why Tax is dispensable for ATL maintenance and why ATL does not use the "oncogene addiction" model of transformation? Will findings of virus-induced senescence, checkpoint inactivation, and cellular genetic damage meld into a coherent chronology of transformation events? In addition to these questions, equally important topics challenge immunologic and clinical HTLV-1 research. Indeed, one can look with anticipation to many intriguing answers that await revelation over the next 30 years.

Acknowledgments

Received 7/10/2007; revised 7/30/2007; accepted 9/10/2007.

References

1. Takatsuki K. Discovery of adult T-cell leukemia. Retrovirology 2005;2:16.
2. Gallo RC. The discovery of the first human retrovirus: HTLV-1 and HTLV-2. Retrovirology 2005;2:17.
3. Yoshida M. Discovery of HTLV-1, the first human retrovirus, its unique regulatory mechanisms, and insights into pathogenesis. Oncogene 2005;24:5931–7.
4. Matsuoka M, Jeang KT. Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. Nat Rev Cancer 2007;7:270–80.
5. Kwaan N, Lee TH, Chafets DM, et al. Long-term variations in human T lymphotropic virus (HTLV)-I and HTLV-II proviral loads and association with clinical data. J Infect Dis 2006;194:1557–64.
6. Asquith B, Zhang Y, Mosley AJ, et al. *In vivo* T lymphocyte dynamics in humans and the impact of human T-lymphotropic virus 1 infection. Proc Natl Acad Sci U S A 2007;104:8035–40.
7. Komori K, Hasegawa A, Kurihara K, et al. Reduction of human T-cell leukemia virus type 1 (HTLV-1) proviral loads in rats orally infected with HTLV-1 by reimmunization with HTLV-1-infected cells. J Virol 2006;80:7375–81.
8. Hasegawa H, Sawa H, Lewis MJ, et al. Thymus-derived leukemia-lymphoma in mice transgenic for the Tax gene of human T-lymphotropic virus type I. Nat Med 2006;12:466–72.
9. Ohsugi T, Kumasaka T, Okada S, Urano T. The Tax protein of HTLV-1 promotes oncogenesis in not only immature T cells but also mature T cells. Nat Med 2007;13:527–8.

VOLUME 25 · NUMBER 34 · DECEMBER 1 2007

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

VCAP-AMP-VECP Compared With Biweekly CHOP for Adult T-Cell Leukemia-Lymphoma: Japan Clinical Oncology Group Study JCOG9801

Kunihiro Tsukasaki, Atae Utsunomiya, Haruhiko Fukuda, Taro Shibata, Takuya Fukushima, Yoshifusa Takatsuka, Shuichi Ikeda, Masato Masuda, Haruhisa Nagoshi, Ryuzo Ueda, Kazuo Tamura, Masayuki Sano, Saburo Momita, Kazunari Yamaguchi, Fumio Kawano, Shuichi Hanada, Kensei Tobinai, Masanori Shimoyama, Tomomitsu Hotta, and Masao Tomonaga

From the Nagasaki University, Nagasaki; Imamura Bun-in Hospital; Kagoshima University, Kagoshima; National Cancer Center Research Institute; National Cancer Center Hospital, Tokyo; Sasebo City General Hospital, Sasebo; Ryukyu University, Nishihara; St Marianna University, Yokohama; Nagoya City University; National Hospital Organization Nagoya Medical Center, Nagoya; Fukuoka University, Fukuoka; Saga University, Saga; National Hospital Organization Nagasaki Medical Center, Ohmura; Kumamoto University; and National Hospital Organization Kumamoto Medical Center, Kumamoto, Japan.

Submitted April 20, 2007; accepted September 6, 2007; published online ahead of print at www.jco.org on October 29, 2007

Supported by Grants-in-Aid No. 2S-1, 5S-1, 8S-1, 11S-1, 14S-1, 17S-1, 1-1, 4-5, 7-29, and 9-10 from the Cancer Research from the Ministry of Health, Labor and Welfare of Japan (1990 to present), for the Second-Term Ten-Year Strategy for Cancer Control from the Ministry of Health and Welfare (1994 to 2004) and for Basic Research from the Science and Technology Agency (1991 to 1993).

Presented in part at the 47th Annual Meeting of the American Society of Hematology, December 10-13, 2005, Atlanta, GA.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Address reprint requests to Kunihiro Tsukasaki, MD, PhD, Nagasaki University Graduate School of Biomedical Science, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan; e-mail: tsukasak@net.nagasaki-u.ac.jp.

© 2007 by American Society of Clinical Oncology

0732-183X/07/2534-5458/$20.00

DOI: 10.1200/JCO.2007.11.9958

ABSTRACT

Purpose

Our previous phase II trial for treating human T-lymphotropic virus type I–associated adult T-cell leukemia-lymphoma (ATLL) with vincristine, cyclophosphamide, doxorubicin, and prednisone (VCAP), doxorubicin, ranimustine, and prednisone (AMP), and vindesine, etoposide, carboplatin, and prednisone (VECP) showed promising results. To test the superiority of VCAP-AMP-VECP over biweekly cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP), we conducted a randomized controlled trial exclusively for ATLL.

Patients and Methods

Previously untreated patients with aggressive ATLL were assigned to receive either six courses of VCAP-AMP-VECP every 4 weeks or eight courses of biweekly CHOP. Both treatments were supported with granulocyte colony-stimulating factor and intrathecal prophylaxis.

Results

A total of 118 patients were enrolled. The complete response (CR) rate was higher in the VCAP-AMP-VECP arm than in biweekly CHOP arm (40% *v* 25%, respectively; $P = .020$). Progression-free survival rate at 1 year was 28% in the VCAP-AMP-VECP arm compared with 16% in the CHOP arm ($P = .100$, two-sided $P = .200$). Overall survival (OS) at 3 years was 24% in the VCAP-AMP-VECP arm and 13% in the CHOP arm ($P = .085$, two-sided $P = .169$). For VCAP-AMP-VECP versus biweekly CHOP, grade 4 neutropenia, grade 4 thrombocytopenia, and grade 3 or 4 infection rates were 98% *v* 83%, 74% *v* 17%, and 32% *v* 15%, respectively. There were three toxic deaths in the VCAP-AMP-VECP arm.

Conclusion

The longer OS at 3 years and higher CR rate with VCAP-AMP-VECP compared with biweekly CHOP suggest that VCAP-AMP-VECP might be a more effective regimen at the expense of higher toxicities, providing the basis for future investigations in the treatment of ATLL.

J Clin Oncol 25:5458-5464. © 2007 by American Society of Clinical Oncology

INTRODUCTION

Adult T-cell leukemia-lymphoma (ATLL) is a distinct peripheral T-lymphocytic malignancy associated with human T-cell lymphotropic virus type I.[1-3] The diverse clinical features of this disease have led to its classification.[4] Aggressive ATLL (ie, acute or lymphoma type) has usually been treated as a subtype of aggressive non-Hodgkin's lymphoma (NHL), whereas indolent ATLL (ie, chronic or smoldering type) has been managed as a subtype of chronic lymphoid leukemia.[5-7] Aggressive ATLL generally has a poor prognosis compared with aggressive B-cell lymphoma and peripheral T-cell lymphoma excluding ATLL.[7-9] Median survival time (MST) of patients with aggressive ATLL is approximately 8 months because of the multidrug resistance (MDR) phenotype of malignant cells, rapid proliferation of the cells, a large tumor burden with multiorgan failure, hypercalcemia, and/or frequent infectious complications.[4-7,10]

In the two previous multicenter trials for advanced NHL, Japan Clinical Oncology Group (JCOG) 8101 (1981 to 1983) and JCOG8701 (1987 to 1991) evaluating the efficacy of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) –like regimens and a multiagent regimen of the second generation, respectively, a significantly

Information downloaded from jco.ascopubs.org and provided by TOKYO DAIGAKU IKAGAKU KENKYUJO on March 22, 2009 from 157.82.101.48.
Copyright © 2007 by the American Society of Clinical Oncology. All rights reserved.

shorter survival time was demonstrated for ATLL patients than for other NHL patients.[11,12] Thus, the first trial exclusively applied to aggressive ATLL, JCOG9109, was started (1991 to 1993). The chemotherapy protocol involved the use of deoxycoformycin, an inhibitor of adenosine deaminase, which was found to be effective against refractory ATLL.[13] However, there were no improvements in overall response rate (ORR) or survival time compared with the previous trials.[14]

The next phase II study (JCOG9303, 1994 to 1996), with the chemotherapy protocol LSG15 against aggressive ATLL consisting of a dose-intensified multiagent chemotherapy with vincristine, cyclophosphamide, doxorubicin, and prednisone (VCAP), doxorubicin, ranimustine, and prednisone (AMP), and vindesine, etoposide, carboplatin, and prednisone (VECP) with granulocyte colony-stimulating factor (G-CSF) and intrathecal prophylaxis, showed promising results with complete response (CR) and partial response rates of 36% and 45%, respectively, and an MST of 13 months at the expense of toxicities.[15] Ranimustine, an alkylating agent crossing the blood-brain barrier, and intrathecal prophylaxis with methotrexate and prednisone were incorporated because ATLL frequently involves the CNS.[4,16] Carboplatin and ranimustine were incorporated because the activity of these agents is not affected by the expression of P-glycoprotein, a product of the MDR gene *MDR1*, which is frequently expressed on ATLL cells.[17] The promising results of JCOG9303 prompted us to conduct a randomized controlled trial comparing the LSG15 regimen with biweekly CHOP against aggressive ATLL. Dose intensification of CHOP with prophylactic use of G-CSF was expected to improve survival among patients with aggressive NHL, and our randomized phase II study (JCOG9505) comparing biweekly CHOP with dose-escalated CHOP to treat aggressive NHL excluding ATLL revealed biweekly CHOP to be more promising.[18] Therefore, we regarded biweekly CHOP as a standard treatment for NHL including aggressive ATLL at the time of designing this phase III study.

PATIENTS AND METHODS

Patients

Previously untreated patients with aggressive ATLL (ie, acute-, lymphoma-, or unfavorable chronic-type ATLL) were eligible. Unfavorable chronic-type ATLL (defined by at least one of the following three factors: low serum albumin, high lactate dehydrogenase, or high blood urea nitrogen concentration) had an unfavorable prognosis similar to acute- or lymphoma-type ATLL.[6] The diagnosis of ATLL was made based on seropositivity for human T-cell lymphotropic virus type I and histologically and/or cytologically proven peripheral T-cell malignancy.

Eligibility criteria, which were identical to those for the previous studies JCOG9109 and JCOG9303, included no prior chemotherapy, age 15 to 69 years, preserved organ functions, and performance status (PS) of 0 to 3 or 4 as a result of hypercalcemia caused by ATLL.[14,15] The study protocol and the informed consent document were approved by both the JCOG Protocol Review Committee and the institutional review board of each institution.

Registration

Registration involved a telephone call or facsimile from the participating physicians to the JCOG Data Center, National Cancer Center, Tokyo, Japan. After an evaluation of eligibility, the patient was assigned to receive either modified (m) LSG15 or mLSG19 with a minimization method for balancing PS (0 or 1 *v* 2, 3, or 4) and institution.

Treatment

mLSG15 in JCOG9801 was a modified version of LSG15 in JCOG9303, consisting of the following three regimens: VCAP (vincristine 1 mg/m^2, maximum, 2 mg; cyclophosphamide 350 mg/m^2; doxorubicin 40 mg/m^2; and prednisone 40 mg/m^2) on day 1, AMP (doxorubicin 30 mg/m^2, ranimustine 60 mg/m^2, prednisone 40 mg/m^2) on day 8, and VECP (vindesine 2.4 mg/m^2 on day 15, etoposide 100 mg/m^2 on days 15 to 17, carboplatin 250 mg/m^2 on day 15, and prednisone 40 mg/m^2 on days 15 to 17) on days 15 to 17; the next course was started on day 29.[15] The modifications to mLSG15 compared with LSG15 were as follows: the total number of cycles was reduced from seven to six because of progressive cytopenia, especially thrombocytopenia, after repeating the VCAP-AMP-VECP therapy; and cytarabine 40 mg was used with methotrexate 15 mg and prednisone 10 mg for prophylactic intrathecal administration at the recovery phases of courses 1, 3, and 5 after confirmation of a platelet recovery of more than 70×10^9/L within 2 days before the next systemic chemotherapy because of the high frequency of CNS relapse in the JCOG9303 study.

mLSG19, a modified version of LSG19, consisted of eight cycles of CHOP (cyclophosphamide 750 mg/m^2; doxorubicin 50 mg/m^2; vincristine 1.4 mg/m^2, with a maximum of 2 mg, on day 1; and prednisone 100 mg on days 1 to 5) every 2 weeks.[18] The modification was an intrathecal administration identical to that in mLSG15.

Neutrophil count was checked twice a week for G-CSF use during the protocol treatment. When a serious infection occurred as a result of severe neutropenia, the doses of cyclophosphamide, doxorubicin, ranimustine, vindesine, etoposide, and carboplatin were decreased to 75% thereafter. If a second infection occurred, treatment was stopped.

Supportive Therapy

Supportive therapy for opportunistic infections was administered as in JCOG9303.[15] When the neutrophil count decreased to less than 1×10^9/L, G-CSF was administered subcutaneously every day until recovery to more than 5×10^9/L was achieved. Each course of mLSG19, VCAP in mLSG15, or AMP/VECP in mLSG15 was started after confirmation of a neutrophil count of more than 1.2×10^9/L, more than 1.0×10^9/L, or more than 0.5×10^9/L, respectively. Administration of G-CSF was discontinued on the day of chemotherapy and the day before. In cases when the hemoglobin level was less than 8 g/dL or platelet count was less than 20×10^9/L, an RBC or platelet transfusion was administered, respectively. Erythropoietin was not recommended for supportive care.

Response and Toxicity Evaluation

Response was judged using our own criteria for ATLL as described.[14,15] Toxicity was graded according to the JCOG toxicity criteria, an expanded version of the National Cancer Institute Common Toxicity Criteria version 1.0.[19]

Statistical Analysis

This trial was designed as a multicenter prospective randomized controlled trial. All analyses were performed on an intent-to-treat basis. The primary end point was overall survival (OS), and the secondary end points were progression-free survival (PFS), CR rate, and toxicity. The planned duration of accrual was 3 years, and the planned follow-up time was 2 years. The study was designed as a superiority trial, with the one-sided hypothesis according that the superiority of the control arm to the mLSG15 arm was out of concern a priori. This is because mLSG15 was expected to be associated with frequent and severe toxicities compared with the control arm.[15,18] The required sample size was 114 eligible patients in total, for 80% power to detect a hazard ratio of 0.6 under the assumption that survival times were exponentially distributed (corresponding to a 15% difference in the 3-year survival rate when the rate in the mLSG19 arm is 10%) with a one-sided type I error of 0.05. The planned sample size was 130 randomly assigned patients, with the expectation that 10% would be ineligible. The duration of accrual and the follow-up time were amended to 5 years and 1 year, respectively, in 2001 because of slow accrual.

OS was defined as the time from random assignment until death from any cause or until the last follow-up for patients who were alive. PFS was defined as the time from random assignment until death from any cause,

Information downloaded from jco.ascopubs.org and provided by TOKYO DAIGAKU IKAGAKU KENKYUJO on March 22, 2009 from 157.82.101.48.
Copyright © 2007 by the American Society of Clinical Oncology. All rights reserved.

relapse, or progressive disease or until the last follow-up for patients who were alive. The CR rate and ORR were defined as the proportion of patients with CR and with CR or partial response, respectively, of all randomly assigned patients. Survival estimates were calculated using the Kaplan-Meier method and compared by stratified log-rank test for all randomly assigned patients, with PS as a stratification factor. An analysis of adverse events was conducted for all patients who received the protocol treatment, whether partially or completely. As a sensitivity analysis, the Cox regression was carried out. In accordance with the hypothesis, all of the *P* values are presented as one sided, except for when explicitly stated as two sided. All analyses were performed with SAS software Release 8.2 (SAS Institute, Cary, NC).

The JCOG Data Center collected and managed case report forms. In-house interim monitoring for quality control was performed at the center, and the monitoring reports were semiannually submitted to and reviewed by the JCOG Data and Safety Monitoring Committee. One interim analysis was planned after half of the planned number of patients had been off treatment with an adjustment for multiplicity by the alpha-spending function of O'Brien-Fleming.[20]

RESULTS

Patient Characteristics

Between July 1998 and October 2003, 118 patients were enrolled from 27 participating institutions (Fig 1). In June 2001, an interim analysis was performed according to the protocol and did not meet the prespecified stopping criteria ($\alpha = .00022$), and the study was continued. The final analyses were performed in February 2005 based on the follow-up data from December 2004 ($\alpha = .04992$). Fifty-seven patients were assigned to the mLSG15 arm (VCAP-AMP-VECP), and the remaining 61 patients were assigned to the mLSG19 arm (biweekly CHOP). The characteristics of the 118 patients are listed in Table 1. Two patients, one from each arm, were deemed ineligible after random assignment because they were judged to have organ dysfunctions not caused by the invasion of ATLL cells by the case report form review. Age, sex, and subtypes of ATLL were well balanced between the arms. However, there were some imbalances in prognostic factors. Although patients were stratified by PS of 0 or 1 versus 2, 3, or 4 at random assignment, there was an imbalance between PS 0 and 1. PS 0 was more frequent in the biweekly CHOP arm than the VCAP-AMP-VECP arm. Also, bulky mass (> 5 cm in diameter) was less frequent in the biweekly CHOP arm. In contrast, "B" symptoms were more frequent in the biweekly CHOP arm.

Table 1. Characteristics of Randomly Assigned Patients

Characteristic	VCAP-AMP-VECP (n = 57)* No. of Patients	VCAP-AMP-VECP (n = 57)* %	Biweekly CHOP (n = 61)* No. of Patients	Biweekly CHOP (n = 61)* %
Age, years				
Median	56		58	
Range	36-69		33-69	
Sex				
Male	27	47	34	56
Female	30	53	27	44
Subtypes of ATLL				
Acute	40	70	41	67
Lymphoma	12	21	14	23
Unfavorable chronic	5	9	6	10
PS				
0	19	33	30	49
1	27	47	19	31
2	8	14	10	16
3	2	4	2	3
4	1	2	0	0
B symptoms				
Absent	39	68	34	56
Present	18	32	27	44
Bulky mass, cm				
< 5	36	63	49	80
≥ 5	17	30	9	15
≥ 10	4	7	3	5

Abbreviations: VCAP, vincristine, cyclophosphamide, doxorubicin, and prednisone; AMP, doxorubicin, ranimustine, and prednisone; VECP, vindesine, etoposide, carboplatin, and prednisone; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone; ATLL, adult T-cell leukemia-lymphoma; PS, performance status.
*Two patients, one in each arm, were ineligible because of organ dysfunction.

Response and Survival

Responses in all randomly assigned patients are listed in Table 2. The CR rate, including uncertified CR, was higher in the VCAP-AMP-

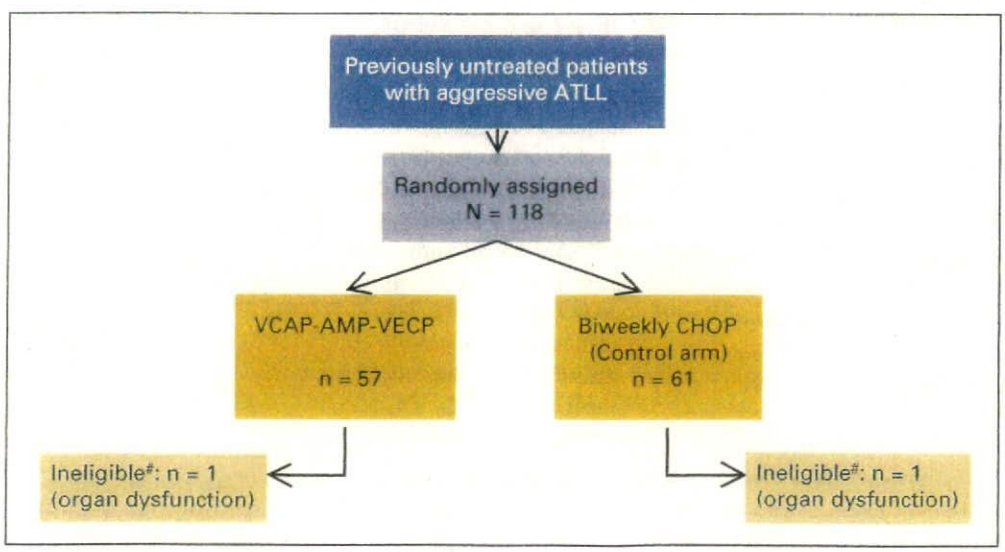

Fig 1. Enrollment and treatment of patients. (#) Included in the analysis. ATLL, adult T-cell leukemia-lymphoma; VCAP, vincristine, cyclophosphamide, doxorubicin, and prednisone; AMP, doxorubicin, ranimustine, and prednisone; VECP, vindesine, etoposide, carboplatin, and prednisone; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone.

Information downloaded from jco.ascopubs.org and provided by TOKYO DAIGAKU IKAGAKU KENKYUJO on March 22, 2009 from 157.82.101.48.
Copyright © 2007 by the American Society of Clinical Oncology. All rights reserved.

Table 2. Response to Treatment

Response	% of Patients: VCAP-AMP-VECP (n = 57)	Biweekly CHOP (n = 61)	P
CR	40	21	
CRu	0	3	
PR	32	41	
NR	9	16	
PD	18	16	
Not assessable	2	2	
CR + CRu	40	25	.020*
95% CI	27.6 to 54.2	14.5 to 37.3	
CR + CRu + PR	72	66	NS*
95% CI	58.5 to 83.0	52.3 to 77.3	

Abbreviations: VCAP, vincristine, cyclophosphamide, doxorubicin, and prednisone; AMP, doxorubicin, ranimustine, and prednisone; VECP, vindesine, etoposide, carboplatin, and prednisone; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone; CR, complete response; CRu, unconfirmed complete response; PR, partial response; NR, no response; PD, progressive disease; NS, not significant.
*Fisher's exact test (one sided).

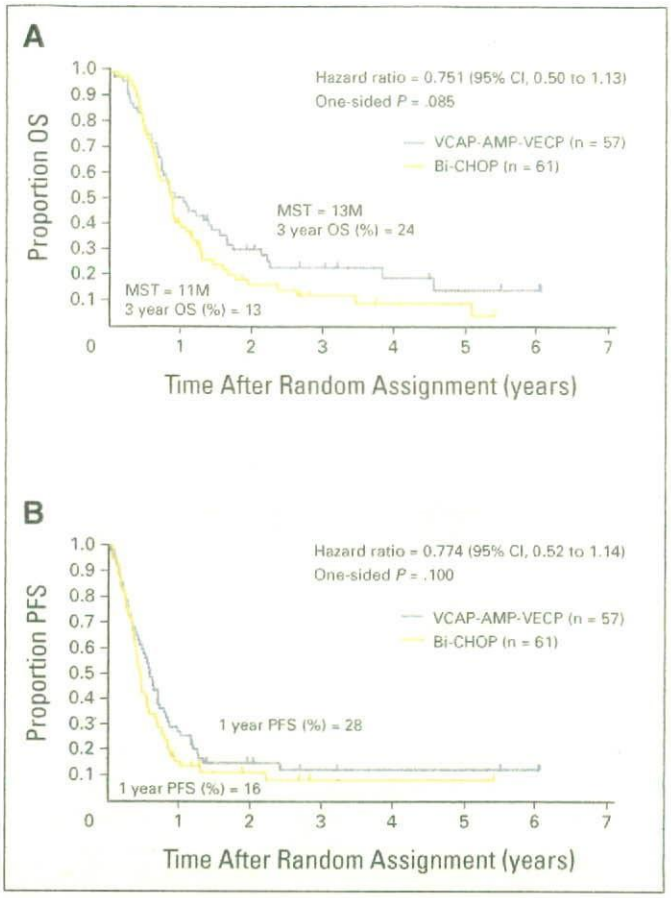

Fig 2. (A) Kaplan-Meier estimate of overall survival (OS) for all randomly assigned patients. (B) Kaplan-Meier estimate of progression-free survival (PFS) for all randomly assigned patients. VCAP, vincristine, cyclophosphamide, doxorubicin, and prednisone; AMP, doxorubicin, ranimustine, and prednisone; VECP, vindesine, etoposide, carboplatin, and prednisone; Bi-CHOP, biweekly cyclophosphamide, doxorubicin, vincristine, and prednisone.

VECP arm (40%) than the biweekly CHOP arm (25%; $P = .020$), and ORRs were similar (72% *v* 66%, respectively) between the arms.

Median follow-up time for all randomly assigned patients was 10.9 months. The MST and OS rate at 3 years without censoring the transplantation patients were 12.7 months and 24%, respectively, in the VCAP-AMP-VECP arm and 10.9 months and 13%, respectively, in the biweekly CHOP arm (Fig 2A). For OS, the preplanned, one-sided, log-rank $P = .085$ (two-sided $P = .169$), and the hazard ratio was 0.75 (95% CI, 0.50 to 1.13). A Cox regression analysis with PS (0 *v* 1 *v* 2 to 4) as stratum for baseline hazard functions was performed to evaluate the effect on OS of the factors of age, B symptoms, subtypes of ATLL, lactate dehydrogenase, blood urea nitrogen, bulky mass, and treatment arms. According to this analysis, the hazard ratio and *P* value for the treatment arms were 0.62 (95% CI, 0.38 to 1.01) and $P = .028$ (two-sided $P = .056$), respectively. The difference between the crude analysis and this result was because of unbalanced prognostic factors, such as PS 0 versus 1, and the presence or absence of bulky lesions between the treatment arms. The median PFS time and PFS rate at 1 year were 7.0 months and 28% in the VCAP-AMP-VECP arm and 5.4 months and 16% in the biweekly-CHOP arm, respectively ($P = .100$, two-sided $P = .200$; hazard ratio = 0.77; 95% CI, 0.52 to 1.14; Fig 2B).

The rate of completion of the planned treatment was 32% in the VCAP-AMP-VECP arm and 49% in the biweekly CHOP arm. Progressive disease or relapse, as a reason for discontinuation of treatment, was observed in 40% of patients in the VCAP-AMP-VECP arm and 31% in the CHOP arm. These results seem to be associated with the periods of treatment (ie, 24 weeks in the VCAP-AMP-VECP arm and 16 weeks in the biweekly CHOP arm) because OS and PFS were better in the VCAP-AMP-VECP arm. Reasons for going off treatment, such as toxicity, were relatively numerous in the VCAP-AMP-VECP arm.

The period needed to complete each course of chemotherapy and proceed to the next course was stable in the biweekly CHOP arm (median, 15 days in courses 1 to 2 and 14 days in courses 7 to 8). In contrast, the more advanced the therapy, the more time that was required, especially after course 3 as a result of bone marrow suppression, mainly neutropenia, despite using G-CSF, in the VCAP-AMP-VECP arm (median, 30 days in courses 1 to 2 and 42 days in courses 5 to 6). The average duration of G-CSF use was 12.9 days and 5.4 days per course in the VCAP-AMP-VECP and biweekly CHOP arms, respectively.

Toxicities

Excluding one patient in the biweekly CHOP arm who refused protocol chemotherapy, 117 patients were assessable for toxicity (Table 3). The major toxicities in both arms were cytopenia and infection. In general, toxicity was more severe in the VCAP-AMP-VECP arm. In the VCAP-AMP-VECP arm versus biweekly CHOP arm, rates of grade 4 neutropenia, grade 4 thrombocytopenia, and grade 3 or 4 infection were 98% *v* 83%, 74% *v* 17%, and 32% *v* 15%, respectively. Three treatment-related deaths, two from sepsis and one from interstitial pneumonitis related to neutropenia, were reported in the VCAP-AMP-VECP arm. Two cases of myelodysplastic syndrome were reported, one each in both arms.

Subgroup Analysis

As shown in Figure 3, there was interaction between the treatment arms and PS, which suggests that the intensive

Information downloaded from jco.ascopubs.org and provided by TOKYO DAIGAKU IKAGAKU KENKYUJO on March 22, 2009 from 157.82.101.48.
Copyright © 2007 by the American Society of Clinical Oncology. All rights reserved.

Table 3. Hematologic and Nonhematologic Toxicities in 117 Treated Patients

Toxicity	Grade	% of Patients	
		VCAP-AMP-VECP (n = 57)	Biweekly CHOP (n = 60)
Neutropenia	4	98	83
Thrombocytopenia	4	74	17
T-bilirubin	3 + 4	5	2
ALT	3 + 4	11	5
Hyperglycemia	3 + 4	13	4
Hyponatremia	3 + 4	5	5
Hypokalemia	3 + 4	12	2
Stomatitis	3 + 4	7	2
Dyspnea	3 + 4	7	5
Infection	3 + 4	32	15
Neuropathy	3 + 4	2	7
Treatment-related deaths, No.	—	3*	0

Abbreviations: VCAP, vincristine, cyclophosphamide, doxorubicin, and prednisone; AMP, doxorubicin, ranimustine, and prednisone; VECP, vindesine, etoposide, carboplatin, and prednisone; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone.
*Two patients died of sepsis, and one died of interstitial pneumonitis.

VCAP-AMP-VECP regimen is more likely to benefit patients with a poor PS, possibly reflecting a more advanced stage of ATLL. Similarly, the younger population was more likely to benefit from VCAP-AMP-VECP (Fig 4). In contrast, no such interaction was observed in the analysis concerning disease type and bulky tumor (data not shown). Fourteen patients, seven in each group, received allogeneic hematopoietic stem-cell transplantation (alloHSCT). Four and three patients received the transplantation before progressive disease in the VCAP-AMP-VECP arm and biweekly CHOP arm, respectively. The estimated OS rates at 2 years for patients receiving alloHSCT were similar (43% for both arms).

DISCUSSION

To our knowledge, this trial, JCOG9801, comparing the efficacy and safety of VCAP-AMP-VECP and biweekly CHOP, is the first phase III trial exclusively conducted for ATLL in the world. We found a better OS in patients with aggressive ATLL treated with VCAP-AMP-VECP compared with biweekly CHOP, as well as a higher CR rate and longer PFS. Although the primary analysis of OS failed to show statistical significance (hazard ratio = 0.75, $P = .085$), a sensitivity analysis demonstrated the consistent result even after an adjustment of imbalance in baseline prognostic factors (hazard ratio = 0.62, $P = .028$). We consider the longer OS at 3 years and higher CR rate of VCAP-AMP-VECP than biweekly CHOP in this trial to be clinically meaningful and the former to be recommended as the first choice for patients with this disease despite higher toxicities.

Hematologic toxicity and infections were more frequent in the VCAP-AMP-VECP arm than the biweekly CHOP arm, which are similar findings to the previous JCOG9303 study with the original VCAP-AMP-VECP regimen.[15] Although both regimens were supported with G-CSF, four more drugs were incorporated in VCAP-AMP-VECP compared with biweekly CHOP, with a dose-dense and long period of chemotherapy. Three treatment-related deaths, which were related to severe neutropenia despite using G-CSF, were reported in the VCAP-AMP-VECP arm. Although VCAP-AMP-VECP caused remarkable thrombocytopenia, no serious hemorrhagic events were documented, possibly because platelet transfusion was encouraged without modifying the schedule of chemotherapy based on the decrease in the platelet count. ATLL patients treated with VCAP-AMP-VECP should be carefully monitored for complications, especially cytopenia and infections,

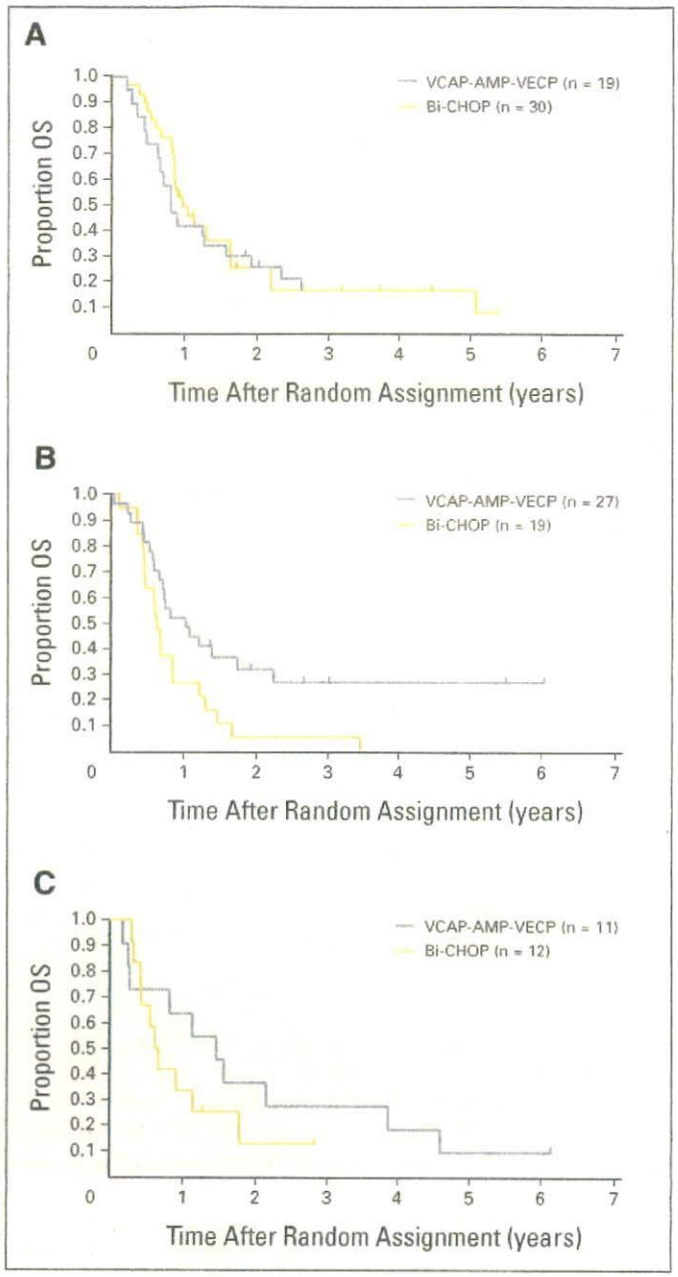

Fig 3. Kaplan-Meier estimate of overall survival (OS) for all randomly assigned patients according to performance status (PS) at diagnosis: (A) PS of 0; (B) PS of 1; and (C) PS of 2, 3, or 4. VCAP, vincristine, cyclophosphamide, doxorubicin, and prednisone; AMP, doxorubicin, ranimustine, and prednisone; VECP, vindesine, etoposide, carboplatin, and prednisone; Bi-CHOP, biweekly cyclophosphamide, doxorubicin, vincristine, and prednisone.

Information downloaded from jco.ascopubs.org and provided by TOKYO DAIGAKU IKAGAKU KENKYUJO on March 22, 2009 from 157.82.101.48.
Copyright © 2007 by the American Society of Clinical Oncology. All rights reserved.

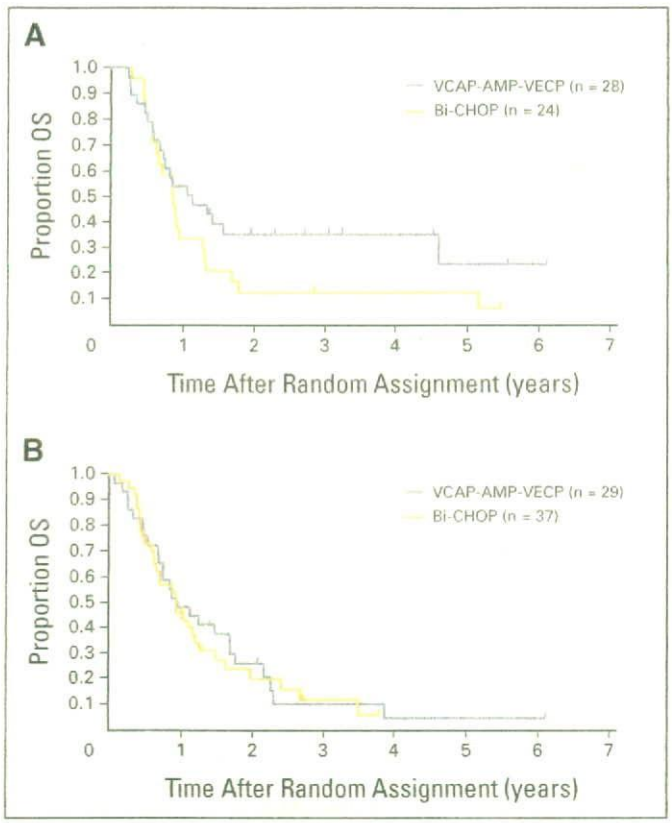

Fig 4. Kaplan-Meier estimate of overall survival (OS) for all randomly assigned patients according to age at diagnosis: (A) age < 56 years and (B) age ≥ 56 years. VCAP, vincristine, cyclophosphamide, doxorubicin, and prednisone; AMP, doxorubicin, ranimustine, and prednisone; VECP, vindesine, etoposide, carboplatin, and prednisone; Bi-CHOP, biweekly cyclophosphamide, doxorubicin, vincristine, and prednisone.

with supportive care such as platelet transfusion, use of G-CSF, and the prophylaxis of opportunistic infection.[4-7]

Because of the geographically limited distribution and rarity of this disease, large-scale trials for the treatment of ATLL have rarely been performed. There have been trials of intensive chemotherapy and a unique combination chemotherapy consisting of interferon alfa and zidovudine.[21,22] However, the results were inferior to those of VCAP-AMP-VECP.[23] The following factors might explain the reasons for the superiority of VCAP-AMP-VECP compared with biweekly CHOP against aggressive ATLL. Four more drugs were incorporated in the VCAP-AMP-VECP arm, compared with the biweekly CHOP arm, with a dose-dense and long period of chemotherapy. Furthermore, carboplatin and ranimustine were incorporated, and these agents are not affected by P-glycoprotein, which is frequently expressed in ATLL cells at onset.[17]

According to the results of subgroup analyses, VCAP-AMP-VECP may be more beneficial in patients with more advanced ATLL or in younger patients. However, because this study excluded patients with a PS of 4 not caused by hypercalcemia, the results would not be applicable to patients with a PS of 4 caused by an opportunistic infection or organ involvement by ATLL.

Our previous studies in advanced NHL revealed that CR rate and OS were poorer in ATLL than in other aggressive NHL.[10-12] In the recent two studies, biweekly CHOP was better than or not inferior to standard CHOP for aggressive NHL excluding ATLL.[24,25] VCAP-AMP-VECP was superior to biweekly CHOP for ATLL in this study. A G-CSF–supported dose-intensified regimen with carboplatin and ranimustine might have overcome the characteristics of ATLL cells, such as rapid proliferation and P-glycoprotein expression.

AlloHSCT is now considered as promising for the treatment of young patients with ATLL.[26] Thus, despite higher toxicities and poor completion rate of the planned course of therapy, VCAP-AMP-VECP, which provided a higher CR rate and a probable survival advantage, is promising as induction chemotherapy preceding upfront alloHSCT for aggressive ATLL. However, to prove the effectiveness of this strategy, further studies are needed.

In conclusion, the results of the present phase III study suggest that VCAP-AMP-VECP is a more effective chemotherapy regimen for patients with newly diagnosed aggressive ATLL even with higher toxicity profiles. However, the MST of 13 months is still not satisfactory. We are now planning a phase II study of myeloablative alloHSCT after induction therapy with VCAP-AMP-VECP for young patients with this disease.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Kunihiro Tsukasaki, Atae Utsunomiya, Haruhiko Fukuda, Takuya Fukushima, Masayuki Sano, Kensei Tobinai, Masanori Shimoyama, Masao Tomonaga
Administrative support: Tomomitsu Hotta
Provision of study materials or patients: Atae Utsunomiya, Takuya Fukushima, Yoshifusa Takatsuka, Shuichi Ikeda, Masato Masuda, Haruhisa Nagoshi, Ryuzo Ueda, Kazuo Tamura, Masayuki Sano, Saburo Momita, Kazunari Yamaguchi, Fumio Kawano, Shuichi Hanada, Kensei Tobinai, Tomomitsu Hotta
Collection and assembly of data: Haruhiko Fukuda, Takuya Fukushima, Yoshifusa Takatsuka, Shuichi Ikeda, Masato Masuda, Haruhisa Nagoshi, Ryuzo Ueda, Kazuo Tamura, Masayuki Sano, Saburo Momita, Kazunari Yamaguchi, Fumio Kawano, Shuichi Hanada, Kensei Tobinai, Tomomitsu Hotta
Data analysis and interpretation: Haruhiko Fukuda, Taro Shibata, Takuya Fukushima, Masao Tomonaga
Manuscript writing: Kunihiro Tsukasaki, Haruhiko Fukuda, Taro Shibata, Kensei Tobinai, Masanori Shimoyama, Tomomitsu Hotta, Masao Tomonaga
Final approval of manuscript: Kunihiro Tsukasaki, Haruhiko Fukuda, Takuya Fukushima, Tomomitsu Hotta, Masao Tomonaga

Information downloaded from jco.ascopubs.org and provided by TOKYO DAIGAKU IKAGAKU KENKYUJO on March 22, 2009 from 157.82.101.48.
Copyright © 2007 by the American Society of Clinical Oncology. All rights reserved.

REFERENCES

1. Uchiyama T, Yodoi J, Sagawa K, et al: Adult T-cell leukemia: Clinical and hematologic features of 16 cases. Blood 50:481-492, 1977
2. Poiesz BJ, Ruscetti FW, Gazdar AF, et al: Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. Proc Natl Acad Sci U S A 77:7415-7419, 1980
3. Yoshida M, Miyoshi I, Hinuma Y: Isolation and characterization of retrovirus from cell lines of human adult T cell leukemia and its implication in the disease. Proc Natl Acad Sci U S A 79:2031-2035, 1982
4. Shimoyama M: Diagnostic criteria and classification of clinical subtypes of adult T- cell leukaemia-lymphoma: A report from the Lymphoma Study Group (1984-87). Br J Haematol 79:428-437, 1991
5. Pawson R, Mufti GJ, Pagliuca A: Management of adult T-cell leukaemia/lymphoma. Br J Haematol 100:453-458, 1998
6. Tobinai K, Watanabe T: Adult T-cell leukemia-lymphoma, in Abeloff MD, Armitage JO, Niederhuber JE, et al (eds): Clinical Oncology (ed 3). Philadelphia, PA, Elsevier Churchill Livingstone, 2004, pp 3109-3130
7. Shimoyama M: Chemotherapy of ATL, in Takatsuki K (ed): Adult T-Cell Leukemia. Oxford, United Kingdom, Oxford University Press, 1994, pp 221-237
8. A predictive model for aggressive non-Hodgkin's lymphoma: The International Non-Hodgkin's Lymphoma Prognostic Factors Project. N Engl J Med 329:987-994, 1993
9. Gallamini A, Stelitano C, Calvi R, et al: Peripheral T-cell lymphoma unspecified (PTCL-U): A new prognostic model from a retrospective multicentric clinical study. Blood 103:2474-2479, 2004
10. Lymphoma Study Group: Major prognostic factors of patients with adult T-cell leukemia-lymphoma: A cooperative study—Lymphoma Study Group (1984-1987). Leuk Res 15:81-90, 1991
11. Shimoyama M, Ota K, Kikuchi M, et al: Chemotherapeutic results and prognostic factors of patients with advanced non-Hodgkin's lymphoma treated with VEPA or VEPA-M. J Clin Oncol 6:128-141, 1988
12. Tobinai K, Shimoyama M, Minato K, et al: Japan Clinical Oncology Group phase II trial of second-generation "LSG4 protocol" in aggressive T- and B-lymphoma: A new predictive model for T-and B-lymphoma. Proc Am Soc Clin Oncol 13:378, 1994 (abstr)
13. Tobinai K, Shimoyama M, Inoue S, et al: Phase I study of YK-176 (2'-deoxycoformycin) in patients with adult T-cell leukemia-lymphoma. Jpn J Clin Oncol 22:164-171, 1992
14. Tsukasaki K, Tobinai K, Shimoyama M, et al: Deoxycoformycin-containing combination chemotherapy for adult T-cell leukemia-lymphoma: Japan Clinical Oncology Group Study (JCOG9109). Int J Hematol 77:164-170, 2003
15. Yamada Y, Tomonaga M, Fukuda H, et al: A new G-CSF-supported combination chemotherapy, LSG15, for adult T-cell leukemia-lymphoma: Japan Clinical Oncology Group Study 9303. Br J Haematol 113:375-382, 2001
16. Tsukasaki K, Ikeda S, Murata K, et al: Characteristics of chemotherapy-induced clinical remission in long survivors with aggressive adult T-cell leukemia/lymphoma. Leuk Res 17:157-166, 1993
17. Kuwazuru Y, Hanada S, Furukawa T, et al: Expression of p-glycoprotein in adult T-cell leukemia cells. Blood 76:2065-2071, 1990
18. Itoh K, Ohtsu T, Fukuda H, et al: Randomized phase II study of biweekly CHOP and dose-escalated CHOP with prophylactic use of lenograstim (glycosylated G-CSF) in aggressive non-Hodgkin's lymphoma: Japan Clinical Oncology Group Study 9505. Ann Oncol 13:1329-1330, 2002
19. Tobinai K, Kohno A, Shimada Y, et al: Toxicity grading criteria of the Japan Clinical Oncology Group: The Clinical Trial Review Committee of the Japan Clinical Oncology Group. Jpn J Clin Oncol 23:250-257, 1993
20. Lan KKG, DeMets DL: Discrete sequential boundaries for clinical trials. Biometrika 70:659-663, 1983
21. Taguchi H, Kinoshita KI, Takatsuki K, et al: An intensive chemotherapy of adult T-cell leukemia/lymphoma: CHOP followed by etoposide, vindesine, ranimustine, and mitoxantrone with granulocyte colony-stimulating factor support. J Acquir Immune Defic Syndr Hum Retrovirol 12:182-186, 1996
22. Gill PS, Harrington W Jr, Kaplan MH, et al: Treatment of adult T-cell leukemia-lymphoma with a combination of interferon alpha and zidovudine. N Engl J Med 332:1744-1748, 1995
23. Tobinai K, Kobayashi Y, Shimoyama M: Interferon alpha and zidovudine in adult T-cell leukemia-lymphoma: Lymphoma Study Group of the Japan Clinical Oncology Group. N Engl J Med 333:1285, 1995
24. Pfreundschuh M, Trumper L, Kloess M, et al: Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of elderly patients with aggressive lymphomas: Results of the NHL-B2 trial of the DSHNHL. Blood 104:634-641, 2004
25. Hotta T, Shimakura Y, Ishizuka N, et al: Randomized phase III study of standard CHOP (S-CHOP) versus biweekly CHOP (Bi-CHOP) in aggressive non-Hodgkin's lymphoma (NHL): Japan Clinical Oncology Group study, JCOG9809. Proc Am Soc Clin Oncol 22:565, 2003 (abstr 2271)
26. Fukushima T, Miyazaki Y, Honda S, et al: Allogeneic hematopoietic stem cell transplantation provides sustained long-term survival for patients with adult T-cell leukemia/lymphoma. Leukemia 19: 829-834, 2005

Acknowledgment

The Acknowledgment is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Information downloaded from jco.ascopubs.org and provided by TOKYO DAIGAKU IKAGAKU KENKYUJO on March 22, 2009 from 157.82.101.48.
Copyright © 2007 by the American Society of Clinical Oncology. All rights reserved.

Expression Profiling of Immature Thymocytes Revealed a Novel Homeobox Gene That Regulates Double-Negative Thymocyte Development[1]

Masahito Kawazu,*† Go Yamamoto,* Mayumi Yoshimi,* Kazuki Yamamoto,* Takashi Asai,* Motoshi Ichikawa,* Sachiko Seo,* Masahiro Nakagawa,* Shigeru Chiba,† Mineo Kurokawa,* and Seishi Ogawa[2]*‡§¶

Intrathymic development of CD4/CD8 double-negative (DN) thymocytes can be tracked by well-defined chronological subsets of thymocytes, and is an ideal target of gene expression profiling analysis to clarify the genetic basis of mature T cell production, by which differentiation of immature thymocytes is investigated in terms of gene expression profiles. In this study, we show that development of murine DN thymocytes is predominantly regulated by largely repressive rather than inductive activities of transcriptions, where lineage-promiscuous gene expression in immature thymocytes is down-regulated during their differentiation. Functional mapping of genes showing common temporal expression profiles implicates previously uncharacterized gene regulations that may be relevant to early thymocytes development. A small minority of genes is transiently expressed in the CD44^{low}CD25^{+} subset of DN thymocytes, from which we identified a novel homeobox gene, *Duxl*, whose expression is up-regulated by Runx1. *Duxl* promotes the transition from CD44^{high}CD25^{+} to CD44^{low}CD25^{+} in DN thymocytes, while constitutive expression of *Duxl* inhibits expression of TCR β-chains and leads to impaired β selection and greatly reduced production of CD4/CD8 double-positive thymocytes, indicating its critical roles in DN thymocyte development. *The Journal of Immunology*, **2007, 179: 5335–5345.**

Intrathymic development of thymocytes from their bone marrow progenitors is a critical process for the generation of mature T cells, through which the immature thymocyte progenitors, as identified by the absence of mature T cell markers (CD4/CD8 double-negative (DN)[3] T cells), differentiate into the CD4/CD8 double-positive (DP) cells, and finally produce CD4 or CD8 single-positive T cells. Although accounting for <5% of total thymocytes in mice, the DN thymocytes undergo a dynamic developmental process that is essential for the subsequent expansion into the DP population (1). Among these DN thymocytes, the earliest chronological subset (DN1) is recognized as a CD44^{high}C-Kit^{+}CD25^{-} population (2, 3). In the first wave of cytokine-dependent pre-T cell expansion, the DN1 cells begin to proliferate with concomitant up-regulation of CD25, giving rise to the DN2 population showing the CD44^{high}C-Kit^{+}CD25^{+} phenotype (4). The DN2 thymocytes then start to rearrange their *TCR* genes and down-regulate the CD44 expression to generate the CD44^{low}CD25^{+} DN3 subset (5). In the DN3 stage, thymocytes are subjected to a process called β selection and only those DN3 cells that have productively rearranged their *TCRβ* gene can survive and transit into CD44^{low}CD25^{-} DN4 thymocytes, followed by rapid expansion into the DP population (6–8).

During these early developmental processes, immature thymocyte progenitors lose their multilineage plasticity and exclusively commit to T cell differentiation. Under appropriate conditions, the DN1 population can generate multilineage hemopoietic components in vitro (9, 10), although there still remains some controversy with regard to their potential to B lineage differentiation (11). The potential of the multilineage commitment is more restricted in the DN2 stage, where cells can still give rise to NK cells and thymic dendritic cells, but no more B cells (10, 12–14), and after the DN2/DN3 transition and the succeeding β selection, thymocytes mostly lose their potential to multilineage plasticity and totally commit to T cell lineage (10).

Because these processes are thought to take place under tightly controlled gene expression, it is of particular importance to clarify the nature of this gene regulation and the key regulators involved in that regulation. To date, a number of molecules have been identified that regulate these developmental processes (15). *Notch1*-deficient thymocytes, for example, are not able to produce T cells but differentiate into B cells (16), while *pTα*, *TCRβ*, *Lck*, *SLP76*, and *Lat* are shown to be indispensable for β selection and their knockout mice show a severe maturational block at the DN3/DN4 transition (17). Similarly, the DN2/DN3 transition is completely blocked in *Runx1*-deficinet mice (18, 19) and also in double knockout mice of *pTα* and *common cytokine receptor γ-chain* genes (20). In contrast, it is well anticipated that these developmental processes in DN thymocytes should involve regulation of much

*Department of Hematology and Oncology, †Cell Therapy and Transplantation Medicine, and ‡Regeneration Medicine for Hematopoiesis; §21st Century Core of Excellence (COE) program, Graduate School of Medicine, University of Tokyo, University of Tokyo Hospital, Tokyo, Japan; and ¶Core Research for Evolutional Science and Technology, Japan Science and Technology Agency

Received for publication June 5, 2007. Accepted for publication July 23, 2007.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by Research on Measures for Intractable Diseases, Health and Labor, Sciences Research Grants, Ministry of Health, Labor and Welfare, Research on Health Sciences focusing on Drug Innovation, the Japan Health Sciences Foundation, and Core Research for Evolutional Science and Technology, Japan Science and Technology Agency.

[2] Address correspondence and reprint requests to Dr. Seishi Ogawa, Department of Regeneration Medicine for Hematopoiesis, Graduate School of Medicine, University of Tokyo, Bunkyo-ku, Tokyo, Japan. E-mail address: sogawa-tky@umin.ac.jp

[3] Abbreviations used in this paper: DN, double negative; DP, double positive; FL, fetal liver; rh, recombinant human; shRNA, small hairpin RNA; iTCRβ, intracellular TCRβ; qPCR, quantitative PCR; NGFRt, nerve growth factor receptor; OP9-DL1, OP9-delta-like-1.

Copyright © 2007 by The American Association of Immunologists, Inc. 0022-1767/07/$2.00

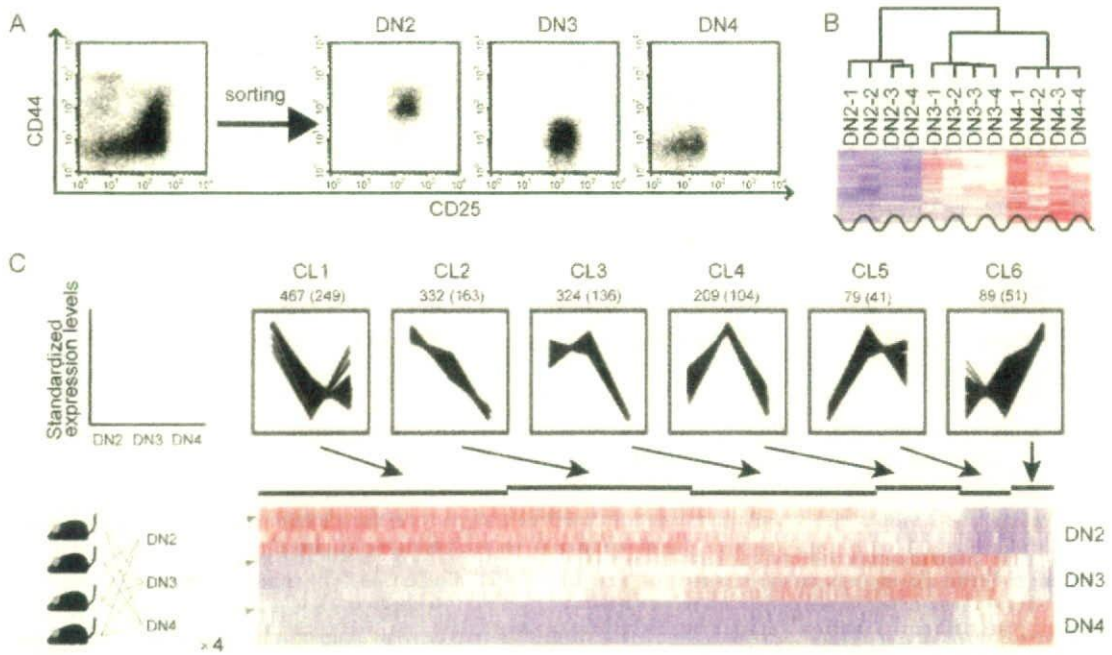

FIGURE 1. Microarray analysis and clustering of genes differentially expressed during early thymocyte development. *A*, DN2, DN3, and DN4 thymocytes from four C57BL/6 mice were FACS sorted and subjected to microarray analysis of gene expression profiles. *B*, Hierarchical clustering of 12 microarray data obtained from four independent experiments. Only branch portion is presented. *C*, One thousand five hundred differentially expressed genes were grouped into six clusters showing discrete temporal expression profiles by K-means clustering methods (*upper panels*), which are also presented in heat map (*lower panel*). The expression values for a gene across all samples were standardized to have a mean 0 and SD 1. Standardized expression levels of genes are indicated in graphs (*upper panels*) and in heat map (*lower panel*). Vertical scale of the graphs is from −1.5 to +1.5. Each column of heat map represents a gene and each row represents a sample. Red indicates high expression and blue low expression. The numbers of genes within clusters are indicated along with those for which human counterparts were identified (in the parentheses).

larger numbers of genes than those related to these known molecules. To understand the molecular mechanisms of DN thymocyte development, it may be also of use to clarify how these developmental processes are regulated in terms of their entire gene expression, to which cell differentiation is ultimately ascribed.

In the current study, we approached this issue by investigating gene expression profiles in discrete subsets of DN thymocytes under development in which DN2, DN3, and DN4 thymocytes were sorted and subjected to expression profiling analysis with high-density oligonucleotide microarrays. Clustering of differentially expressed genes among these DN thymocyte subsets demonstrated that during DN development, regulation of gene expression is predominantly repressive rather than inductive, in which multiple lineage-affiliated genes expressed in immature thymocytes are down-regulated during the course of DN thymocyte development. Functional mapping of clustered genes also revealed a possible involvement of previously uncharacterized functional gene regulations in thymocyte development. Finally, we identified a novel homeobox gene (*Duxl*), transiently expressed in DN3 thymocytes. We showed that *Duxl* is induced by Runx1 and regulates DN thymocyte development by promoting the DN2/DN3 transition, while deregulated expression of *Duxl* resulted in impaired β selection and severely compromised production of DP thymocytes, indicating its critical roles in DN thymocyte production.

Materials and Methods

Cell sorting and RNA extraction

All Abs used for cell sorting were purchased from BD Pharmingen. Thymocytes were harvested from 5- to 6-wk-old female C57BL/6 mice. Four independent cell sortings were performed and four mice were sacrificed for each experiment. Before cell sorting, $CD4^+$ cells and $CD8^+$ cells were depleted using the MACS LD system (Miltenyi Biotec). The remaining fraction was stained with anti-CD44 and anti-CD25 Abs conjugated to FITC or PerCP-Cy5.5, respectively, and also with PE-conjugated Abs to CD4, CD8, CD3, NK1.1, and TCR$\gamma\delta$ and sorted using a FACSAria cell sorter (BD Biosciences). DN2, DN3, and DN4 subsets were identified as $FITC^+PE^-PerCP^-Cy5.5^+$, $FITC^-PE^-PerCP^-Cy5.5^+$, and $FITC^-PE^-PerCP^-Cy5.5^-$ populations, respectively. For expression analysis of various hemopoietic lineages, mononuclear cells were separated from a single-cell suspension of bone marrow of 5- to 6-wk-old female C57BL/6 mice by centrifugation on a Histopaque-1083 (Sigma-Aldrich). $c\text{-}kit^+$ cells were obtained by positive selection for the c-kit Ag with MACS magnetic beads. The remaining fraction was stained with FITC-conjugated Ab to Ter119, PerCP-conjugated Ab to B220, and PE-conjugated Abs to Mac1, and $Ter119^+$ fraction, $B220^+$ fraction, and $Mac1^+$ fraction were sorted. B220+ splenocytes and CD3+ splenocytes were collected by positive selection of splenocytes for the B220 Ag or CD3 Ag with MACS magnetic beads. RNA was extracted from sorted cells using an RNeasy Mini kit (Qiagen) according to the manufacturer's instruction.

Microarray experiments

Biotin-labeled cRNA probes were prepared using a Two-Cycle cDNA Synthesis Kit (Affymetrix). Following fragmentation, biotin-labeled cRNA was hybridized to the Mouse Genome 430 2.0 Array (Affymetrix) for 16 h at 45°C as recommended by the manufacturer. Washing was performed using an automated fluidics workstation, and the array was immediately scanned with GeneChip Scanner 3000 7G. Expression data were extracted from image files produced on Affymetrix GeneChip Operating software 1.0 (GCOS). The absolute detection call (present, absent, or marginal) for each probe set was determined on GCOS. Normalization and expression value calculation were performed using a DNA-Chip Analyzer (www.dchip.org) (21). The invariant set normalization method (22) was used to normalize arrays at probe cell level to make them comparable and the model-based method (22) was used for computing expression values. These expression levels were attached with SEs as measurement accuracy, which were subsequently used to compute 90% confidence intervals of fold changes in two group comparisons (22). The lower confidence bounds of fold changes

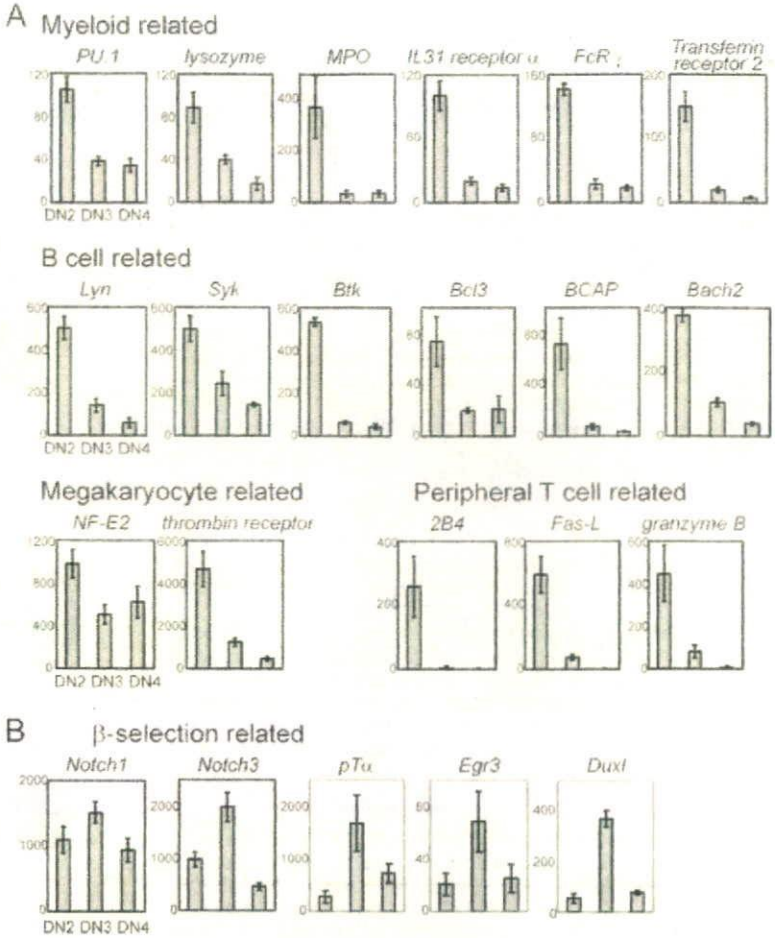

FIGURE 2. Expression patterns of lineage-affiliated genes and genes involved in β selection. Signal intensities on microarray for representative genes affiliated with various hemopoietic lineages are shown in *A* and for those relevant to β selection in *B* by their mean values ($\pm$SD), where the outlier data were excluded from the calculation. Expression data for *Duxl* are also presented in *B*.

were conservative estimate of the real fold changes. Differentially expressed probe sets were identified as those sets whose mean signals showed >1.5-fold difference between DN2 and DN3 and DN3 and DN4. The probability of false discoveries with this threshold was calculated by random permutations as described previously (23). Raw microarray data can be found at http://www.ncbi.nlm.nih.gov/geo/, GEO accession no. GSE7784.

Clustering and pathway analysis

One thousand five hundred differentially expressed genes that were identified as described above were clustered using a K-means (24) server on Gene Expression Pattern Analysis Suite (http://gepas.bioinfo.cnio.es/) (25, 26). Molecular network of each cluster was analyzed by KeyMolnet software, which was developed by the Institute of Medicinal Molecular Design, Inc. (IMMD) (27). Known molecular data were curated by IMMD and the obtained gene list of each cluster was combined with this software and shown as molecular networks. Significance of a determined pathway to obtained networks was determined. To ascertain whether any molecular pathways determined by IMMD annotate a relation between molecules in the networks at a frequency greater than that would be expected by chance, this software calculates a p value using the hypergeometric distribution as described previously (28).

Cloning of Duxl cDNA and construction of retrovirus plasmid

cDNAs of Duxl and FLAG-tagged Duxl having *Not*I and *Xho*I sites on their 5′ and 3′ terminus, respectively, were PCR amplified from template cDNA prepared from total thymic RNA using TaKaRa LA *Taq* (Takara Bio) with the following sets of primers: 5′-AAAAGCGGCCGCATCGATACCATGGAGCTGAGCTGCAGTACT-3′ (for Duxl sense), 5′-AAAAGCGGCCGCATCGATACCATGGACTACAAGGACGACGATGACAAGATGGAGCTGAGCTGCAGTACT-3′ (for FLAG-tagged Duxl sense), and 5′-AAAACTCGAGCTACGGAGTTTGGTGTGCTT-3′ (for the common antisense for both). Each PCR product was digested with *Not*I and *Xho*I and cloned into the *Not*I-*Xho*I site located at the 5′ upstream of IRES-GFP or IRES-NGFRt of the pGCDNsam (a gift from Dr. H. Nakauchi, University of Tokyo, Tokyo, Japan) retrovirus vector. Nucleotide sequences of these plasmids were confirmed by resequencing. Retrovirus was prepared by transfecting the PlatE (a gift from Dr. T. Kitamura, University of Tokyo) cell line with each construct.

Quantitative PCR analysis

Total cellular RNA was converted into cDNAs by reverse transcriptase (Superscript III; Invitrogen Life Technologies) with random primers. cDNAs were amplified in triplicate for 40 cycles at 95°C for 15 s and 60°C for 60 s using an Applied Biosystems PRISM 7000 Sequence Detection System according to the manufacturer's instructions. Predesigned TaqMan primer and probe sets for 1110051B16Rik (Mm00841823_m1; Applied Biosystems) (see Fig. 4), for PU.1 (Mm00488140_m1), and for 18S rRNA (no. 4308329; Applied Biosystems) were used for the assay. PCR amplification of GAPDH, 1110051B16Rik (see Fig. 5*B*) and Rag1 was performed using Platinum SYBR Green qPCR SuperMix-UDG with ROX (Invitrogen Life Technologies) and the following primer sequences: GAPDH (forward, 5′-GAATCTACTGGAGTCTTCACC-3′; reverse, 5′-GTCATGAGCCCTTCCACGATGC-3′), 1110051B16Rik (forward, 5′-GGGAAAACTGGCTCAACAA-3′; reverse, 5′-GTGTTCTGTCCTGGGTCTGG-3′), Rag1 (forward, 5′-CTGAAGCTCAGGGTAGACGG-3′; reverse, 5′-CAACCAAGCTGCAGACATTC-3′). Significant PCR fluorescent signals were normalized for each sample to a PCR fluorescent signal obtained using GAPDH or 18S rRNA as control.

Coculture of fetal liver (FL) cells with OP9-delta-like-1 (OP9-DL1) stromal cells

OP9-DL1 is a bone marrow stromal cell line that expresses a Notch ligand, Delta-like 1, and supports development of DP thymocytes from FL-derived hemopoietic progenitors (29) and was provided by Dr. J. C. Zúñiga-Pflücker (University of Toronto, Toronto, Canada). In our OP9-DL1 assay, FL cells were harvested from E14.5 embryos and cultured on OP9-DL1 cells

in combination with retroviral gene transfer as previously described (19). Briefly, mononuclear cells were separated from FL cells of C57BL/6 mice. In brief, 5×10^4 mononuclear cells were cultured on confluent OP9-DL1 cells in flat-bottom 24-well culture plates with 500 ml of MEM (Invitrogen Life Technologies) supplemented with 20% FCS, penicillin/streptomycin, and 5 ng/ml recombinant human (rh) IL-7 (Techne Laboratories). After 5 or 6 days of culture, 5×10^4 cells were passed onto newly prepared OP9-DL1 cells in the presence of 5 ng/ml rhIL-7, and retrovirus infection was performed using polybrene (final concentration, 8 mg/ml), followed by another 5 or 6 days of culture. In brief, 5×10^4 cells were again passed onto newly prepared OP9-DL1 cells and cultured for another 5 or 6 days, but in rhIL-7-free culture medium.

PCR for TCRβ rearrangement

PCR for TCRβ rearrangements was performed as described elsewhere (30) on DNA isolated from FL cells cultured on OP9-DL1 using the following primers: Dβ2, GTAGGCACCTGTGGGGAAGAAACT; Jβ2, TGAGAG CTGTCTCCTACTATCGATT; and Vβ5.1, GTCCAACAGTTTGATGAC TATCAC. After 40 cycles of amplification (10 s at 98°C, 2 min at 68°C), PCR products were separated on a 4% agarose gel.

RNA interference

The vector backbone was RNAi-Ready pSIREN-RetroQ-ZsGreen (BD Clontech). The RNA interference target sequences were GGAGCAG GATAAACCTAGA (sequence 1), GACTGATATTCTAATTGAA (sequence 2), and GTTCCAGACTGATATTCTA (sequence 3). The small hairpin RNA (shRNA) were designed by a shRNA design algorithm, which was developed by Dr. M. Miyagishi (31). Oligonucleotides used for construction were GATCCGGGGTAGGATAAACTTAGAACGTGTGCTGT CCGTTCTAGGTTTATCCTGCTCCTTTTTACGCGTG (oligonucleotide 1, sense), AATTCACGCGTAAAAAGGAGCAGGATAAACCTAGAAC GGACAGCACACGTTCTAAGTTTATCCTACCCCG (oligonucleotide 1, antisense), GATCCGATTGATGTTCTAGTTGAAACGTGTGCTGTC CGTTTCAATTAGAATATCAGTCTTTTTACGCGTG (oligonucleotide 2, sense), AATTCACGCGTAAAAAGACTGATATTCTAATTGAAACG GACAGCACACGTTTCAACTAGAACATCAATCG (oligonucleotide 2, antisense), GATCCGTTTCAGATTGATGTTCTAACGTGTGCTGTCCG TTAGAATATCAGTCTGGAACTTTTTACGCGTG (oligonucleotide 3, sense), and AATTCACGCGTAAAAAGTTCCAGACTGATATTCTAAC GGACAGCACACGTTAGAACATCAATCTGAAACG (oligonucleotide 3, antisense). Retrovirus was prepared by transfecting the PlatE cell line with the knockdown vectors.

Results

Clustering of differentially expressed genes during DN thymocyte development

The DN2, DN3, and DN4 populations were FACS sorted from DN thymocytes harvested from four C57BL/6 mice and analyzed by an Affymetrix Mouse Genome 430 2.0 Array for gene expression (Fig. 1*A*). Four independent experiments were performed using 16 mice. After normalizing the array signals using an invariant gene set (22), we extracted differentially expressed probe sets whose signals showed ≥1.5-fold difference between DN2 and DN3 or between DN3 and DN4. With this threshold, 1,901 probe sets (or 1,500 nonredundant genes) were extracted as "differentially expressed" from a total of 27,330 probes (16,131 genes) expressed in either of the three DN subsets, with a false discovery rate of 0.007 as determined by random permutation tests (23). In hierarchical clustering, the four independent array data sets for each subset were correctly clustered into the same clusters, validating the reproducibility across the experiments (Fig. 1*B*). In contrast, the K-means clustering of the 1,500 differentially expressed genes identified six clusters, CL1–CL6, showing discrete temporal expression profiles: genes expressed higher in DN2 and down-regulated in DN3 (CL1), those expressed higher in DN2 and gradually down-regulated in DN3 and DN4 (CL2), those expressed in DN2 and DN3 but down-regulated in DN4 (CL3), those only transiently expressed in DN3 (CL4), those expressed both in DN3 and DN4 (CL5), and those showing low expression in DN2 and DN3 and up-regulated in DN4 (CL6) (Fig. 1*C*). The lists of genes in these clusters were presented in supplementary Table S1, a–f.[4]

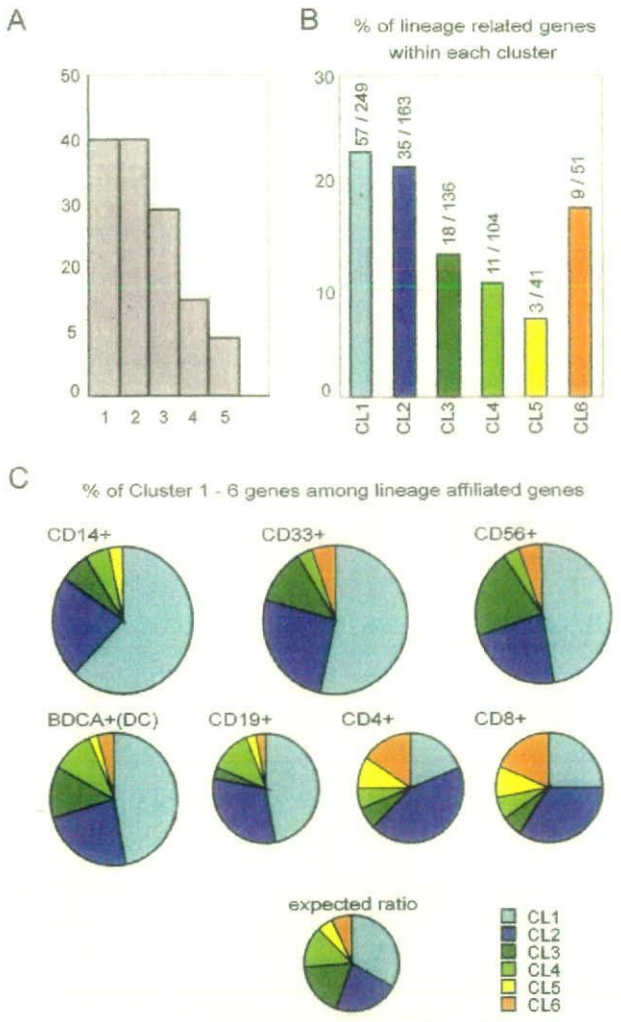

FIGURE 3. Temporal expression profiles of hemopoietic lineage-affiliated genes. *A*, Histogram of the number of lineage-affiliated genes with regard to the number of lineages to which each gene is assigned. *B*, Percentages of lineage-affiliated genes within each cluster showing discrete temporal expression profiles. Absolute numbers of lineage-affiliated genes and all genes in each cluster are shown above the bar. *C*, Number and composition of different lineage-affiliated genes with regard to clusters. Area of each circular graph represents the number of genes affiliated with each lineage.

Predominantly repressive gene regulation in DN thymocytes and their lineage-promiscuous gene expression

With regard to the temporal profiles of gene expression in DN thymocytes, our first note is that 1123 (74.9%) of the 1500 differentially expressed genes are initially expressed in DN2 but eventually down-regulated during the course of DN thymocyte development (CL1–3), while only 79 (5.27%) and 89 (5.93%) genes are up-regulated in DN3 and DN4 (CL5 and CL6), respectively (Fig. 1*C*). The other set of genes (CL4) are transiently expressed in DN3 but have low expression in DN2 and DN4. Thus, gene regulation during DN thymocyte development is largely repressive rather than inductive. Our next note was that these down-regulated genes

[4] The online version of this article contains supplemental material.

Table I. *Significantly relevant pathways in each cluster*

Rank	Name	p
CL1		
1	Ets (Spi subfamily) regulation	3.84×10^{-9}
2	IL-4 signaling	1.35×10^{-8}
3	IgG signaling	1.39×10^{-4}
4	Integrin signaling	6.94×10^{-4}
5	c-kit signaling	1.47×10^{-3}
6	G protein (G12/13) signaling	1.55×10^{-3}
7	Ets (TCF subfamily) regulation	3.11×10^{-3}
8	Rho family signaling	8.30×10^{-3}
9	CD44 signaling	8.60×10^{-3}
CL2		
1	RUNX regulation	2.19×10^{-7}
2	Endoplasmic reticulum regulation	9.72×10^{-5}
3	Platelet-derived growth factor signaling	4.94×10^{-4}
4	Prolactin signaling	7.09×10^{-4}
5	PI3K signaling	1.23×10^{-3}
6	γ-Aminobutyric acid signaling	1.23×10^{-3}
7	IL-2 signaling	1.78×10^{-3}
8	CCR10 signaling	4.66×10^{-3}
9	G protein (Gq/11) signaling	5.87×10^{-3}
10	Human growth factor signaling	9.55×10^{-3}
CL3		
1	NFAT regulation	9.67×10^{-8}
2	AP-1 regulation	3.12×10^{-5}
3	Caspase signaling	3.59×10^{-3}
4	PPARα regulation	5.24×10^{-3}
CL4		
1	TCRαβ signaling	2.35×10^{-6}
2	Ets (Ets family) regulation	1.40×10^{-5}
3	Integrin signaling	1.87×10^{-3}
4	NF-κB regulation	7.58×10^{-3}
5	Notch signaling	8.23×10^{-3}
CL5 and 6		
1	Integrin signaling	1.67×10^{-6}
2	TCRαβ signaling	7.23×10^{-6}
3	Ets (Ets family) regulation	3.17×10^{-5}
4	RB/E2F regulation	2.64×10^{-4}
5	TCF regulation	4.53×10^{-4}

contain a variety of genes whose expression is relatively characteristic to one or more hemopoietic lineages (lineage affiliated), including *NF-E2* and *thrombin receptor* (megakaryocytes), *PU.1*, *lysozyme*, and *myeloperoxidase* (myeloid lineages), *Lyn*, *Syk*, *Btk*, *Bcl3*, *BCAP*, and *Bach2* (B cells), and *2B4*, *Fas ligand*, and *granzyme B* (mature T cells) (Fig. 2*A*). This suggested that "promiscuous" expression of multiple lineage-affiliated genes in early thymocyte progenitors and their repression during the course of their commitment to mature T cells could be one of the characteristics of immature thymocytes. To confirm this in more detail, we identified those genes whose expression was thought to characterize a variety of mature hemopoietic lineages and tracked their expression levels during the course of DN thymocyte development. Because no comprehensive gene expression database was available for mice, we did this using a human tissue expression database at the Genomics Institute of the Novartis Research Foundation web site (http://symatlas.gnf.org/SymAtlas/) (32) by translating mouse genes into their human counterparts. A gene is considered to be affiliated with a lineage if its human counterpart shows >10 times higher expression in that lineage than their median expression among 79 different tissues. Among the 1500 differentially expressed genes, the human counterparts were uniquely identified for 744 (49.6%) genes (Fig. 1*C*), of which 133 (17.9%) satisfied the above criteria for being affiliated with one or more hemopoietic cell lineages (Fig. 3*A*), and lists of genes affiliated with respective lineages are presented in supplementary Table S2, a–g.[4] Among the 133 lineage-affiliated genes, 110 (83%) are expressed in DN2 and down-regulated during the course of the DN thymocyte development (CL1–CL3), accounting for 20% of 548 down-regulated genes assigned to CL1–3 (Fig. 3*B*). More genes are rapidly down-regulated during the DN2/DN3 transition (CL1), whereas as indicated from the relatively high CL3 component, down-regulation of genes affiliated with NK cells seems to occur more slowly (Fig. 3*C*). Although most of these down-regulated genes are affiliated with lineages other than T cells, many T cell-affiliated genes are also prematurely expressed in immature thymocytes and undergo down-regulation before they are definitely expressed later in thymocyte maturation (Figs. 2*A* and 3*C*). Only 12 (9%) lineage-affiliated genes were newly up-regulated until the DN4 stages and mostly related to T cells (Fig. 3, *B* and *C*). Note that our criteria for lineage relatedness may be too conservative, since many of genes presumed to be specific to mature T cells, such as *2B4*, *Fas ligand*, and *granzyme B*, were not extracted as T cell-related genes.

Pathway analysis of gene clusters during DN thymocytes development

The functional features of genes showing discrete expressional time courses are of interest to understand the regulation of DN development. To address this, we statistically explored possible functional links among genes within each cluster by evaluating a probability that a given set of genes on the known functional pathway was grouped together by chance within that cluster, showing a similar temporal expression profile using KeyMolnet software

Table II. *List of genes in CL4 with presumed DNA-binding capacity*

Entrez Gene Identification	Gene Title	GO Molecular Function Description
26972	Sporulation protein, meiosis-specific, SPO11 homolog (*Saccharomyces cerevisiae*)	DNA binding, DNA topoisomerase (ATP-hydrolyzing) activity, ATP binding, hydrolase activity
18131	Notch gene homolog 3 (*Drosophila*)	DNA binding, transcription factor activity, receptor activity, calcium ion binding, protein binding
72693	Zinc finger, CCHC domain containing 12	Nucleic acid binding, zinc ion binding, metal ion binding
278672	RIKEN cDNA 1110051B16 gene	DNA binding, transcription factor activity, sequence-specific DNA binding
13655	Early growth response 3	Nucleic acid binding, DNA binding, zinc ion binding, metal ion binding
210104	cDNA sequence BC043301	Zinc ion binding, metal ion binding, nucleic acid binding
320790	Chromodomain helicase DNA binding protein 7	—
67344	Tctex1 domain containing 1	—
15463	HIV-1 Rev-binding protein	DNA binding, zinc ion binding, metal ion binding
272382	SpiB transcription factor (Spi1/PU.1 related)	DNA binding, transcription factor activity, sequence-specific DNA binding, transcription factor activity

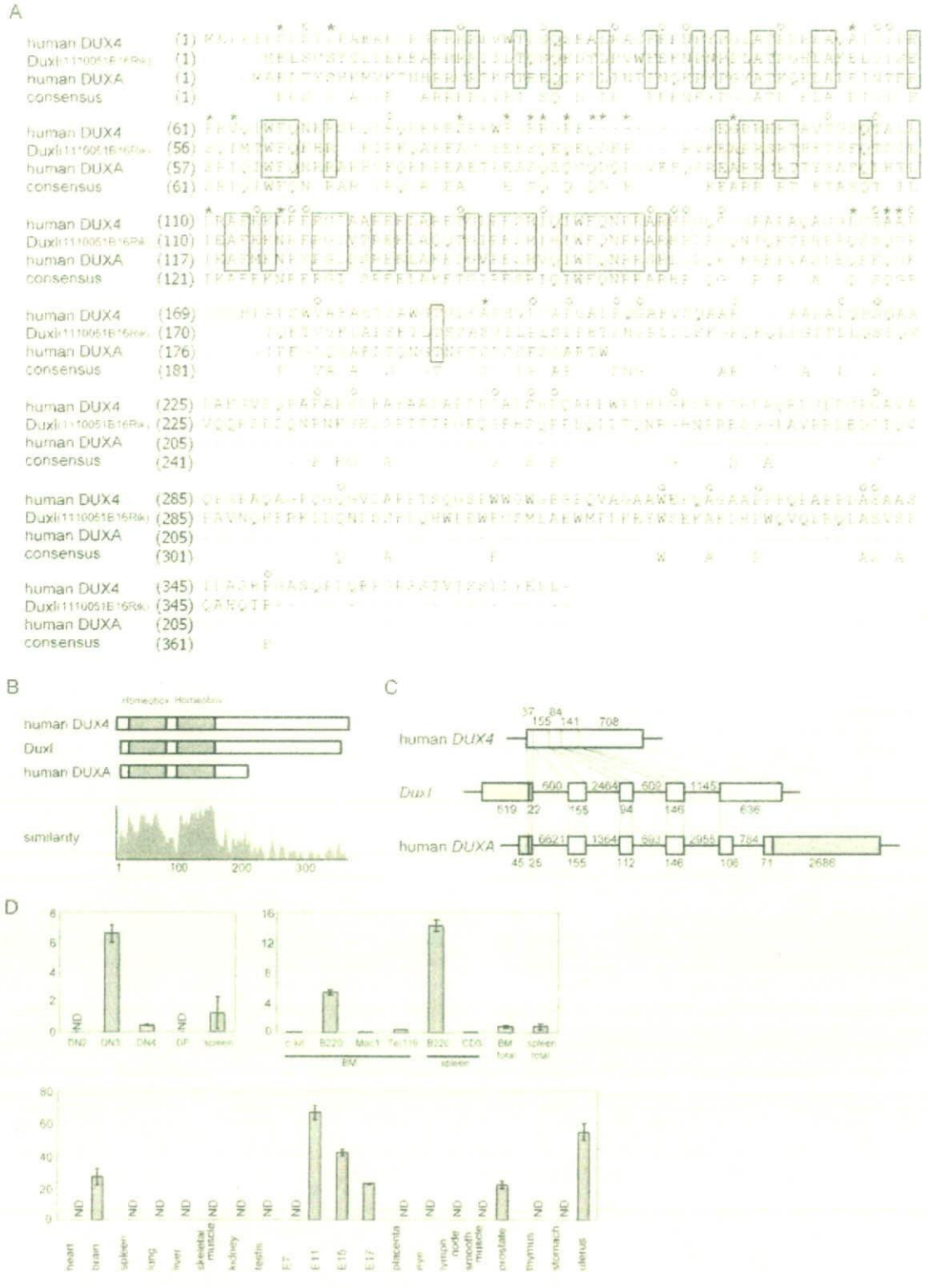

FIGURE 4. Structure of Duxl and similarity to its candidate orthologs. *A.* Homology of Duxl to its putative human orthologs (DUXA and DUX4) in amino acid sequences in which boxes indicate the identical amino acids among the three genes, asterisks indicate the common amino acids between Duxl and DUXA, and open circles indicate the common amino acids between Duxl and DUX4. *B.* Structure of predicted Duxl protein with its similarity to DUXA and DUX4. *C.* Gene structures of *Duxl* in mouse and *DUXA* and *DUX4* in humans. *D.* Distribution of *Duxl* expression in various tissues and cell lineages as determined by qPCR. The amount of transcript of *Duxl* was normalized to the amount of 18S rRNA in each tissue/population and is shown relative to levels in the total splenocytes in the *upper right panel*. Data shown are the average ± SD from triplicate samples.

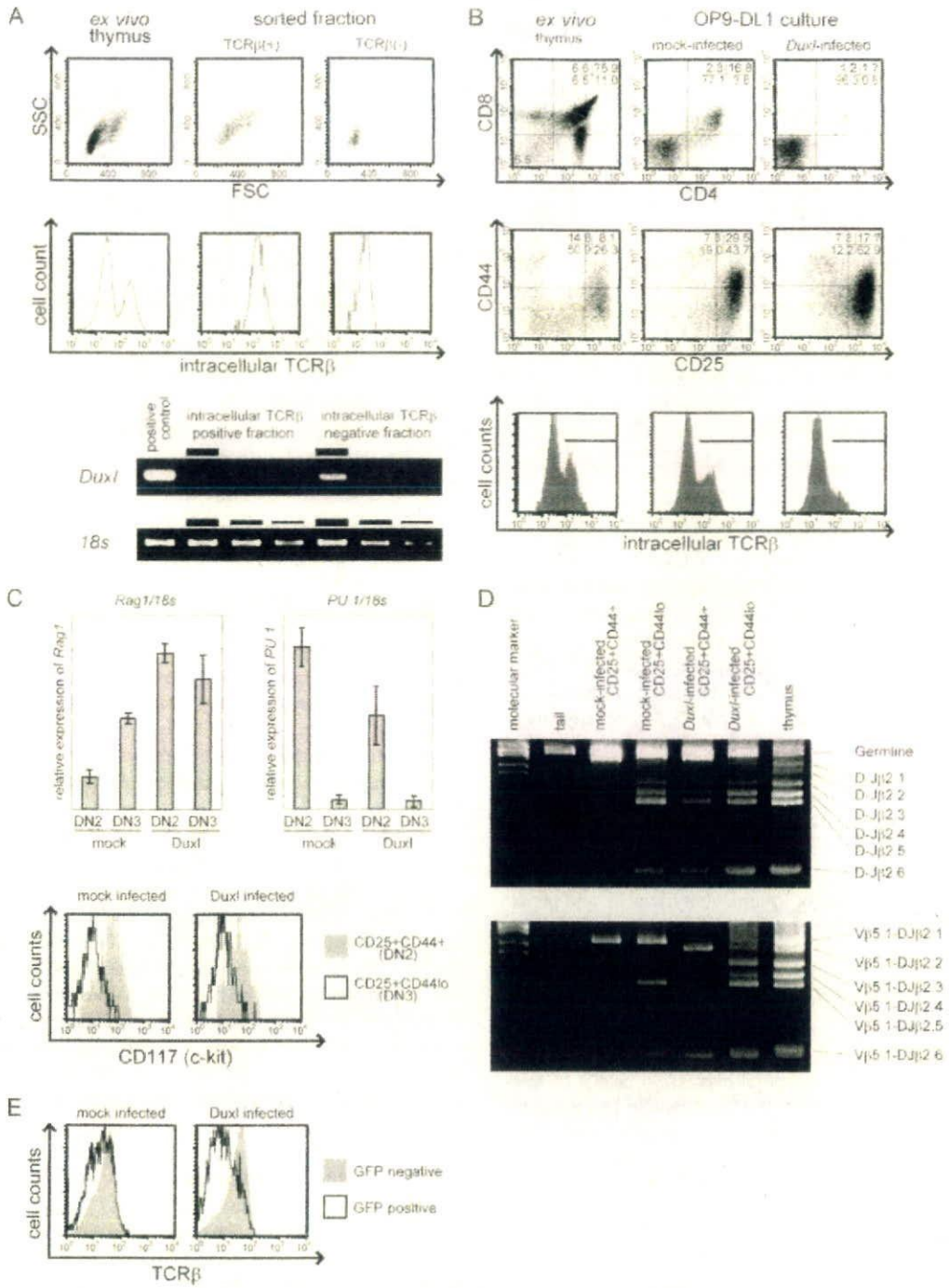

FIGURE 5. Effect of *Dux1* transduction on thymocyte development in OP9-DL culture. *A*, DN3 cells from normal mice were FACS sorted according to the expression of iTCRβ chain (*upper two panels*) and expression of *Dux1* in iTCRβ$^+$ and iTCRβ$^-$ cells was examined by qPCR analysis. Resultant PCR products were electrophoresed on 4% agarose gel (*lower two panels*). *B*, *Dux1* was transduced into FL cells and the development into thymocytes was examined by FACS analysis for their expression of CD4/CD8 (*upper panels*), CD25/CD44 (*middle panels*), and iTCRβ chains (*bottom panels*) in ex vivo thymus (*left*), mock-infected cultured FL cells (*middle*), and *Dux1*-transduced cultured FL cells (*right*). *C*, Expression levels of *Rag1* and *PU.1* were examined by qPCR using RNA isolated from the sorted GFP$^+$CD25$^+$CD44$^+$ fraction and GFP$^+$CD25$^+$CD44^{low} fraction of mock- or *Dux1*-transduced FL cells cultured on OP9-DL1 (*upper panels*). The amount of transcript of *Rag1* and *PU.1* was normalized to the amount of 18S rRNA in each population and is shown as relative values. Data shown are the average ± SD from triplicate samples. Expression of CD117 was analyzed with FACS (*lower panels*). *D*, TCRβ rearrangement status was analyzed by PCR using DNA isolated from the sorted GFP$^+$CD25$^+$CD44$^+$ fraction and GFP$^+$CD25$^+$CD44^{low} fraction of mock- or *Dux1*-transduced FL cells cultured on OP9-DL1. *E*, *Dux1* was overexpressed in FL cells or AKR1 cells using retrovirus vector, and the expression levels of TCRβ in the GFP$^+$ fraction were examined by FACS.

(IMMD) (27). Because of the small numbers of genes within the clusters CL5 and CL6, both clusters were analyzed after being combined together. In this pathway analysis, several sets of functionally related genes or pathways were extracted from each cluster. Table I shows a list of pathways extracted in high significance values ($p < 0.01$). For examples, genes involved in c-kit signaling

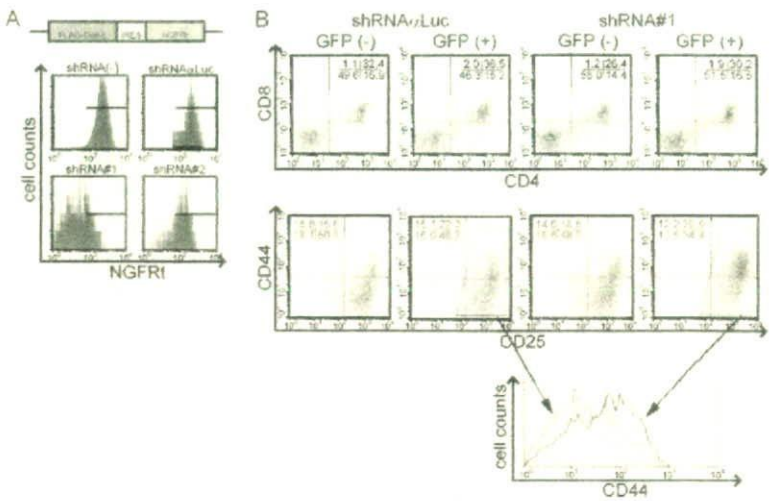

FIGURE 6. Effect of *Duxl* knockdown on expression levels of CD44. *A*, NIH3T3 cells stably expressing FLAG-Duxl-IRES-NGFRt (*upper left*) were infected with retroviruses encoding shRNA directed against a luciferase sequence (αLuc: *upper right*) or shRNA against a DuxL sequence #1 (*lower left*) and sequence #2 (*lower right*). Transduction of shRNA was tracked by GFP. NGFRt expression on GFP^+ cells was analyzed by flow cytometry. *B*, FL cells infected with Luc-shRNA retroviruses (αLuc) or Duxl-shRNA retroviruses (sequence #1) were cultured on OP9-DL1 cells and expression of CD4/CD8 (*upper panels*) and CD25/CD44 (*lower panels*) was analyzed after 17 days. Dot plots are gated on $GFP^-CD4^-CD8^-$ cells or $GFP^+CD4^-CD8^-$ cells.

($p = 0.00147$) and CD44 signaling ($p = 0.0086$) were extracted from the CL1 cluster, reflecting the fact that c-*kit* and *CD44* are down-regulated in DN3 and also indicating that the pathways extracted from CL1 are expected to be inactivated in DN3. Other pathways that are known to undergo dynamic regulations during DN thymocyte development were also extracted from different clusters, including the Runx1 pathway from CL2 (gradually down-regulated), NFAT and AP-1 pathways from CL3 (down-regulated in DN4), TCRαβ, NF-κB, and Notch pathways from CL4 (transiently activated in DN3), and Ets and TCRαβ pathways from CL5/CL6 (up-regulated in DN3 or DN4). The other pathways that were extracted with high significance values but have not been previously implicated in T cell development include those related to platelet-derived growth factor, γ-aminobutyric acid, and peroxisome proliferator-activated receptor α.

Expression and structure of 110051B16Rik *or* Duxl *gene*

In view of clarifying gene regulations that operate during thymocyte development, of particular interest are genes included in CL4, because they show a unique temporal expression profile of transient induction or up-regulation in DN3 and contain key genes for the thymocyte development or β selection, including *pre-TCRα*, *SpiB*, *Egr3*, and *Notch3* (33–37) (Fig. 2*B*). Especially, we were interested in genes that were thought to be involved in transcriptional regulations (Table II). Among these genes, we focused on a gene, *1110051B16Rik*, which encodes a putative transcription factor with unknown functions (Table II and Fig. 4*A*). It has an open reading frame of 1050 nt and the predicted protein shares structural features and sequence similarities with the families of double homeobox proteins that are thought to have rapidly diverged during evolution (Fig. 4*B*) (38). Although currently no definite human ortholog is uniquely identified due to incomplete sequence homology to human sequences, it shows the highest similarity with human DUXA or DUX4 in their amino acids sequence and gene structure with DUXA (Fig. 4, *B* and *C*), and thus, was named as *Duxl* (*Dux* in lymphoid lineage). In hemopoietic compartments, expression of *Duxl* is largely restricted to DN3 thymocytes and B220-positive B cells in bone marrow and spleen, implicating their functional roles in both subsets of lymphocytes, although it is also expressed in embryos from mid to late gestation, as well as in other non-hemopoietic adult organs, including brain, prostate, and uterus (Fig. 4*D*).

Function of Duxl *in thymocyte development*

To get an insight into the role of *Duxl* in thymocyte development, we first examined its expression among the subpopulations of DN3 thymocytes by FACS sorting DN3 cells into two populations according to their expression of intracellular TCRβ (iTCRβ) chains (Fig. 5*A*, *upper panel*) (8, 39). In quantitative PCR (qPCR) analysis, *Duxl* transcripts were detected in the $iTCR\beta^-$ fraction (DN3a) but not in the $iTCR\beta^+$ fraction (DN3b), indicating that the *Duxl* expression is largely restricted to the DN3a fraction (Fig. 5*A*, *lower panel*).

To explore the effect of constitutively expressed *Duxl* on DN thymocyte differentiation, we transduced the *Duxl* cDNA into mouse FL cells using a bicistronic retrovirus vector with a marker GFP cDNA in the second cistron, and the *Duxl*-transduced FL cells were then assayed on the OP9-Delta-1 (OP9-DL1) culture system that can mimic intrathymic development of DP thymocytes from the FL hemopoietic progenitors (29). As previously described, the mock-infected FL cells generated substantial numbers of GFP^+ DP thymocytes after 15 days of culture on OP9-DL1 (Fig. 5*B*, *upper central panel*) (19). In contrast, the number of GFP^+ DP cells was dramatically reduced in the culture from the *Duxl*-transduced FL cells (Fig. 5*B*, *upper right panel*). Although the total cell numbers were not significantly different between both cultures, the GFP^+ cells from the *Duxl*-transduced culture generated an increased proportion of the $CD25^+CD44^{low}$ cells, compared with mock-infected cells (Fig. 5*B*, *middle panels*) that produced a similar proportion of $CD25^+CD44^{low}$ cells as GFP^- cells (data not shown). The increased proportion of $CD25^+CD44^{low}$ in *Duxl*-transduced culture was accompanied with a reduced DN2 population. The $GFP^+CD25^+CD44^{low}$ cells in the *Duxl*-transduced FL cells had a $Thy1^+$ $CD11c^-$ $Mac1^-NK1.1^-B220^-$ phenotype. We extracted total RNAs from mock-infected DN2 and DN3 and Duxl-introduced DN2 and DN3 cells cultured on OP9-DL1 and analyzed their expression of *Rag1*, *PU.1*, and c-kit using real-time PCR and flow cytometry, respectively. In Duxl-introduced DN2 and DN3 cells, expression of *Rag1* was enhanced and expression of *PU.1* was reduced, whereas no difference was observed in c-kit expression between mock-infected and Duxl-introduced cells (Fig. 5*C*). Therefore, Duxl is supposed to promote the DN2/DN3 transition in some respects, while not in others.

Mock-infected and *Duxl*-transduced $CD25^+CD44^{low}$ cells were sorted and their genomic DNA was analyzed by PCR for TCRβ

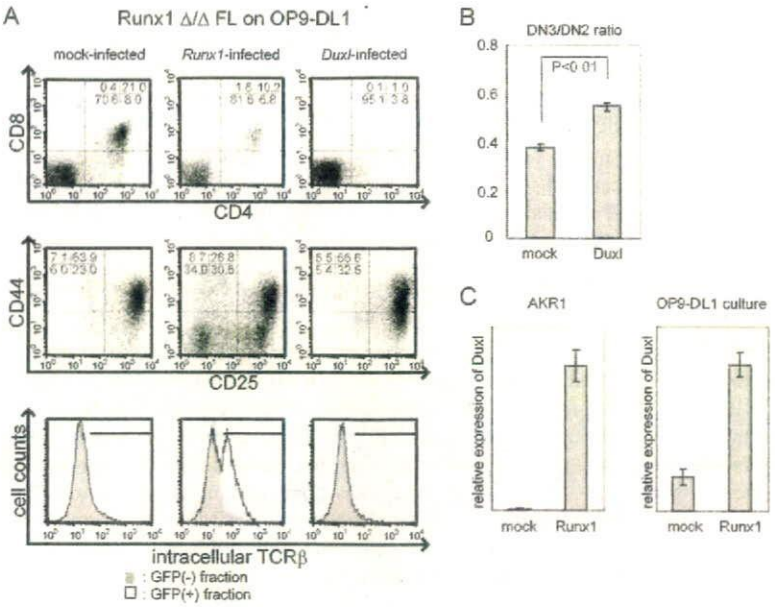

FIGURE 7. Effect of *Duxl* transduction on *Runx1*-deficient FL cells cultured on OP9-DL1. *A*, Expression of CD4/CD8 (*upper panels*), CD25/CD44 (*middle panels*), and iTCRβ (*lower panels*, solid line; GFP$^+$, gray shade; GFP$^-$) was examined in OP9-DL1 culture of FL cells harvested from the conditional knockout mice for the *Runx1* gene, in which two floxed alleles of *Runx1* are deleted by Cre recombinase induced from the *Lck* promoter. FL cells were transfected with mock virus (*left*) or with retrovirus expressing *Runx1* (*middle*) or *Duxl* (*right*). Representative FACS analyses are shown. *B*, Average DN3:DN2 ratio of mock-infected and *Duxl*-introduced cells in three independent cultures are shown with $\pm$SD. The two groups were compared using Student's *t* test. *C*, *Runx1* was overexpressed in FL cells or AKR1 cells using retrovirus vector, and the expression levels of *Duxl* in the GFP$^+$ fraction were examined by qPCR. The amount of transcript of *Duxl* was normalized to the amount of GAPDH RNA in each population and is shown as relative values. Data shown are the average $\pm$ SD from triplicate samples.

rearrangement. Interestingly, although the fraction of iTCRβ-positive cells in *Duxl*-transduced DN cells was substantially reduced compared with that in control DN cells (Fig. 5*B*, *lower panels*), the extent of DJβ and V-DJβ rearrangement in *Duxl*-transduced CD25$^+$CD44^{low} cells was comparable to those of control cells or rather slightly accelerated compared with control cells (Fig. 5*D*), indicating that Duxl actively represses expression of rearranged *TCRβ* genes. Indeed, expression of surface TCRβ in the mouse thymoma cell line AKR1, which exhibits the DP phenotype (CD4$^+$CD8$^+$TCR^{int}), was reduced after *Duxl* was introduced using retroviral vector (Fig. 5*E*). Taken together, these data suggest that the increased CD25$^+$CD44^{low} population in the *Duxl*-transduced cells show a DN3a-like phenotype and that constitutive expression of Duxl promotes the DN2/DN3 transition but compromises the process of β selection.

To further investigate the role of Duxl in DN2/DN3 transition, we knocked down the expression of *Duxl* in FL cells using RNA interference, in which FL cells were transfected with a retrovirus that produces shRNA directed against *Duxl* and examined for their development in the OP9-DL1 culture. Among three different shRNAs (sequences 1–3), only sequence 1 shRNA showed significant RNA interference as confirmed by the reduced cell surface expression of a truncated form of nerve growth factor receptor (NGFRt) in NIH3T3, where the NGFRt was translated from a bicistronic message of the *Duxl-IRES-NGFRt* (Fig. 6*A*). When transduced with sequence 1 shRNA, FL cells showed a reduced DN2/DN3 transition in OP9-DL1 culture as determined by the higher levels of CD44 expression in GFP$^+$CD25$^+$ cells compared with GFP$^-$CD25$^+$ cells within the same culture or with GFP$^+$CD25$^+$ cells in the α*Luc* (mock)-infected culture (Fig. 6*B*).

It was of particular interest to explore a possible functional link between Duxl and Runx1, because we showed that Runx1 is essential for the DN2/DN3 transition (18). Thus, we tested whether Duxl can rescue the phenotype of Runx1 deficiency in developing thymocytes by transducing *Duxl* into FL cells from *Runx1*-deficient mice in OP9-DL1 culture. Although mock-transduced *Runx1*-deficinet FL cells showed the defective DN2/DN3 transition with impaired down-regulation of CD44 (19), *Duxl* transduction clearly increased the CD44^{low}CD25$^+$ population (Fig. 7, *A* and *B*), although the latter population still did not express intracellular TCRβ chains (Fig. 7*A*, *lower panels*). Thus, *Duxl* can induce down-regulation of CD44 in a *Runx1*-independent manner. We further examined whether *Duxl* is induced by Runx1 or not, where *Runx1* was transduced into normal FL cells in OP9-DL1 culture or an AKR1 cell line, and expression of *Duxl* was measured by quantitative PCR 72 h after the transduction. In both experiments, *Runx1*-transduced GFP$^+$ cells showed a marked increase in *Duxl* expression (Fig. 7*C*). These results indicated that Runx1 promotes the DN2/DN3 transition by, at least in part, regulating expression of *Duxl*, although it is not clear whether this regulation is direct or indirect.

Discussion

Several groups have investigated the gene expression profiles of developing thymocytes (40–43). Despite the differences in analytical methods, panels of genes to be examined, microarray used, thymocyte subpopulation analyzed, or species, the reported expression pattern of some representative genes such as *pTα*, *Nothc3*, *Egr3*, or *SpiB* were reproduced in our study (15, 40, 42). Furthermore, our study revealed a new aspect that was not described so far and adds to knowledge of the development of thymocytes. Intrathymic development of thymocytes is a dynamic process, during which immature thymocytes progressively commit to mature T cell differentiation accompanied by explosive expansions of their population. As such, we initially expected that this process be driven by induction of a large number of regulatory genes. Unexpectedly, however, our expression profiling clearly showed that gene regulation during early thymocyte development is characterized by mostly repressive activities of transcription, where 90% of the differentially expressed genes were eventually down-regulated. This means that as thymocytes differentiate from their progenitors, they become to use more limited sets of genes at least before differentiating into DP cells.

Among these down-regulated genes, a notable subset is a group of genes that show lineage-promiscuous expression. According to our somewhat arbitrary criteria, these genes account for a substantial proportion (20%) of down-regulated genes in DN2. In other words, immature thymocytes in the DN2 subset simultaneously

express genes from different lineage-affiliated programs. Such lineage-promiscuous states were previously implicated for multipotent hemopoietic progenitors based on the observation for limited numbers of lineage-affiliated genes (44–46). We confirmed this for DN thymocytes by analyzing expression of a large number of genes and also evaluated their temporal changes during the course of their development. Of note is that mature T cell-related genes expressed in DN2 are also transiently repressed during DN2 to DN4 progression. Because our criteria of >10 times higher expression than the median could be too strict, more DN2 genes undergoing down-regulation thereafter could be explained under this framework.

The interpretation of this lineage-promiscuous gene expression in very immature thymocytes is elusive, but it may be speculated that these immature cells are conditioned or primed before their differentiation into particular lineages and that these "primed" states are represented by expression of multiple lineage-affiliated genes. In this scenario, the lineage-promiscuous expression of DN thymocytes could be related to plasticity of these cells. DN2 thymocytes are committed to T lineage to a certain extent but still can give rise to other lineages such as NK cells (10, 12–14), which might reflect the observation that NK cell-affiliated genes were more slowly down-regulated during DN thymocyte development (Fig. 3*C*). Alternatively, lineage-promiscuous gene expression could be explained by the heterogeneous lineage potential at the level of cellular complexity. Because $CD44^{+}LCD25^{low}$ DN1 cells are composed of multiple subsets with distinct differentiation capacity that could be identified with additional c-kit and CD24 staining (11), DN1 cells express a large number of lineage-promiscuous genes down-regulated at the DN1/DN2 transition (40). This may be also the case with DN2 cells we have sorted without c-kit and CD24 staining in this study. Although, the observation that differentiation capacity of DN1 cells with lysozyme expression was similar to that of DN1 cells without lysozyme expression (47) strongly supports the former interpretation, it is necessary to examine the gene expression profiles of each single DN cell, as well as more precise sorting with c-kit and CD24 staining, to know which interpretation is more accurate.

Computational mapping of differentially expressed genes showing similar temporal profiles (CL1–CL6) to known functional pathways seems to be effective to extract the relevant pathways in DN thymocyte development, because it successfully extracted well-characterized pathways in developing thymocytes, such as those related to *Runx1*, *TCRβ*, *NF-κB*, *NFAT*, and *Notch* genes. In addition, it also implicated the presence of several previously unknown pathways that might be regulated during thymocyte development. Although further evaluations are required, finding of these potentially relevant pathways will provide valuable clues to the exploring molecular mechanisms of regulation of DN thymocyte development.

During DN thymocyte development, only a minority of genes was newly induced, among which those included in the CL4 cluster were of our particular interest, because they were transiently expressed within a window in DN3 and contained several well-known pathways and genes that are critical for thymocyte development. From a few candidate genes within this cluster that are potentially involved in transcriptional regulations, we identified a novel double homeobox gene, named *Duxl*, which is transiently expressed in DN3a and could have critical roles in regulation of DN thymocyte development. When constitutively expressed in FL cells in OP9-DL1 culture, Duxl enhances the proportion of $CD44^{low}CD25^{+}$ thymocytes that mimic the phenotype of DN3a cells with severely reduced DP cells, while the knockdown of Duxl in FL cells impaired the proper DN2/DN3 transition. Thus, Duxl is thought to play an important role in promoting the DN2/DN3 transition in the development of DN thymocytes.

In early thymocyte development, Duxl has several functional similarities to SpiB, in that it is highly expressed in DN3a as well as B cells, and that normal thymocyte development is impaired when constitutively expressed or knocked down in early thymocytes (33). These observations may implicate a possible functional link between SpiB and Duxl, which should be addressed in further analysis. For example, SpiB also regulates B cell differentiation, implicating that Duxl is also involved in development of B cells.

Of particular interest is the finding that *Duxl* is induced by Runx1 expression and can partially rescue the *Runx1*-deficient phenotype with regard to the down-regulation of cell surface CD44, indicating that Duxl is one of the effecter molecules involved in the DN2/DN3 transition that is regulated by Runx1. The detailed molecular mechanism of the DN2/DN3 transition is largely unknown because only limited numbers of mice strains, namely, *Runx1*-deficient mice (18) and *pTα/common cytokine receptor γ-chain* double knock-out mice (20), exhibit the maturational block between the DN2 and the DN3 stage. However, considering the fact that during the transition from the DN2 to the DN3 stage, the *TCRβ* gene is rearranged and αβT cell, γδT cell, and NK cell lineages begin to diverge (10, 48, 49), the DN2/DN3 transition, on which our findings could shed light, is supposed to be a crucial developmental step.

On the other hand, the interpretation of the severely reduced production of DP cells associated with constitutive expression of *Duxl* in FL cells is complicated in a context of its physiological roles. Since the *Duxl*-transduced FL cells seem to show a maturational block at the transition between DN3a and DN3b, during which Duxl undergoes down-regulation in vivo, it may be postulated that *Duxl* being normally down-regulated in this step is important for the β selection and subsequent DP thymocyte production. We may safely conclude that the precise regulation of Duxl expression is essential for DP thymocytes to be normally generated in OP9 culture, but its physiological functions in β selection and DP thymocyte generation are still elusive. To address this, more sophisticated experimental approaches with precisely targeted expression of *Duxl* in vivo should be required.

In conclusion, through the gene expression profiling of chronologically discrete subsets of DN thymocytes, we demonstrated the predominantly repressive gene regulation during DN thymocyte development along with its implication in lineage-promiscuous gene expression in immature thymocytes. Among the gene cluster showing transient expression in DN3, we identified *Duxl*, a novel double homeobox gene, that is induced by Runx1 and involved in regulation of DN thymocyte development.

Acknowledgments

We thank J. C. Zúñiga-Pflücker for OP9-DL1 stromal cells, H. Nakauchi for pGCDNsam retroviral vector, M. Miyagishi for designing shRNA, and T. Kitamura for PlatE packaging cells.

Disclosures

The authors have no financial conflict of interest.

References

1. Ceredig, R., and T. Rolink. 2002. A positive look at double-negative thymocytes. *Nat. Rev. Immunol.* 2: 888–897.
2. Pearse, M., L. Wu, M. Egerton, A. Wilson, K. Shortman, and R. Scollay. 1989. A murine early thymocyte developmental sequence is marked by transient expression of the interleukin 2 receptor. *Proc. Natl. Acad. Sci. USA* 86: 1614–1618.
3. Godfrey, D. I., J. Kennedy, T. Suda, and A. Zlotnik. 1993. A developmental pathway involving four phenotypically and functionally distinct subsets of $CD3^{-}CD4^{-}CD8^{-}$ triple-negative adult mouse thymocytes defined by CD44 and CD25 expression. *J. Immunol.* 150: 4244–4252.

4. Penit, C., B. Lucas, and F. Vasseur. 1995. Cell expansion and growth arrest phases during the transition from precursor ($CD4^-8^-$) to immature ($CD4^+8^+$) thymocytes in normal and genetically modified mice. *J. Immunol.* 154: 5103–5113.
5. Godfrey, D. I., J. Kennedy, P. Mombaerts, S. Tonegawa, and A. Zlotnik. 1994. Onset of TCR-β gene rearrangement and role of TCR-β expression during $CD3^-CD4^-CD8^-$ thymocyte differentiation. *J. Immunol.* 152: 4783–4792.
6. Mombaerts, P., A. R. Clarke, M. A. Rudnicki, J. Iacomini, S. Itohara, J. J. Lafaille, L. Wang, Y. Ichikawa, R. Jaenisch, M. L. Hooper, et al. 1992. Mutations in T-cell antigen receptor genes α and β block thymocyte development at different stages. *Nature* 360: 225–231.
7. Shinkai, Y., S. Koyasu, K. Nakayama, K. M. Murphy, D. Y. Loh, E. L. Reinherz, and F. W. Alt. 1993. Restoration of T cell development in RAG-2-deficient mice by functional TCR transgenes. *Science* 259: 822–825.
8. Hoffman, E. S., L. Passoni, T. Crompton, T. M. Leu, D. G. Schatz, A. Koff, M. J. Owen, and A. C. Hayday. 1996. Productive T-cell receptor β-chain gene rearrangement: coincident regulation of cell cycle and clonality during development in vivo. *Genes Dev.* 10: 948–962.
9. Allman, D., A. Sambandam, S. Kim, J. P. Miller, A. Pagan, D. Well, A. Meraz, and A. Bhandoola. 2003. Thymopoiesis independent of common lymphoid progenitors. *Nat. Immunol.* 4: 168–174.
10. Schmitt, T. M., M. Ciofani, H. T. Petrie, and J. C. Zúñiga-Pflücker. 2004. Maintenance of T cell specification and differentiation requires recurrent notch receptor-ligand interactions. *J. Exp. Med.* 200: 469–479.
11. Porritt, H. E., L. L. Rumfelt, S. Tabrizifard, T. M. Schmitt, J. C. Zúñiga-Pflücker, and H. T. Petrie. 2004. Heterogeneity among DN1 prothymocytes reveals multiple progenitors with different capacities to generate T cell and non-T cell lineages. *Immunity* 20: 735–745.
12. Ikawa, T., H. Kawamoto, S. Fujimoto, and Y. Katsura. 1999. Commitment of common T/natural killer (NK) progenitors to unipotent T and NK progenitors in the murine fetal thymus revealed by a single progenitor assay. *J. Exp. Med.* 190: 1617–1626.
13. Wu, L., C. L. Li, and K. Shortman. 1996. Thymic dendritic cell precursors: relationship to the T lymphocyte lineage and phenotype of the dendritic cell progeny. *J. Exp. Med.* 184: 903–911.
14. Manz, M. G., D. Traver, T. Miyamoto, I. L. Weissman, and K. Akashi. 2001. Dendritic cell potentials of early lymphoid and myeloid progenitors. *Blood* 97: 3333–3341.
15. David-Fung, E. S., M. A. Yui, M. Morales, H. Wang, T. Taghon, R. A. Diamond, and E. V. Rothenberg. 2006. Progression of regulatory gene expression states in fetal and adult pro-T-cell development. *Immunol. Rev.* 209: 212–236.
16. Radtke, F., A. Wilson, G. Stark, M. Bauer, J. van Meerwijk, H. R. MacDonald, and M. Aguet. 1999. Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity* 10: 547–558.
17. Michie, A. M., and J. C. Zúñiga-Pflücker. 2002. Regulation of thymocyte differentiation: pre-TCR signals and β-selection. *Semin. Immunol.* 14: 311–323.
18. Ichikawa, M., T. Asai, T. Saito, G. Yamamoto, S. Seo, I. Yamazaki, T. Yamagata, K. Mitani, S. Chiba, H. Hirai, et al. 2004. AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nat. Med.* 10: 299–304.
19. Kawazu, M., T. Asai, M. Ichikawa, G. Yamamoto, T. Saito, S. Goyama, K. Mitani, K. Miyazono, S. Chiba, S. Ogawa, et al. 2005. Functional domains of Runx1 are differentially required for $CD4^+$ repression, TCRβ expression, and CD4/8 double-negative to CD4/8 double-positive transition in thymocyte development. *J. Immunol.* 174: 3526–3533.
20. Di Santo, J. P., I. Aifantis, E. Rosmaraki, C. Garcia, J. Feinberg, H. J. Fehling, A. Fischer, H. von Boehmer, and B. Rocha. 1999. The common cytokine receptor γ chain and the pre-T cell receptor provide independent but critically overlapping signals in early $\alpha\beta$ T cell development. *J. Exp. Med.* 189: 563–574.
21. Li, C., and W. H. Wong. 2001. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA* 98: 31–36.
22. Li, C., and W. H. Wong. 2001. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome Biol.* 2: 0032.0031–0032.0011.
23. Tusher, V. G., R. Tibshirani, and G. Chu. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* 98: 5116–5121.
24. de Hoon, M. J., S. Imoto, J. Nolan, and S. Miyano. 2004. Open source clustering software. *Bioinformatics* 20: 1453–1454.
25. Herrero, J., F. Al-Shahrour, R. Diaz-Uriarte, A. Mateos, J. M. Vaquerizas, J. Santoyo, and J. Dopazo. 2003. GEPAS: A web-based resource for microarray gene expression data analysis. *Nucleic Acids Res.* 31: 3461–3467.
26. Herrero, J., J. M. Vaquerizas, F. Al-Shahrour, L. Conde, A. Mateos, J. S. Diaz-Uriarte, and J. Dopazo. 2004. New challenges in gene expression data analysis and the extended GEPAS. *Nucleic Acids Res.* 32: W485–W491.
27. Sato, H., S. Ishida, K. Toda, R. Matsuda, Y. Hayashi, M. Shigetaka, M. Fukuda, Y. Wakamatsu, and A. Itai. 2005. New approaches to mechanism analysis for drug discovery using DNA microarray data combined with KeyMolnet. *Curr. Drug Discov. Technol.* 2: 89–98.
28. Boyle, E. I., S. Weng, J. Gollub, H. Jin, D. Botstein, J. M. Cherry, and G. Sherlock. 2004. GO: Termfinder–open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes. *Bioinformatics* 20: 3710–3715.
29. Schmitt, T. M., and J. C. Zúñiga-Pflücker. 2002. Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. *Immunity* 17: 749–756.
30. Anderson, S. J., K. M. Abraham, T. Nakayama, A. Singer, and R. M. Perlmutter. 1992. Inhibition of T-cell receptor β-chain gene rearrangement by overexpression of the non-receptor protein tyrosine kinase $p56^{lck}$. *EMBO J.* 11: 4877–4886.
31. Miyagishi, M., and K. Taira. 2003. Strategies for generation of an siRNA expression library directed against the human genome. *Oligonucleotides* 13: 325–333.
32. Su, A. I., T. Wiltshire, S. Batalov, H. Lapp, K. A. Ching, D. Block, J. Zhang, R. Soden, M. Hayakawa, G. Kreiman, M. P. Cooke, et al. 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. USA* 101: 6062–6067.
33. Lefebvre, J. M., M. C. Haks, M. O. Carleton, M. Rhodes, G. Sinnathamby, M. C. Simon, L. C. Eisenlohr, L. A. Garrett-Sinha, and D. L. Wiest. 2005. Enforced expression of Spi-B reverses T lineage commitment and blocks β-selection. *J. Immunol.* 174: 6184–6194.
34. Carleton, M., M. C. Haks, S. A. Smeele, A. Jones, S. M. Belkowski, M. A. Berger, P. Linsley, A. M. Kruisbeek, and D. L. Wiest. 2002. Early growth response transcription factors are required for development of $CD4^-CD8^-$ thymocytes to the $CD4^+CD8^+$ stage. *J. Immunol.* 168: 1649–1658.
35. Miyazaki, T. 1997. Two distinct steps during thymocyte maturation from $CD4^-CD8^-$ to $CD4^+CD8^+$ distinguished in the early growth response (Egr)-1 transgenic mice with a recombinase-activating gene-deficient background. *J. Exp. Med.* 186: 877–885.
36. Xi, H., R. Schwartz, I. Engel, C. Murre, and G. J. Kersh. 2006. Interplay between RORγt, Egr3, and E proteins controls proliferation in response to pre-TCR signals. *Immunity* 24: 813–826.
37. Bellavia, D., A. F. Campese, A. Vacca, A. Gulino, and I. Screpanti. 2003. Notch3, another Notch in T cell development. *Semin. Immunol.* 15: 107–112.
38. Booth, H. A., and P. W. Holland. 2007. Annotation, nomenclature, and evolution of four novel homeobox genes expressed in the human germ line. *Gene* 387: 7–14.
39. Taghon, T., M. A. Yui, R. Pant, R. A. Diamond, and E. V. Rothenberg. 2006. Developmental and molecular characterization of emerging β- and $\gamma\delta$-selected pre-T cells in the adult mouse thymus. *Immunity* 24: 53–64.
40. Tabrizifard, S., A. Olaru, J. Plotkin, M. Fallahi-Sichani, F. Livak, and H. T. Petrie. 2004. Analysis of transcription factor expression during discrete stages of postnatal thymocyte differentiation. *J. Immunol.* 173: 1094–1102.
41. Huang, Y. H., D. Li, A. Winoto, and E. A. Robey. 2004. Distinct transcriptional programs in thymocytes responding to T cell receptor, Notch, and positive selection signals. *Proc. Natl. Acad. Sci. USA* 101: 4936–4941.
42. Dik, W. A., K. Pike-Overzet, F. Weerkamp, D. de Ridder, E. F. de Haas, M. R. Baert, P. van der Spek, E. E. Koster, M. J. Reinders, J. J. van Dongen, et al. 2005. New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. *J. Exp. Med.* 201: 1715–1723.
43. Hoffmann, R., L. Bruno, T. Seidl, A. Rolink, and F. Melchers. 2003. Rules for gene usage inferred from a comparison of large-scale gene expression profiles of T and B lymphocyte development. *J. Immunol.* 170: 1339–1353.
44. Akashi, K., D. Traver, T. Miyamoto, and I. L. Weissman. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* 404: 193–197.
45. Hu, M., D. Krause, M. Greaves, S. Sharkis, M. Dexter, C. Heyworth, and T. Enver. 1997. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev.* 11: 774–785.
46. Nutt, S. L., B. Heavey, A. G. Rolink, and M. Busslinger. 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* 401: 556–562.
47. Laiosa, C. V., M. Stadtfeld, H. Xie, L. de Andres-Aguayo, and T. Graf. 2006. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBPα and PU.1 transcription factors. *Immunity* 25: 731–744.
48. Ciofani, M., G. C. Knowles, D. L. Wiest, H. von Boehmer, and J. C. Zúñiga-Pflücker. 2006. Stage specific and differential notch dependency at the $\alpha\beta$ and $\gamma\delta$ T lineage bifurcation. *Immunity* 25: 105–116.
49. Prinz, I., A. Sansoni, A. Kissenpfennig, L. Ardouin, M. Malissen, and B. Malissen. 2006. Visualization of the earliest steps of $\gamma\delta$ T cell development in the adult thymus. *Nat. Immunol.* 7: 995–1003.

Human Molecular Genetics, 2007, Vol. 16, No. 20 2494–2505
doi:10.1093/hmg/ddm205
Advance Access published on July 31, 2007

Evaluation of genome-wide power of genetic association studies based on empirical data from the HapMap project

Yasuhito Nannya[1,2,4], Kenjiro Taura[3], Mineo Kurokawa[1], Shigeru Chiba[2] and Seishi Ogawa[2,4,*]

[1]Department of Hematology/Oncology, [2]Department of Cell Therapy and Transplantation Medicine, Graduate School of Medicine and [3]Department of Information and Communication Engineering, Graduate School of Information Science, University of Tokyo, Tokyo 113-8655, Japan and [4]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Saitama 332-0012, Japan

Received April 7, 2007; Revised and Accepted July 22, 2007

With recent advances in high-throughput single nucleotide polymorphism (SNP) typing technologies, genome-wide association studies have become a realistic approach to identify the causative genes that are responsible for common diseases of complex genetic traits. In this strategy, a trade-off between the increased genome coverage and a chance of finding SNPs incidentally showing a large statistics becomes serious due to extreme multiple-hypothesis testing. We investigated the extent to which this trade-off limits the genome-wide power with this approach by simulating a large number of case-control panels based on the empirical data from the HapMap Project. In our simulations, statistical costs of multiple hypothesis testing were evaluated by empirically calculating distributions of the maximum value of the χ^2 statistics for a series of marker sets having increasing numbers of SNPs, which were used to determine a genome-wide threshold in the following power simulations. With a practical study size, the cost of multiple testing largely offsets the potential benefits from increased genome coverage given modest genetic effects and/or low frequencies of causal alleles. In most realistic scenarios, increasing genome coverage becomes less influential on the power, while sample size is the predominant determinant of the feasibility of genome-wide association tests. Increasing genome coverage without corresponding increase in sample size will only consume resources without little gain in power. For common causal alleles with relatively large effect sizes [genotype relative risk $\geq$1.7], we can expect satisfactory power with currently available large-scale genotyping platforms using realistic sample size ($\sim$1000 per arm).

INTRODUCTION

Genome-wide association studies have been proposed as a strategy to identify genetic factors with small to moderate genetic effects in the development of human diseases, as typically assumed for a common disease common variant (CDCV) model (1). In this strategy, a disease-associated locus is identified through single nucleotide polymorphisms (SNPs) that show 'significantly' different allele frequencies between affected (cases) and unaffected (controls) individuals, and a large number of SNPs are tested for association in an attempt to realistically identify such SNPs (2,3). Although only a theoretical perspective a decade ago (1), with the unprecedented advance in large-scale genotyping technologies (4–6), it has now become a realistic approach to exploring the genetic basis of human disease (7,8). In addition, recent efforts in the International HapMap Project to understand the genetic diversity among human populations (9,10) have greatly contributed to clarifying the extent to which the number of marker SNPs could be reduced to achieve given genome coverage, or how much genome coverage can be obtained with a given marker SNP set by optimally 'tagging' untyped SNPs based on the linkage disequilibrium (LD) observed in the human genome (11–16).

*To whom correspondence should be addressed to: Department of Cell Therapy and Transplantation Medicine, The 21st Century COE Program, Graduate School of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. Tel: +81 358008741; Fax: +81 358046261; Email: sogawa-tky@umin.ac.jp

© *The Author 2007. Published by Oxford University Press. All rights reserved.*
For Permissions, please email: journals.permissions@oxfordjournals.org