as by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

1 Redman CW, Sacks GP, Sargent IL: Preeclampsia: an excessive maternal inflammatory response to pregnancy. *Am J Obstet Gynecol* 1999; 180:499–506.
2 Saito S, Shiozaki A, Nakashima A, Sakai M, Sasaki Y: The role of the immune system in preeclampsia. *Mol Aspects Med* 2007; 28:192–209.
3 Saito S, Sakai M: Th1/Th2 balance in preeclampsia. *J Reprod Immunol* 2003; 59:161–173.
4 Saito S, Nakashima A, Shima T, Ito M: Th1/Th2/Th17 and regulatory T-cell paradigm in pregnancy. *Am J Reprod Immunol* 2010; 63:601–610.
5 Saito S: Th17 cells and regulatory T cells: new light on pathophysiology of preeclampsia. *Immunol Cell Biol* 2010; 88:615–617.
6 Basso AS, Cheroutre H, Mucida D: More stories on Th17 cells. *Cell Res* 2009; 19:399–411.
7 Santner-Nanan B, Peek MJ, Khanam R, Richarts L, Zhu E, Fazekas de St Groth B, Nanan R: Systemic increase in the ratio between Foxp3+ and IL-17-producing CD4+ T cells in healthy pregnancy but not in preeclampsia. *J Immunol* 2009; 183:7023–7030.
8 Toldi G, Rigo J Jr, Stenczer B, Vasarhelyi B, Molvarec A: Increased prevalence of IL-17-producing peripheral blood lymphocytes in preeclampsia. *Am J Reprod Immunol* 2011; 66:223–229.
9 Sasaki Y, Darmochwal-Kolarz D, Suzuki D, Sakai M, Ito M, Shima T, Shiozaki A, Rolinski J, Saito S: Proportion of peripheral blood and decidual CD4(+) CD25(bright) regulatory T cells in pre-eclampsia. *Clin Exp Immunol* 2007; 149:139–145.
10 Darmochwal-Kolarz D, Saito S, Rolinski J, Tabarkiewicz J, Kolarz B, Leszczynska-Gorzelak B, Oleszczuk J: Activated T lymphocytes in pre-eclampsia. *Am J Reprod Immunol* 2007; 58:39–45.
11 Prins JR, Boelens HM, Heimweg J, Van der Heide S, Dubois AE, Van Oosterhout AJ, Erwich JJ: Preeclampsia is associated with lower percentages of regulatory T cells in maternal blood. *Hypertens Pregnancy* 2009; 28:300–311.
12 Steinborn A, Haensch GM, Mahnke K, Schmitt E, Toermer A, Meuer S, Sohn C: Distinct subsets of regulatory T cells during

pregnancy: is the imbalance of these subsets involved in the pathogenesis of preeclampsia? *Clin Immunol* 2008; 129:401–412.

13 Toldi G, Svec P, Vásárhelyi B, Mészáros G, Rigó J, Tulassay T, Treszl A: Decreased number of FoxP3+ regulatory T cells in preeclampsia. *Acta Obstet Gynecol Scand* 2008; 87:1229–1233.

14 Paeschke S, Chen F, Horn N, Fotopoulou C, Zambon-Bertoja A, Sollwedel A, Zenclussen ML, Casalis PA, Dudenhausen JW, Volk HD, Zenclussen AC: Pre-eclampsia is not associated with changes in the levels of regulatory T cells in peripheral blood. *Am J Reprod Immunol* 2005; 54:384–389.

15 Hu D, Chen Y, Zhang W, Wang H, Wang Z, Dong M: Alteration of peripheral CD4+CD25+ regulatory T lymphocytes in pregnancy and pre-eclampsia. *Acta Obstet Gynecol Scand* 2008; 87:190–194.

16 Li L, Boussiotis VA: Molecular and functional heterogeneity of T regulatory cells. *Clin Immunol* 2011; 141:244–252.

17 Steinborn A, Schmitt E, Kisielewicz A, Rechenberg S, Seissler N, Mahnke K, Schaier M, Zeier M, Sohn C: Pregnancy-associated diseases are characterized by the composition of the systemic regulatory T cell (T(reg)) pool with distinct subsets of T(regs). *Clin Exp Immunol* 2012; 167:84–98.

18 Hori S, Nomura T, Sakaguchi S: Control of regulatory T cell development by the transcription factor Foxp3. *Science* 2003; 299:1057–1061.

19 Fontenot JD, Gavin MA, Rudensky AY: Foxp3 programs the development and function of CD4+ CD25+ regulatory T cells. *Nat Immunol* 2003; 4:330–336.

20 Nishioka T, Shimizu J, Iida R, Yamazaki S, Sakaguchi S: CD4+CD25+Foxp3+ T cells and CD4+CD25-Foxp3+ T cells in aged mice. *J Immunol* 2006; 176:6586–6593.

21 Yang ZZ, Novak AJ, Ziesmer SC, Witzig TE, Ansell SM: CD70+ non-Hodgkin lymphoma B cells induce Foxp3 expression and regulatory function in intratumoral CD4+CD25 T cells. *Blood* 2007; 110:2537–2544.

22 Miyara M, Yoshioka Y, Kitoh A, Shima T, Wing K, Niwa A, Parizot C, Taflin C, Heike T, Valeyre D, Mathian A, Nakahata T, Yamaguchi T, Nomura T, Ono M, Amoura Z, Gorochov G, Sakaguchi S: Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. *Immunity* 2009; 30:899–911.

23 Heikkinen J, Möttönen M, Alanen A, Lassila O: Phenotypic characterization of regulatory T cells in the human decidua. *Clin Exp Immunol* 2004; 136:373–378.

24 Somerset DA, Zheng Y, Kilby MD, Sansom DM, Drayson MT: Normal human pregnancy is associated with an elevation in the immune suppressive CD25+ CD4+ regulatory T-cell subset. *Immunology* 2004; 112:38–43.

25 Miko E, Szereday L, Barakonyi A, Jarkovich A, Varga P, Szekeres-Bartho J: Immunoactivation in preeclampsia: Vdelta2+ and regulatory T cells during the inflammatory stage of disease. *J Reprod Immunol* 2009; 80:100–108.

Journal of Human Genetics (2012) 57, 787–795

www.nature.com/jhg

ORIGINAL ARTICLE

The history of human populations in the Japanese Archipelago inferred from genome-wide SNP data with a special reference to the Ainu and the Ryukyuan populations

Japanese Archipelago Human Population Genetics Consortium (Consortium members: Timothy Jinam[1,18], Nao Nishida[2,19], Momoki Hirai[3,19], Shoji Kawamura[3,19], Hiroki Oota[4,19], Kazuo Umetsu[5,19], Ryosuke Kimura[6,19], Jun Ohashi[7,19], Atsushi Tajima[8,19], Toshimichi Yamamoto[9,19], Hideyuki Tanabe[10,19], Shuhei Mano[11,19], Yumiko Suto[12,19], Tadashi Kaname[13], Kenji Naritomi[13], Kumiko Yanagi[13], Norio Niikawa[14], Keiichi Omoto[15,19], Katsushi Tokunaga[2,19] and Naruya Saitou[1,16,17,19])

The Japanese Archipelago stretches over 4000 km from north to south, and is the homeland of the three human populations; the Ainu, the Mainland Japanese and the Ryukyuan. The archeological evidence of human residence on this Archipelago goes back to >30 000 years, and various migration routes and root populations have been proposed. Here, we determined close to one million single-nucleotide polymorphisms (SNPs) for the Ainu and the Ryukyuan, and compared these with existing data sets. This is the first report of these genome-wide SNP data. Major findings are: (1) Recent admixture with the Mainland Japanese was observed for more than one third of the Ainu individuals from principal component analysis and frappe analyses; (2) The Ainu population seems to have experienced admixture with another population, and a combination of two types of admixtures is the unique characteristics of this population; (3) The Ainu and the Ryukyuan are tightly clustered with 100% bootstrap probability followed by the Mainland Japanese in the phylogenetic trees of East Eurasian populations. These results clearly support the dual structure model on the Japanese Archipelago populations, though the origins of the Jomon and the Yayoi people still remain to be solved.
Journal of Human Genetics (2012) **57**, 787–795; doi:10.1038/jhg.2012.114; published online 8 November 2012

Keywords: admixture; Ainu; Japanese Archipelago; population; Ryukyuan; SNP

INTRODUCTION

The origins of the people living in the Japanese Archipelago have been studied for a long time (for review, see Suzuki[1] and Saitou[2]). The standard theory based on craniofacial data is the dual structure model proposed by Hanihara.[3] According to this model, the first migrants to the Japanese Archipelago came from somewhere in Southeast Asia in the Upper Paleolithic age, who were ancestors of the Jomon people. The second wave of migration took place later in

[1]Department of Genetics, School of Life Science, Graduate University for Advanced Studies (SOKENDAI), Mishima, Japan; [2]Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [3]Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan; [4]Laboratory of Anatomy and Physical Anthropology, Kitasato University School of Medicine, Sagamihara, Japan; [5]Department of Forensic Medicine, Yamagata University School of Medicine, Yamagata, Japan; [6]Transdisciplinary Research Organization for Subtropical and Island Studies (TRO-SIS), University of the Ryukyus, Nishihara, Japan; [7]Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan; [8]Department of Human Genetics, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima, Japan; [9]Department of Legal Medicine and Bioethics, Nagoya University School of Medicine, Nagoya, Japan; [10]Department of Evolutionary Studies of Biosystems, School of Advanced Sciences, Graduate University for Advanced Studies (SOKENDAI), Hayama, Japan; [11]Department of Mathematical Analysis and Statistical Inference, The Institute of Statistical Mathematics, Tachikawa, Tokyo, Japan; [12]Research Center for Radiation Emergency Medicine, National Institute of Radiological Sciences, Chiba, Japan; [13]Department of Medical Genetics, Graduate School of Medicine, University of the Ryukyus, Nishihara, Japan; [14]Health Sciences University of Hokkaido, Tobetsu, Japan; [15]Department of Anthropology, Faculty of Science, The University of Tokyo, Tokyo, Japan; [16]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan and [17]Division of Population Genetics, National Institute of Genetics, Mishima, Japan
[18]Current address: Division of Human Genetics, National Institute of Genetics, Mishima, Japan.
[19]Members of the 'Asian Archival DNA Repository Consortium'.
Correspondence: Dr N Saitou, Division of Population Genetics, National Institute of Genetics, 1111 Yata, Mishima 411-8540, Japan.
E-mail: saitounr@lab.nig.ac.jp
or Dr K Tokunaga, Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan.
E-mail: tokunaga@m.u-tokyo.ac.jp
or Dr K Omoto, Faculty of Science, Department of Anthropology, The University of Tokyo, Tokyo 113-0033, Japan.
E-mail: kocolias@msg.biglobe.ne.jp
Received 17 March 2012; revised 27 August 2012; accepted 30 August 2012; published online 8 November 2012

the Yayoi period, and the people came in this time from Northeast Asia. The indigenous Jomon people and the new migrants in and after the Yayoi period gradually mixed with each other. This model provides a reasonable explanation for the morphological similarity between the Ainu people of Hokkaido, the northernmost main island of the Japanese Archipelago, and the Ryukyuan (or Okinawan) people in the Southwest Archipelago, despite a large geographical distance. The similarity of these peoples was already noticed 101 years ago by von Baelz[4] as the Ainu–Ryukyuan common origin theory.

A series of studies on genetic polymorphisms of classic markers, such as blood groups, serum proteins and red cell enzymes, were carried out for human populations in the Japanese Archipelago from the 1960s to the 1970s, and the Ainu and the Ryukyuan populations were also studied (for example, Misawa and Hayashida;[5,6] Omoto and Harada;[7] Nakajima *et al.*;[8] Omoto *et al.*[9]). Omoto[10,11] computed genetic distances among various populations of the world, and by constructing a phylogenetic tree he concluded that the Ainu population may have originated in East Asia, in spite of their unique morphological characters somewhat resembling West Eurasians. Omoto *et al.*[9] estimated genetic distances among the Ainu, the Ryukyuan, the Mainland Japanese and the Chinese populations. Although they did not show a phylogenetic tree, distance relationships indicated a clustering of the Ainu and the Ryukyuan if we apply the neighbor-joining method.[12] Omoto[13] constructed a phylogenetic tree for various populations of the Mainland Japan, the Ainu and the Ryukyuan, and showed that the Ainu and the Ryukyuan were clustered together. Seven serum protein polymorphism data were used for this tree construction, and this tree was the first one to suggest the genetic similarity between the Ainu and the Ryukyuan. Nei[14] constructed a neighbor-joining[12] tree of 15 Eurasian populations based on D_A distances[15] computed from allele frequency data of 18 polymorphic classic markers. The Ainu and the Ryukyuan clustered with 62% bootstrap probability, followed by the Mainland Japanese population.

Omoto and Saitou[16] constructed an unrooted tree of three populations in the Japanese Archipelago (Ainu, Ryukyuan and Mainland Japanese) and Korean from allele frequency data of 25 classic polymorphic marker loci, and showed that Ainu and Ryukyuan clustered with 85% and 74% bootstrap probabilities when D_A distances[15] and D_{st} distances[17] were used, respectively. These probabilities were considerably higher than the random expectation (33% = 1/3), and they considered these results as partially supporting the dual structure hypothesis of Hanihara.[3] This is because the Ainu–Ryukyuan clustering was expected from the contrasting genetic backgrounds of Jomon vs Yayoi, irrespective of their origins. Saitou[18] later constructed a distance-based phylogenetic network[19] of these four populations using D_A distance values obtained by Omoto and Saitou,[16] and found that the length of split separating Ainu and Ryukyuan populations from Mainland Japanese and Korean was much longer than that separating Mainland Japanese and Ryukyuan from the rest of two populations. It confirmed a tree-like structure of the four populations shown by Omoto and Saitou.[16]

Mitochondrial DNA and Y chromosomal DNA examinations became popular for human population studies from the 1980s, and a series of papers were published for human populations of the Japanese Archipelago.[20–28] All of these studies showed some genetic similarity between the Ainu and the Ryukyuan populations. However, mitochondrial DNA and Y chromosomes are both non-recombining, and the genetic information that can be extracted is limited. Meanwhile, Tokunaga and his colleagues[29–32] conducted DNA typings of highly polymorphic human leukocyte antigen loci for human populations of the Japanese Archipelago, and found a clustering of the Ainu and the Ryukyuan populations.

The situation surrounding DNA polymorphism study of human population drastically changed with the sequencing of the human genome.[33] A genome-wide cataloging of single-nucleotide polymorphisms (SNPs) were carried out for four human populations including the Japanese living in Tokyo.[34] Nishida *et al.*[35] also conducted an SNP typing of 400 Mainland Japanese individuals, and Yamaguchi-Kabata *et al.*[36] produced and studied SNP data for 7000 individuals living in the various locations of the Japanese Archipelago. They clearly demonstrated the genetic differences between the Mainland Japanese and the Ryukyuan populations. Worldwide surveys of SNP typing were carried out under the Human Genome Diversity Project (HGDP-CEPH)[37] including the Mainland Japanese, and the human genetic diversity was analyzed for many human populations in Asia under the HUGO Pan-Asian SNP Consortium,[38] including the Mainland and Ryukyu Japanese. However, the Ainu individuals were not included in any of these studies.

One of us (Keiichi Omoto) carried out a series of studies on the genetic polymorphism of many human populations in Asia, and DNA samples have been preserved under his and Momoki Hirai's leadership at the University of Tokyo Kashiwa Campus. We recently formed 'Asian Archival DNA Repository Consortium' to keep and utilize these precious DNA samples. This paper is an outcome of activities of this Consortium. The availability of archival DNA in the Ainu population as well as the advancement of high throughput SNP genotyping for this study allowed us to examine in closer detail the genetic substructure and the evolutionary history of the human populations in the Japanese Archipelago. We conducted both individual-based analyses and population-based analyses, and will suggest that the two major waves of migrations during the paleolithic–Jomon period and the Yayoi and the later period produced the unique features of human populations in the Archipelago.

MATERIALS AND METHODS

Sample data, ethical approval and SNP genotyping

Blood samples of the Ainu people were collected in Biratori Town, Hidaka District of Hokkaido in the early 1980s for analysis of DNA by a group from The University of Tokyo, and have since been archived there. A study to use these samples was approved by the Research Ethics Committee of The University of Tokyo. These Ainu DNA samples were the same used in previous studies on mitochondrial DNAs,[20,21,23,24] on Y chromosome,[24,26,28] and on human leukocyte antigen types.[29–32] Two of us (Naruya Saitou and Timothy Jinam) recently visited Biratori Town, and explained these molecular anthropological studies in the past as well as the current study to the representatives of the Ainu people living in that area. DNA samples of the Ryukyuan from the Okinawa main island were collected from 2004 to 2008 by a group from University of the Ryukyus. A study to use these samples was approved by the Ethics Committee of University of the Ryukyus.

A total of 36 Ainu and 38 Ryukyuan samples were genotyped using the Affymetrix genome-wide SNP 6.0 microarray platform (Affymetrix, Santa Clara, CA, USA). All genotyping experiments and their computational analyses were conducted at the Department of Human Genetics, The University of Tokyo. In addition to the Ainu and the Ryukyuan populations, SNP genotype data, generated using the same method, from 200 Mainland Japanese (first set) mostly from the Kanto area[35] were used. These three groups (Ainu, Ryukyuan and Mainland Japanese) form the Japanese Archipelago population data set, which was further augmented with HapMap data[34] from four populations, namely Yorubans from Africa, Americans of European origins, Han Chinese from Beijing (CHB) and Japanese from Tokyo (JPT). The list of these seven populations is shown in Table 1.

Data filtering and quality checks

SNPs of the mitochondrial DNA, X chromosome and Y chromosome were excluded from the initial SNPs numbering 906 600. Duplicate SNPs and those without a dbSNP ID were also filtered out, resulting in a total of 868 257 remaining SNPs. Individual samples with poor genotyping performances were further filtered out based on the Affymetrix contrast quality control (cQC) threshold of 0.04, as recommended by the manufacturer. Three Ryukyuan and two Mainland Japanese samples were omitted based on this criterion. However, in the Ainu population, only 13 out of 36 individuals passed the cQC threshold. This was probably due to the degradation of DNA quality of the archival samples. To maximize the number of Ainu individuals to be used for downstream analysis, further SNP filtering was done based on confidence scores for each SNP generated during genotype calling using the Affymetrix Birdseed Ver2 algorithm.

In general, SNPs with a confidence score >0.1 are more likely to have failed genotype calling (that is 'no calls'). By visually inspecting genotype cluster graphs of random SNPs with confidence scores ranging from 0.1 to 0.004, a more stringent cutoff of 0.008 was used to exclude under-performing SNPs while retaining the maximum number of individuals. Thus, based on this criterion, 212 448 SNPs were omitted from the set of 36 Ainu individuals, resulting in 656 237 remaining SNPs. The SNP data in the Ainu and all other populations that were generated using the Affymetrix Genome-Wide 6.0 Assay were further filtered to remove SNPs with call rate $<95\%$ and that deviate from Hardy–Weinberg equilibrium ($P<0.001$). For example, 449 SNP loci were further excluded from the Ainu data, resulting in 655 788 SNP loci (see Supplementary Table S1). The filtering steps were done on each population separately and the number of SNPs filtered out is shown in Supplementary Table S1. After merging SNP data from all populations, the final number of SNPs in the seven population data sets was 641 314.

Merging with other population data

We also included SNP data from 30 other East Asian populations available from public databases in addition to the Japanese Archipelago and HapMap population data sets. These included the HGDP-CEPH data set, which consists of 650 000 SNPs from 51 worldwide populations[37] and the Pan-Asian SNP (PASNP) data set, which consists of 54 794 SNPs from 73 populations in Asia.[38] The number of SNPs that overlap between the Japanese Archipelago-HapMap 7-population data sets with the HGDP-CEPH panel was 114 001. After applying filters (excluding SNPs with $<95\%$ genotype call rate and minor allele frequency $<1\%$) in the merged data set, there were 101 562 SNPs remaining. For merging the Japanese Archipelago-HapMap population data sets with the PASNP data, 15 526 overlapping SNPs from both data sets were extracted and merged. After applying the same filtering criteria as above, the number of remaining SNPs was 14 997. The combination of the Japanese Archipelago, HapMap, HGDP-CEPH and PASNP data sets yielded only 4237 overlapping SNPs. All filtering and merging steps were carried out using PLINK software.[39] Supplementary Table S2 shows the list of 16 and 14 populations used from HGDP-CEPH and PASNP data sets, respectively.

Data analysis

Subsequent analysis was carried out using different combinations of the above data sets. For the merged data from all data sets, only populations in East Asia (Supplementary Table S2) were used for analysis. We first conducted individual-based analyses. Principal component analysis (PCA), using the smartpca program in the EIGENSOFT software package,[40] was our main strategy. The program frappe[41] was used to represent alternative views of population structure and admixture patterns. A maximum-likelihood method is used in this program, and it is computationally more efficient than STRUCTURE.[42] Population-based phylogenetic trees were constructed by using CONTML (a maximum-likelihood method for allele frequency data[43] was used) for SNP allele frequency data and NEIGHBOR (the neighbor-joining method[12] was used) for D_{st} distance[17] matrices computed from SNP allele frequency data by using GENDIST programs, all from the PHYLIP package[44] with 5000 bootstrap replicates. Neighbor-net networks[45] were also constructed from D_{st} distance matrices using the software SplitsTree 4.[46]

RESULTS

Individual-based analyses based on over 640 000 SNP data

The PCA result for the individual SNP data of the seven populations listed in Table 1 is shown in Supplementary Figure S1. The African, the West Eurasian and the East Eurasian (five East Asian populations) were located at the apexes of the triangle in this figure. Because African populations are known to be more distantly related from the two Eurasian populations,[47–49] the second PCA was conducted after excluding the African population (Supplementary Figure S2). Now the first principal component (PC1) separates the West Eurasian and the five East Asian populations, whereas the PC2 separates the Ainu population and the remaining four East Asian populations. Interestingly, the Ainu individuals are linearly aligned, suggesting varying degrees of recent admixture with the mainland Japanese population. It also appears that the population genetically closest to the Ainu is the Ryukyuan, despite the fact that these two populations are geographically located on the northernmost and southernmost poles of the Japanese archipelago, respectively.

When the West Eurasian individuals were further eliminated in the third PCA (Figure 1a), the unique feature of the Ainu individuals and the relationship of the other three populations became prominent. The PC1 and PC2 of this figure explained 1.8% and 0.6%, respectively, of the total variances among the SNP data of 356 individuals. The Mainland Japanese individuals and the HapMap Japanese (JPT) individuals clustered together, as expected. Many Ainu individuals were located at the left side and the Han Chinese in Beijing (HapMap CHB) were distributed at the rightmost side, whereas the Ryukyuan and the Mainland Japanese were sandwiched between them. The PC1 and PC2 coordinates were well correlated with Ryukyuan, Mainland Japanese and Han Chinese individuals. This pattern suggests the existence of the south (Ryukyuan) to the north (Han Chinese in Beijing) geographical cline. However, Ainu individuals, distributed in the northernmost Japanese Archipelago, were found to be closer to the Ryukyuan individuals than to the Mainland Japanese, consistent with the result shown in Supplementary Figure S2.

This finding clearly supports the genetic similarity between the Ainu and the Ryukyuan, in spite of their large geographical distance with each other within the Japanese Archipelago.

Another interesting pattern was the substantial interindividual variation among the Ainu individuals compared with those of the other three populations. Three Ainu individuals were within the cluster of the Mainland Japanese population, whereas the other five Ainu individuals in a red circle constituted a distinct cluster.

Table 1 Basic information of three populations in the Japanese Archipelago and the four populations of the HapMap data set

Population	*No. of individuals*	*Sampled location*	*Average heterozygosity*	*Reference*
Ainu	36	Hokkaido, Japan	0.220	This study
Ryukyuan	35	Okinawa, Japan	0.238	This study
Mainland Japanese	198	Tokyo, Japan	0.242	Ref. 35
Japanese (JPT)	45	Tokyo, Japan	0.240	Ref. 34
Han Chinese (CHB)	42	Beijing, China	0.242	Ref. 34
European (CEU)	89	Utah, USA	0.265	Ref. 34
Yoruban (YRI)	89	Ibadan, Nigeria	0.284	Ref. 34

Abbreviations: CEU, Americans of European origins; CHB, Han Chinese from Beijing; JPT, Japanese from Tokyo; YRI, Yorubans from Nigeria.

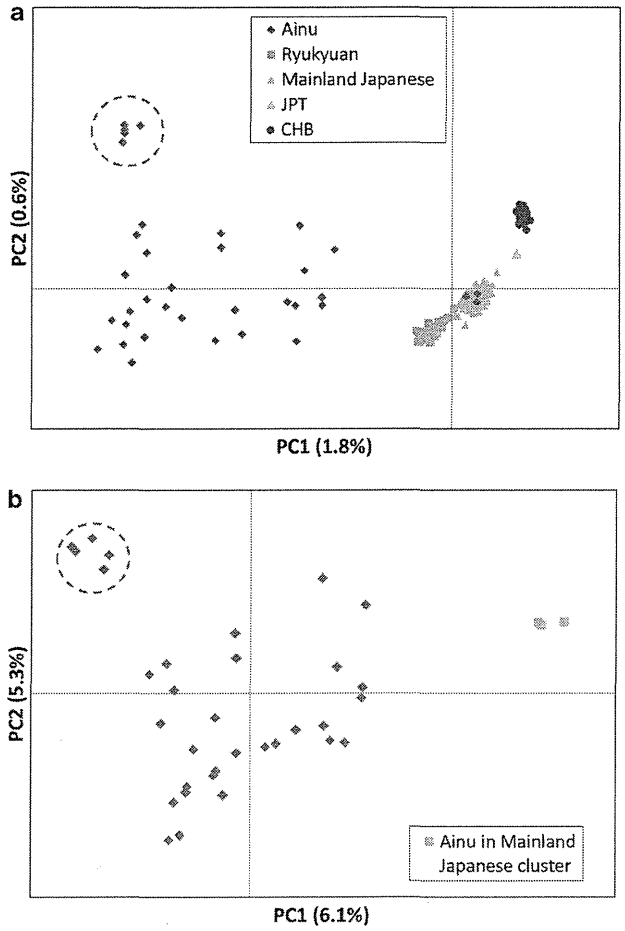


Figure 1 (**a**) A PCA plot of individuals for the three populations in the Japanese Archipelago (Ainu, Ryukyuans and the Mainland Japanese), HAPMAP Japanese (JPT) and HAPMAP Chinese (CHB). (**b**) APCA plot for 36 Ainu individuals only.

To further examine this pattern, PCA was performed only for the 36 Ainu individuals (Figure 1b). Now, the PC1 and PC2 explained 6.1% and 5.3% of the total variances, respectively. Interestingly, the high heterogeneity similar to that of Figure 1a was reproduced. This finding indicates that the wide variation observed for the Ainu individuals in Figure 1a was not because of the coexistence of other three populations, but because of the interindividual relationship inherent to the Ainu population. This result may be caused by recent admixtures involving two different source populations.

To see whether this was the case, we calculated the allele-sharing distances between the Ainu and the Mainland Japanese individuals, and compared these with the PC1 coordinates of the Ainu individuals. There was a clear positive correlation ($r^2 = 0.542$) between the allele-sharing distances and the PC1 coordinates (Supplementary Figure S3). The three Ainu individuals within the Mainland Japanese cluster had the smallest allele-sharing distances, and conversely, the Ainu individuals located farthest from the Mainland Japanese on the PC1 axis tend to have greater allele-sharing distances with the Mainland Japanese population. These observations suggest that the allele sharing between the Ainu and the Mainland Japanese was the result of relatively recent and continuing episodes of gene flow between the two populations. If this is the case, another high degree

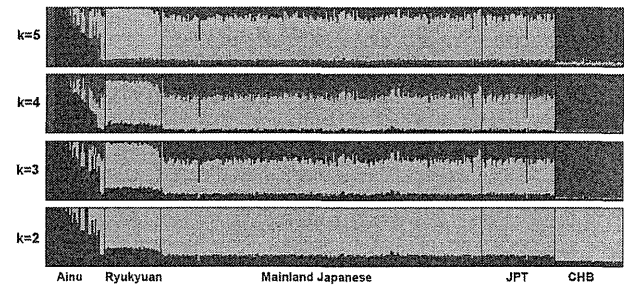


Figure 2 The result of frappe analysis from $k=2$ to $k=5$ for the same individuals of five populations used for the PCA analysis shown in Figure 1a.

of variation within the Ainu individuals regarding PC2 may also be explained by the gene flow from a human population genetically distinct from the Mainland Japanese. In particular, the five Ainu individuals in the red circle having the highest PC2 coordinates may be affected by recent admixture events.

The result of the frappe analysis is shown in Figure 2. When $k=2$, the two ancestry components corresponds to one (dark blue), which is 100% in 16 Ainu individuals and the other (orange), which is the highest ($\sim$90%) in all the CHB individuals. Interestingly, 20 Ainu individuals showed varied proportions of the orange-colored ancestry component, again suggesting the recent admixture. Most of the Ryukyuan and the Mainland Japanese individuals showed about 30% and 20% blue-colored component, respectively. This difference was consistent with the PCA results in Figure 1a, in that the Ainu was closer to the Ryukyuan than to the Mainland Japanese. The three Ainu individuals who were within the Mainland Japanese cluster in Figure 1a are located at the rightmost columns of the Ainu population, and they also showed $\sim$20% blue-colored component, similar to the Mainland Japanese.

As k was increased to 3, the orange-colored component at $k=2$ divided into two (orange and magenta). Now, all the CHB individuals are almost full of magenta ancestry component, and the Mainland Japanese and JPT individuals also consist of $\sim$30% magenta-colored ancestry component. There are two Japanese individuals, one in the Mainland Japanese and the other in HapMap JPT, who have >70% magenta component. These two individuals are also outliers in the PCA analysis shown in Figure 1a. If we consider blue- and magenta-colored ancestry components as the Jomon and the Yayoi factors, the intermediate orange-colored component is not easy to comprehend. It is possible that this does not correspond to a real ancestral population, but an artificially inferred component corresponding to the long-term admixture between the Jomon and the Yayoi genetic components. At $k=4$, the five outlier Ainu individuals observed in the PCA plot (those in red circles in Figures 1a and b) were differentiated from the rest of the Ainu, as indicated in the purple color. The Ryukyuan-specific ancestral component appeared at $k=5$. This is again not easy to interpret, and could be an artificially inferred component.

Generally speaking, the frappe results appear to be consistent with the PCA analysis in terms of the two patterns: (1) varying amounts of admixture in the Ainu with the Mainland Japanese and (2) the possible presence of another source population, which contributed to the genetic structure of the Ainu. A high correlation between PC1 coordinates of Figure 1a and proportions of the blue-colored ancestry from the frappe analysis ($k=2$) shown in Supplementary Figure S4 confirms the pattern (1). The pattern (2) can be supported by a high

correlation between the purple component frequencies for $k=4$ and PC2 coordinates in Figure 1a for the Ainu individuals, as shown in Supplementary Figure S5.

Individual-based analyses merged with HGDP-CEPH and PASNP data sets

We now move to analyses combined with individuals belonging to the 16 HGDP-CEPH populations. Figure 3a shows the PCA result. The overall distribution of individuals in this figure indicates an L shape. Ainu and Yakut individuals are at the two extremes. The Ryukyuan and the Mainland Japanese populations were located as if they were pulled by the Ainu population, whereas the remaining East Asian populations were closer to northern East Asian populations (Yakut, Mongolian and Oroqen). This clear dichotomy with the L-shaped constellation remains when we further added individuals of the 14-population PASNP data sets (Figure 3b). The Ainu population is located at the one extreme followed by the Ryukyuan, whereas another extreme is now the Uyghur population. The Uyghurs are known to be an admixed population with West Eurasians, so as the Yakuts to a lesser extent.[38] This fact indicates that the Ainu, and the two other populations in the Japanese Archipelago (Ryukyuan and Mainland Japanese) to a lesser degree, contain genetic components neither found in the other East Eurasians nor West Eurasians. Another notable addition in Figure 3b is the Korean, who are located between the Mainland Japanese and the Han Chinese in the northern part of China. This result is consistent with that of Tian *et al.*[50]

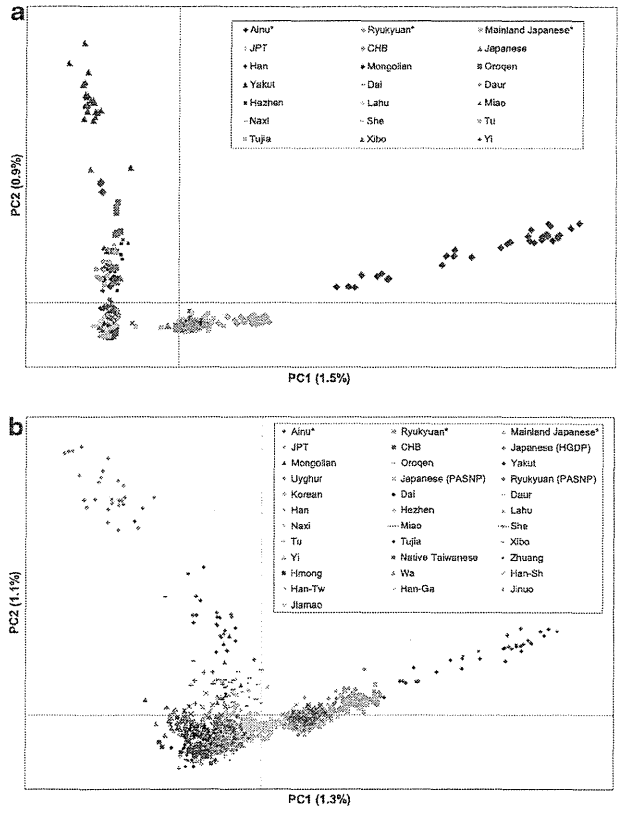


Figure 3 PCA plots of the individuals for the three populations in the Japanese Archipelago (Ainu, Ryukyuans and the Mainland Japanese) and other Asian populations. (**a**) Result with the HGDP-CEPH data set (see Supplementary Table S2 for the list of populations). (**b**) Result with the HGDP-CEPH data set and the PASNP data set (see Supplementary Table S2 for the list of populations).

The result of the frappe analysis for populations corresponding to Figure 3a is shown in Supplementary Figure S6. When $k=2$, the dichotomy pattern is quite similar with that of Figure 2, and additional East Eurasian populations showed similar blue/orange frequencies as in the Han Chinese. As k increased to 3, the light-blue ancestry component, which was almost exclusively found in the Yakut individuals, diverged from the orange component under $k=2$. The green ancestry component further divided from the orange one at $k=4$, and this was dominant in the Ryukyuan and the Mainland Japanese. The five 'outlier' Ainu individuals in Figure 1 were fully covered by the violet ancestry component at $k=5$. The fraction of this component was high only in the Ainu. The same situation continues from $k=2$ to $k=6$ for the blue ancestry component. This result clearly indicates a unique position of the Ainu population in East Asia. Supplementary Figure S7 shows the result of the frappe analysis for populations corresponding to Figure 3b. The overall pattern is similar to that of Supplementary Figure S6.

Population-based analysis

Individual-based analyses presented in the previous section deciphered the complex structure of each population, especially for the Ainu. Populations are still realistic units of modern human evolution, and we thus constructed phylogenetic trees and networks of populations.

We first used SNP data of the Ainu, the Ryukyuan, the Mainland Japanese and the Han Chinese in Beijing (see Table 1), and constructed an unrooted maximum-likelihood tree (Supplementary Figure S8). The Ainu and the Ryukyuan populations clustered with 100% bootstrap probability, and this is consistent with the PCA result (Figure 1a). The overall pattern is similar to the tree shown by Omoto and Saitou[16] based on only 25 polymorphic loci of classic genetic markers, if we equate the Han Chinese in Beijing in the tree of Supplementary Figure S8 with the Korean in their tree; the branch going to the Ainu population is quite long compared with that to the Ryukyuan, whereas the branch going to the Mainland Japanese had zero length.

A neighbor-joining tree was then constructed for the 29 human populations in East Asia, using the merged SNP data with the HGDP-CEPH and PASNP data sets to see the relationships between the three populations in the Japanese Archipelago with other worldwide populations (Figure 4a). Three populations (Japanese, Mainland Japanese and Ryukyu) listed in Supplementary Table S2 were excluded to avoid redundancy of populations. The Korean population is now phylogenetically closest to the human populations in the Japanese Archipelago, though the bootstrap probability to support the clustering of these four populations was 94% (bootstrap value not shown). Four populations (Hezhen, Daur, Oroqen, and Mongolians) that distribute in the northeast Asia as well as the Yakut, the Xibo and the Uyghur formed one cluster in this tree, whereas three populations (Tu, Naxi and Yi) in the southern China formed one cluster, and these two clusters are phylogenetically closer to the Korean and the Japanese Archipelago cluster. In terms of genetic distances, however, some Han Chinese populations (CHB in Beijing and Han-Tw) were closer to populations in the Japanese Archipelago. These smaller genetic distances may be contributed to a smaller random genetic drift caused by large population sizes of these Han populations.

We then selected 14 populations in Figure 4a and constructed a neighbor-joining tree (Figure 4b). Because sample sizes of many ethnic minorities in China were small, we merged the data for

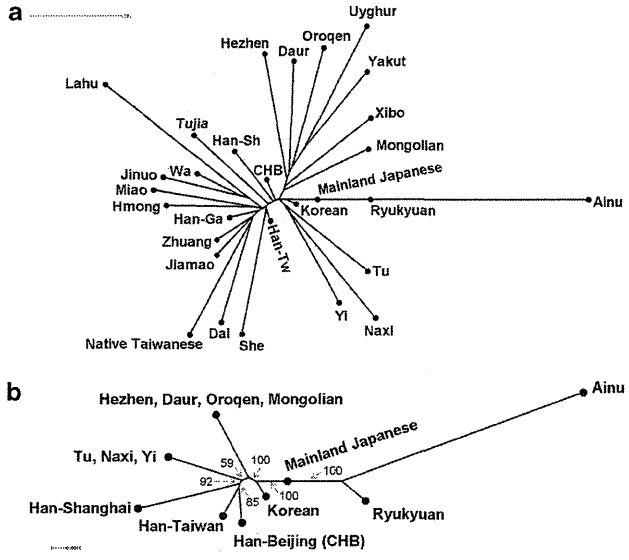


Figure 4 Phylogenetic trees for the three Japanese populations and other Asian populations. (**a**) A neighbor-joining tree for the three Japanese populations and other Asian populations listed in Supplementary Table S2. (**b**) A maximum-likelihood tree for the three Japanese populations with Northeast Asians (Hezhen, Daur, Oroqen, Mongolian), Koreans, Han Chinese and populations from central China (Tu, Naxi, Yi). All bootstrap values shown with arrows were obtained from 5000 replications.

Hezhen, Daur, Oroqen and Mongolians as the northeast Asian population, and three populations in southern China (Tu, Naxi and Yi) were also merged. The Ainu and the Ryukyuan were clustered with 100% bootstrap probability, followed by the Mainland Japanese. The three populations in the Japanese Archipelago clustered with the Korean with 100% bootstrap probability. The Ainu population has a long branch and is clearly different from the other populations in this figure, confirming the unique phylogenetic position of this population in East Asia. The very short, almost nonexistent branch leading to the Mainland Japanese, as well as its intermediate position between the Ainu-Ryukyu and the cluster for the remaining populations suggest that the Mainland Japanese was formed by the result of admixture between these two ancestral population sources, symbolized as the Jomon and the Yayoi. The northeast Asian population was phylogenetically closest to the Korean–Japanese Archipelago population cluster, followed by the populations in southern China. The three Han Chinese populations in Beijing, Taiwan and Shanghai clustered together with 92% bootstrap probability.

Phylogenetic networks were also constructed for the same genetic distance matrices used for constructing trees in Figures 4a and b, as shown in Supplementary Figures S9 and S10, respectively. The overall pattern of Supplementary Figure S9 was similar to that of Figure 4a, except for the intermediate position of the Yakut, which was not only close to the Uyghur but also close to the four northeast Asian populations (Oroqen, Daur, Hezhen and Mongolian). There are two interesting reticulations in Supplementary Figure S10. Although the southern Chinese populations are close to the three Han Chinese population cluster, they are also close to the northeast Asian population. The Korean is also located in an intermediate position, close to the Japanese Archipelago populations, but also phylogenetically close to the Han Chinese.

DISCUSSION

Genetic heterogeneity of the Ainu population

The s.d. (8.3×10^{-3}) of allele-sharing distances for the Ainu population is ~10 times higher than those (0.75×10^{-3} and 0.85×10^{-3}) for the Ryukyuan and the Mainland Japanese, respectively. One possible factor for this variation may be degradation of DNA quality of the Ainu individual samples after preservation of almost 30 years. We thus classified 36 Ainu individuals into two categories (high or low) in terms of Affymetrix Contrast QC (CQC) values. Supplementary Figure S11 shows the distribution of these two categories of individuals on the PCA plot shown in Figure 1a. There is no clear correlation between these two categories and the locations of the Ainu individuals. Furthermore, we performed PCA separately using 13 Ainu individuals who passed the Affymetrix CQC threshold, and the remaining 23 who did not pass the threshold. Supplementary Figure S13 shows that in both cases, the same scattered pattern was observed. We also examined effects of SNP genotyping call heterogeneity among SNP loci. Because the genotyping call rate threshold of 95% was used to produce the PCA plots shown in Figure 1a, we used 90 and 100% call rates and conducted PCA, as shown in Supplementary Figures S12A and S12B. These three PCA results are essentially the same. We further retrieved the top 100 SNPs contributing to PC1 and PC2 in Figure 1a, and a total of 200 SNPs were obtained. After pruning for linkage disequilbrium, 109 SNP loci were left. We used the Digi Tag2 method[51] for 96 SNP loci randomly selected from the 109 SNP loci for technical simplicity. The DNA samples used were 36 Ainu, 33 Ryukyuan and 22 Mainland Japanese. SNP call rates for Ainu, Ryukyuan and Mainland Japanese were 0.950, 0.985 and 0.999, respectively. Although the call rate for Ainu was somewhat lower than those for other two populations, the proportions of identical SNP typing with those obtained by using Affymetrix ver. 6.0 were 99.4%, 99.8% and 99.9% for Ainu, Ryukyuan and Mainland Japanese, respectively. We thus conclude that SNP genotypes estimated by using Affymetrix ver. 6.0 were almost identical with those estimated by using the Digi Tag2 method for all the three populations. These results indicate that the DNA degradation was not the reason for the variation among the Ainu individuals.

Interestingly, the average heterozygosity (0.220) of the Ainu population was the least among the seven populations listed in Table 1, whereas those of the other four East Asian populations were more or less the same (~0.24). This feature is not consistent with a high degree of variation on allele-sharing distances for the Ainu population, if it is a panmictic and isolated population with no gene flow with the surrounding populations. As we showed in this study, however, the Ainu population seems to have experienced gene flow with two different populations, the Mainland Japanese and the yet unknown population in the north. If gene flow happens to occur, the heterogyzosity of admixed individuals should be in the middle of those of the two original populations. To see whether this is the case, the 28 Ainu individuals (three individuals within the Mainland Japanese and five outliers in the red circle in Figure 1a were not included) were divided into two; more admixed and less admixed. When the Ainu individuals were separated according to the normalized PC1 coordinate 0.75 of Supplementary Figure S4, the average heterozygosity of the more admixed 13 Ainu individuals was 0.223, whereas that for the less admixed 15 Ainu individuals was 0.201. When Ainu individuals were separated according to the normalized PC2 coordinate 0.3 of Supplementary Figure S5, average heterozygosities of the less admixed (18), more admixed (10) and outlier (5) Ainu individuals became 0.213, 0.201 and 0.163, respectively. In both cases, the putative admixed Ainu individuals showed the intermediate