Table 6 Subsequent treatments after failure to respond to gefitinib ($n = 28$)

Gefitinib treatment	No. of Patients	1st regimen after gefitinib	No. of patients	2nd regimen after gefitinib	No. of patients
1st line	17	Plt doublet	5	Gem or Doce	2
				Gefitinib[a]	1
		VNR	1	—	-
2nd line[b]	4	Doce	2	Doce	1
		Plt doublet	1	Doce	1
2nd line	5	Doce	1	Gefitinib[a]	1
3rd line	2	—	—	—	—
Total	28		10		
Response			4/10		2/6

Doce = docetaxel; Gem = gemcitabine; Plt = platinum; VNR = vinorelbine. [a]Both patients had an SD response after gefitinib re-treatment. [b]First regimen as systemic chemotherapy after adjuvant treatment.

Table 7 Bronchial alveolar carcinoma (BAC) features and *EGFR* mutation status

	EGFR mutation		
	+	–	*P*-value
Surgically resected adenocarcinoma case	12	24	
BAC component			
Yes	8	17	1.0
No	4	7	
Micropapillary pattern			
Yes	4	12	0.48
No	8	12	
Mucin production			
Yes	1	5	1.0
No	11	19	

EGFR = epidermal growth factor receptor.

background factors. In the present study, *EGFR* mutations were detected in 16 out of 40 (40%) female never smokers with adenocarcinoma who underwent the screening process, and 14 out of these 16 patients (88%) achieved a response after undergoing gefitinib therapy. We could not compare the predictive powers of clinical predictors and the *EGFR* mutation status with regard to the clinical benefits of gefitinib in this study. Thus, the need for *EGFR* mutation testing among clinically favourable patients remains uncertain. Decisions regarding the first-line therapy of choice for patients with *EGFR* mutations or a clinically favourable profile (nonsmoker with adenocarcinoma) must also await the results of an ongoing randomised phase III study in an Asian population (IPASS: Iressa Pan-Asian Study) comparing platinum doublets with gefitinib.

In contrast, 50% of the men, 67% of the smokers and 63% of the men who were smokers achieved a PR in this study. Furthermore, one female nonsmoker with squamous cell carcinoma also responded to gefitinib. The histological type of this tumour was reassigned by a pulmonary pathologist, and the tumour was finally confirmed to be a squamous cell carcinoma. Squamous cell carcinoma harbouring an *EGFR* mutation is rarely seen but has been previously reported (Asahina *et al*, 2006). In a Japanese phase II trial of gefitinib for unselected chemotherapy-naive patients (Niho *et al*, 2006), the response rates among smokers, men, and patients with nonadenocarcinoma were 19, 13 and 10%, respectively. Thus, NSCLC patients who are either smokers, men or have a nonadenocarcinoma histology are unlikely to receive gefitinib treatment as a first-line treatment instead of standard chemotherapies (platinum doublets), which yield a response rate of about 30% (Schiller *et al*, 2002). Therefore, *EGFR* mutation screening may have a higher impact on the selection of responders to gefitinib treatment among these kinds of Asian patient subset (for example, smokers with adenocarcinoma, and nonsmoking men or women with nonadenocarcinoma).

The benefit of chemotherapy in general among patients with *EGFR* mutations, compared with *EGFR* mutation-negative patients, remains uncertain. Previous studies (Bell *et al*, 2005) have suggested that patients with *EGFR* mutations tend to be more sensitive to chemotherapy than those with wild-type *EGFR*. In the present study, 40 and 33% of the patients responded to first- and second-line chemotherapy regimens after gefitinib, respectively. These relatively high response rates for refractory NSCLC suggest that patients with an *EGFR* mutation-positive status are generally sensitive to chemotherapy. Large-scale multivariate analyses, using pooled data from prospective phase II or III trials in which the *EGFR* mutation status was clearly confirmed, are needed to clarify this point.

The toxicities observed in the present study were mostly tolerable. Most of the common adverse events, like rash, diarrhoea or hepatotoxicity, were mild and subsided after gefitinib administration was paused for a short period. One male smoker with adenocarcinoma died of ILD. Thus, even among patients who are selected based on their *EGFR* mutation status, men or smokers may still be at risk for developing ILD; therefore, biomarkers to predict ILD are needed.

Patients with exon 19 mutations tended to have a higher response rate than those with a missense mutation in exon 21, consistent with the findings of previous reports (Jackman *et al*, 2006; Riely *et al*, 2006). The Spanish Lung Cancer Group also reported on a prospective phase II study of erlotinib in advanced NSCLC patients with *EGFR* mutations (Paz-Ares *et al*, 2006). The overall response rate was 82%. They also showed a difference in response rates between patients with mutations in exons 19 and 21 (95 and 67%, respectively). Exon 11 c-kit mutations are more closely correlated with a good prognosis in patients with gastrointestinal stromal tumour, who may benefit from lower doses of imatinib, whereas patients with exon 9 mutations may require higher doses (Debiec-Rychter *et al*, 2006). In the case of *EGFR*, functional differences between mutation types may also exist.

We found no discernible associations between the *EGFR* mutation frequency and the presence of a BAC component. Several reports, including that of Hirsch *et al* (2005) suggest that a higher *EGFR* copy number is correlated with BAC histological features. We also found an association between a high *EGFR* copy number and the presence of a BAC component, even though the number of specimens examined was relatively small. In a study on erlotinib, the presence of a BAC component was clearly associated with *EGFR* amplification. As the *EGFR* mutation rate is lower in western populations than in Asian populations, the *EGFR* gene copy number might be a more useful biomarker in western populations, especially with regard to the use of erlotinib.

In conclusion, gefitinib treatment for patients with advanced NSCLC harbouring an *EGFR* mutation demonstrated a promising activity in patients with a good performance status. Patient screening according to *EGFR* mutation status may be a useful tool in daily practice and will likely have a great impact on the selection of patients who are likely to benefit from gefitinib treatment.

ACKNOWLEDGEMENTS

We thank Masaru Sekijima, PhD (Director, Research Division for Advanced Technology, Kashima Laboratory, Mitsubishi Chemical Institute Ltd) for technical support with the *EGFR* mutation analyses, and Yuki Inoue, Erina Hatashita and Yuki Yamada for data management.

REFERENCES

Asahina H, Yamazaki K, Kinoshita I, Sukoh N, Harada M, Yokouchi H, Ishida T, Ogura S, Kojima T, Okamoto Y, Fujita Y, Dosaka-Akita H, Isobe H, Nishimura M, on behalf of the Hokkaido Lung Cancer Clinical Study Group (2006) A phase II trial of gefitinib as first-line therapy for advanced non-small cell lung cancer with epidermal growth factor receptor mutations. *Br J Cancer* **95:** 998–1004

Ando M, Okamoto I, Yamamoto N, Takeda K, Tamura K, Seto T, Ariyoshi Y, Fukuoka M (2006) Predictive factors for interstitial lung disease, antitumor response and survival in non-small-cell lung cancer patients treated with gefitinib. *J Clin Oncol* **24:** 2549–2556

Bell DW, Lynch TJ, Haserlat SM, Harris PL, Okimoto RA, Brannigan BW, Sgroi DC, Muir B, Riemenschneider MJ, Iacona RB, Krebs AD, Johnson DH, Giaccone G, Herbst RS, Manegold C, Fukuoka M, Kris MG, Baselga J, Ochs JS, Haber DA (2005) Epidermal growth factor receptor mutations and gene amplification in non-small-cell lung cancer: molecular analysis of the IDEAL/INTACT gefitinib trials. *J Clin Oncol* **23:** 8081–8092

Cappuzzo F, Hirsch FR, Rossi E, Bartolini S, Ceresoli GL, Bemis L, Haney J, Witta S, Danenberg K, Domenichini I, Ludovini V, Magrini E, Gregorc V, Doglioni C, Sidoni A, Tonato M, Franklin WA, Crino L, Bunn Jr PA, Varella-Garcia M (2005) Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst* **97:** 643–655

Debiec-Rychter M, Sciot R, Le Cesne A, Schlemmer M, Hohenberger P, van Oosterom AT, Blay JY, Leyvraz S, Stul M, Casali PG, Zalcberg J, Verweij J, Van Glabbeke M, Hagemeijer A, Judson I, EORTC Soft Tissue and Bone Sarcoma Group, The Italian Sarcoma Group, Australasian GastroIntestinal Trial Group (2006) KIT mutations and dose selection for imatinib in patients with advanced gastrointestinal stromal tumours. *Eur J Cancer* **42:** 1093–1103

Fukuoka M, Yano S, Giaccone G, Tamura T, Nakagawa K, Douillard JY, Nishiwaki Y, Vansteenkiste J, Kudoh S, Rischin D, Eek R, Horai T, Noda K, Takata I, Smit E, Averbuch S, Macleod A, Feyereislova A, Dong RP, Baselga J (2003) Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer. *J Clin Oncol* **21:** 2237–2246

Han SW, Kim TY, Lee KH, Hwang PG, Jeon YK, Oh DY, Lee SH, Kim DW, Im SA, Chung DH, Heo DS, Bang YJ (2006) Clinical predictors *versus* epidermal growth factor receptor mutation in gefitinib-treated non-small-cell lung cancer patients. *Lung Cancer* **54:** 201–207

Hirsch FR, Varella-Garcia M, Bunn Jr PA, Di Maria MV, Veve R, Bremmes RM, Barón AE, Zeng C, Franklin WA (2003) Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *J Clin Oncol* **21:** 3798–3807

Hirsch FR, Varella-Garcia M, McCoy J, West H, Xavier AC, Gumerlock P, Bunn Jr PA, Franklin WA, Crowley J, Gandara DR, Southwest Oncology Group (2005) Increased epidermal growth factor receptor gene copy number detected by fluorescence *in situ* hybridization associates with increased sensitivity to gefitinib in patients with bronchioloalveolar carcinoma subtypes: a Southwest Oncology Group Study. *J Clin Oncol* **23:** 6838–6845

Inoue A, Suzuki T, Fukuhara T, Maemondo M, Kimura Y, Morikawa N, Watanabe H, Saijo Y, Nukiwa T (2006) Prospective phase II study of gefitinib for chemotherapy-naïve patients with advanced non-small cell lung cancer with epidermal growth factor receptor gene mutations. *J Clin Oncol* **24:** 3340–3346

Jackman DM, Yeap BY, Sequist LV, Lindeman N, Holmes AJ, Joshi VA, Bell DW, Huberman MS, Halmos B, Rabin MS, Haber DA, Lynch TJ, Meyerson M, Johnson BE, Jänne PA (2006) Exon 19 deletion mutation of epidermal growth factor receptor are associated with prolonged survival in non-small cell lung cancer patients treated with gefitinib or erlotinib. *Clin Cancer Res* **12:** 3908–3914

Kris MG, Natale RB, Herbst RS, Lynch Jr TJ, Prager D, Belani CP, Schiller JH, Kelly K, Spiridonidis H, Sandler A, Albain KS, Cella D, Wolf MK, Averbuch SD, Ochs JJ, Kay AC (2003) Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small-cell lung cancer. *JAMA* **290:** 2149–2158

Kaneda H, Tamura K, Kurata T, Uejima H, Nakagawa K, Fukuoka M (2004) Retrospective analysis of the predictive factors associated with response and survival benefit of gefitinib in patients with advanced non-small cell lung cancer. *Lung Cancer* **46:** 247–254

Kobayashi S, Boggon TJ, Dayaram T, Jänne PA, Kocher O, Meyerson M, Johnson BE, Eck MJ, Tenen DG, Halmos B (2005) EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* **352:** 786–792

Lee DH, Han JY, Yu SY, Kim HY, Nam BH, Hong EK, Kim HT, Lee JS (2006) The role of gefitinib treatment for Korean never-smokers with advanced or metastatic adenocarcinoma of lung: a prospective study. *J Thorac Oncol* **1:** 965–971

Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* **350:** 2129–2139

Miller VA, Kris MG, Shah N, Patel J, Azzoli C, Gomez J, Krug LM, Pao W, Rizzo B, Tyson L, Venkatraman E, Ben-Porat L, Memoli N, Zakowski M, Rusch V, Heelan RT (2004) Bronchioloalveolar pathologic subtype smoking history predicts sensitivity to gefitinib in advanced non-small-cell lung cancer. *J Clin Oncol* **22:** 1103–1109

Niho S, Kubota K, Goto K, Yoh K, Ohmatsu H, Kakinuma R, Saijo N, Nishiwaki Y (2006) First-line single agent treatment with gefitinib in patients with advanced non-small-cell lung cancer: a phase II study. *J Clin Oncol* **24:** 64–69

Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE, Meyerson M (2004) EGFR mutation in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304:** 1497–1500

Paz-Ares L, Sanchez JM, Garcia-Velasco A, Masuti B, Majem M, Lopez-Vivanco G, Provencio M, Montes A, Amador M, Rosell R (2006) A prospective phase II trial if erlotinib in advanced non-small cell lung cancer (NSCLC) patients with mutations in the tyrosine kinase (TK) domain of the epidermal growth factor receptor (EGFR). *Proc Am Soc Clin Onc* **24**(Suppl): abstract 7020

Pao W, Miller VA, Zakowski MF, Doherty J, Politi KA, Sarkaria I, Singh B, Varmus H (2004) EGF receptor gene mutations are common in lung cancers from 'never smokers' and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA* **101:** 13306–13311

Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, Kris MG, Varmus H (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med* **2:** 1–11

Ranson M, Hammond LA, Ferry D, Kris M, Tullo A, Murray PI, Miller V, Averbuch S, Ochs J, Morris C, Feyereislova A, Swaisland H, Rowinsky EK (2002) ZD1839, a selective oral epidermal growth factor receptor-tyrosine kinase inhibitor, is well tolerated and active in patients with solid, malignant tumors: results of a phase I trial. *J Clin Oncol* **20:** 2240–2250

Riely GJ, Pao W, Pham D, Li AR, Rizvi N, Venkatraman ES, Zakowski MF, Kris MG, Ladanyi M, Millar VA (2006) Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clin Cancer Res* **12:** 839–844

Schiller JH, Harrington D, Belani CP, Langer C, Sandler A, Krook J, Zhu J, Jhonson DH, Eastern Cooperative Oncology Group (2002) Comparison of four chemotherapy regimens for advanced non-small cell lung cancer. *N Engl J Med* **346:** 92–98

Shepherd FA, Rodrigues Pereira J, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S, Compos D, Maoleekoonpiroj S, Smylie M, Martins R, van Kooten M, Dediu M, Findlay B, Tu D, Johnston D, Bezjak A, Clark G, Santabarbara P, Seymour L, National Cancer Institute of Canada Clinical Trials Group (2005) Erlotinib in previously treated non-small cell lung cancer. *N Engl J Med* **353:** 123–132

Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, Fong KM, Lee H, Toyooka S, Shimizu N, Fujisawa T, Feng Z, Roth JA, Herz J, Minna JD, Gazdar AF (2005) Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst* **97:** 339–346

Thatcher N, Chang A, Parikh P, Rodrigues Pereira J, Ciuleanu T, von Pawel J, Thongprasert S, Tan EH, Pemberton K, Archer V, Carroll K (2005) Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* **366:** 1527–1537

Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG (2000) New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst* **92:** 205–216

Yatabe Y, Hida T, Horio Y, Kosaka T, Takahashi T, Mitsudomi T (2006) A rapid, sensitive assay to detect EGFR mutation in small biopsy specimens from lung cancer. *J Mol Diagn* **8:** 335–341

Yoshida K, Yatabe Y, Young Ji P, Shimizu J, Horio Y, Matsuo K, Kosaka T, Mitsudomi T, Hida T (2007) Prospective validation for prediction of gefitinib sensitivity by epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer. *J Thoracic Oncol* **2:** 22–28

Oncogene (2007), 1–12

www.nature.com/onc

ORIGINAL ARTICLE

Aberrant expression of Fra-2 promotes CCR4 expression and cell proliferation in adult T-cell leukemia

T Nakayama[1], K Hieshima[1], T Arao[2], Z Jin[1], D Nagakubo[1], A-K Shirakawa[1], Y Yamada[3], M Fujii[4], N Oiso[5], A Kawada[5], K Nishio[2] and O Yoshie[1]

[1]Department of Microbiology, Kinki University School of Medicine, Osaka, Japan; [2]Department of Genome Science, Kinki University School of Medicine, Osaka, Japan; [3]Department of Laboratory Medicine, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan; [4]Division of Virology, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan and [5]Department of Dermatology, Kinki University School of Medicine, Osaka, Japan

Adult T-cell leukemia (ATL) is a mature CD4$^+$ T-cell malignancy etiologically associated with human T-cell leukemia virus type 1 (HTLV-1). Primary ATL cells frequently express CCR4 at high levels. Since HTLV-1 Tax does not induce CCR4 expression, transcription factor(s) constitutively active in ATL may be responsible for its strong expression. We identified an activator protein-1 (AP-1) site in the CCR4 promoter as the major positive regulatory element in ATL cells. Among the AP-1 family members, Fra-2, JunB and JunD are highly expressed in fresh primary ATL cells. Consistently, the Fra-2/JunB and Fra-2/JunD heterodimers strongly activated the CCR4 promoter in Jurkat cells. Furthermore, Fra-2 small interfering RNA (siRNA) or JunD siRNA, but not JunB siRNA, effectively reduced CCR4 expression and cell growth in ATL cells. Conversely, Fra-2 or JunD overexpression promoted cell growth in Jurkat cells. We identified 49 genes, including c-Myb, BCL-6 and MDM2, which were downregulated by Fra-2 siRNA in ATL cells. c-Myb, BCL-6 and MDM2 were also downregulated by JunD siRNA. As Fra-2, these proto-oncogenes were highly expressed in primary ATL cells but not in normal CD4$^+$ T cells. Collectively, aberrantly expressed Fra-2 in association with JunD may play a major role in CCR4 expression and oncogenesis in ATL.
Oncogene advance online publication, 10 December 2007; doi:10.1038/sj.onc.1210984

Keywords: adult T-cell leukemia; CCR4; Fra-2; JunD; c-Myb; MDM2; BCL-6

Introduction

Adult T-cell leukemia (ATL) is a highly aggressive malignancy of mature CD4$^+$CD25$^+$ T cells etiologically associated with human T-cell leukemia virus type 1 (HTLV-1; Yamamoto and Hinuma, 1985). HTLV-1 encodes a potent viral transactivator Tax that activates the HTLV-1 long terminal repeat (LTR) and also induces the expression of various cellular target genes, including those encoding cytokines, cytokine receptors, chemokines, cell adhesion molecules and nuclear transcriptional factors, collectively leading to the strong promotion of cell proliferation (Yoshida, 2001; Grassmann *et al.*, 2005). However, ATL develops after a long period of latency, usually several decades, during which oncogenic progression is considered to occur through the accumulation of multiple genetic and epigenetic changes (Matsuoka, 2003). Furthermore, circulating ATL cells usually do not express Tax and are considered to be independent of Tax (Matsuoka, 2003). Previously, Mori *et al.* have demonstrated the strong constitutive activation of nuclear factor kappa B (NF-κB) and activator protein-1 (AP-1) in primary ATL cells (Mori *et al.*, 1999, 2000). However, the molecular mechanisms of ATL oncogenesis still remain largely unknown.

CCR4 is a chemokine receptor known to be selectively expressed by Th2 cells, regulatory T cells (Treg) and skin-homing effector/memory T cells (Imai *et al.*, 1999; Iellem *et al.*, 2001; Yoshie *et al.*, 2001). Previously, we and others showed that ATL cells in the majority of cases are strongly positive for surface CCR4 (Yoshie *et al.*, 2002; Ishida *et al.*, 2003; Nagakubo *et al.*, 2007). Ishida *et al.* have also demonstrated a significant correlation of CCR4 expression with skin involvement and poor prognosis in ATL patients (Ishida *et al.*, 2003). Furthermore, several groups have reported that FOXP3, a forkhead/winged helix transcription factor and a specific marker of Treg (Hori *et al.*, 2003), is frequently expressed in ATL (Karube *et al.*, 2004; Matsubara *et al.*, 2005), supporting the notion that at least a fraction of ATL cases are derived from Treg.

It is also notable that primary ATL cells express CCR4 at levels much higher than normal resting CD4$^+$CD25$^+$ T cells (Nagakubo *et al.*, 2007). Given

Correspondence: Professor O Yoshie, Department of Microbiology, Kinki University School of Medicine, 377-2, Ohono-Higashi, Osaka-Sayama, Osaka 589-8511, Japan.
E-mail: o.yoshie@med.kindai.ac.jp
Received 29 May 2007; revised 29 October 2007; accepted 6 November 2007

that CCR4 is not inducible by Tax (Yoshie *et al.*, 2002), transcription factor(s) constitutively active in ATL cells may be responsible for CCR4 expression. Here, we demonstrate that Fra-2, one of the AP-1 family members (Shaulian and Karin, 2002; Eferl and Wagner, 2003), is aberrantly expressed in primary ATL cells. We further demonstrate that the Fra-2/JunD heterodimer plays a major role in both CCR4 expression and cell proliferation in ATL cells. Furthermore, we demonstrate that the proto-oncogenes c-Myb, BCL-6 and MDM2 (Oh and Reddy, 1999; Pasqualucci *et al.*, 2003; Vargas *et al.*, 2003) are the downstream target genes of the Fra-2/JunD heterodimer and are highly expressed in primary ATL cells. Thus, aberrantly expressed Fra-2 in association with JunD may be involved in ATL oncogenesis.

Results

Analysis of CCR4 promoter activity in ATL-derived cell lines

To examine the transcriptional regulation of CCR4 expression in ATL, we constructed a reporter plasmid carrying the CCR4 promoter region from −983 to +25 bp (the major transcriptional initiation site, +1) fused with the luciferase reporter gene. As shown in Figure 1a, pGL3-CCR4 (−983/+25) showed much stronger promoter activities in ATL cell lines (HUT102 and ST1) than in control human T-cell lines (MOLT-4 and Jurkat). We therefore generated a series of 5′-truncated promoter plasmids and examined their activity in ATL cell lines. As shown in Figure 1b, the promoter region from −151 to −96 bp was the major positive regulatory region in both cell lines. The TFSEARCH program (http://mbs.cbrc.jp/research/db/TFSEARCH.html) revealed various potential transcriptional elements in this region (Figure 1c). To identify the actual regulatory elements, we introduced a mutation in each potential element and examined the promoter activity in ATL cell lines. As shown in Figure 1d, a mutation at the AP-1 site or the GATA-3 site significantly reduced the promoter activity. Moreover, double mutations targeting both sites further reduced the promoter activity.

Constitutive expression of Fra-2, JunB and JunD in primary ATL cells

AP-1 is known to be involved in tumorigenesis (Shaulian and Karin, 2002; Eferl and Wagner, 2003), while GATA-3 regulates Th2-type gene expression (Rengarajan *et al.*, 2000). Therefore, we focused on AP-1 in the subsequent study. AP-1 constitutes a heterodimer of a member of the Fos family (c-Fos, FosB, Fra-1 and Fra-2) and a member of the Jun family (c-Jun, JunB and JunD) or a homodimer of the Jun family (Shaulian and Karin, 2002; Eferl and Wagner, 2003). Even though AP-1 was shown to be constitutively active in primary ATL cells (Mori *et al.*, 2000), it has not been clarified which members of AP-1 are actually

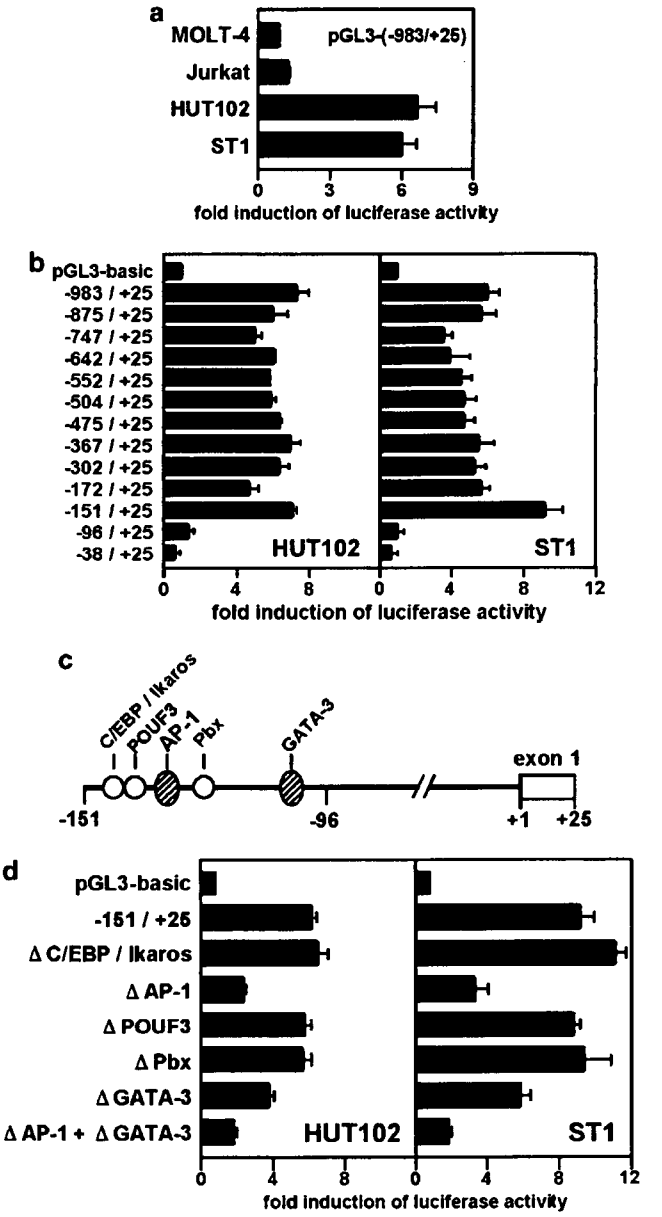


Figure 1 Identification of regulatory elements in the CCR4 promoter. Cells were transfected with pSV-β-galactosidase and pGL3-basic or pGL3-basic inserted with the CCR4 promoter regions as indicated. After 24–27 h, luciferase assays were performed. Promoter activation was expressed by the fold induction of luciferase activity in cells transfected with the CCR4 promoter–luciferase constructs versus cells transfected with the control pGL3-basic. Transfection efficiency was normalized by β-galactosidase activity. Each bar represents the mean ± s.e.m. from three separate experiments. (**a**) Selective activation of the CCR4 promoter in adult T-cell leukemia (ATL) cell lines. MOLT-4 and Jurkat: control human T-cell lines; HUT102 and ST1: ATL cell lines. (**b**) Deletion analysis. The promoter region from −151 to −96 bp is necessary and sufficient for reporter gene expression in the two ATL cell lines. (**c**) The schematic depiction of potential regulatory elements in the promoter region from −151 to −96 bp. (**d**) Mutation analysis. ΔC/EBP/Ikaros (from TCTTGGGAAA TGA to TCTTGCAAAATGA), ΔAP-1 (from AATGACTAAGA to AATGTCAAAGA), ΔPOUF3 (from CTTGGGAAATGA to CTTGGGAGGTGA), ΔPbx (from AAGAATCAT to AAGA CCCAT) and ΔGATA-3 (from TTCTATCAA to TTCTGACAA). The potential AP-1 and GATA-3 sites present within the −151 to −96 bp region are the major elements for CCR4 promoter activation in the two ATL cell lines.

expressed in primary ATL cells. We therefore first examined the mRNA expression of the AP-1 family members in primary ATL cells freshly isolated from patients in comparison with normal CD4+ T cells in resting, activated and Th1/Th2-polarized conditions (Figure 2a). As reported previously (Yoshie *et al.*, 2002; Nagakubo *et al.*, 2007), primary ATL cells consistently expressed CCR4 at levels much higher than various normal CD4+ T-cell populations, including Th2-polarized cultured T cells. Furthermore, primary ATL cells consistently expressed Fra-2 in sharp contrast to various normal CD4+ T-cell populations that were essentially negative for Fra-2 expression. Similar to various normal CD4+ T-cell populations, primary ATL

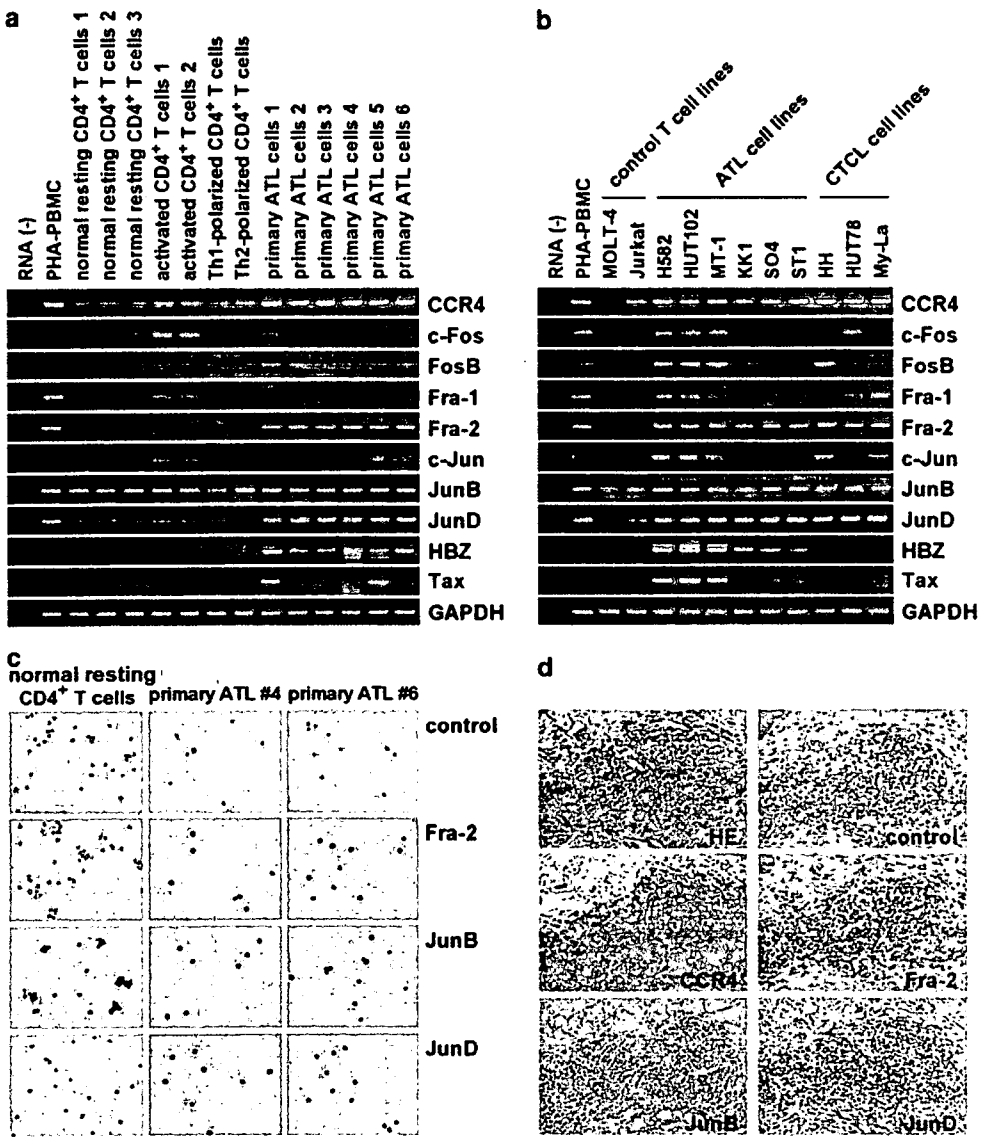


Figure 2 Constitutive expression of Fra-2, JunB and JunD in adult T-cell leukemia (ATL). (**a**) Reverse transcription (RT)–PCR analysis for the expression of the AP-1 family in normal T cells and primary ATL cells. Normal resting CD4+ T cells (purity, >96%) from healthy donors ($n=3$), activated CD4+ T cells from normal donors ($n=2$), Th1-polarized cultured CD4+ T cells, Th2-polarized cultured CD4+ T cells and freshly isolated primary ATL cells (>90% leukemic cells) from patients ($n=6$) were examined as indicated. Normal peripheral blood mononuclear cells treated with phytohemagglutinin (PHA-PBMC) served as a positive control. GAPDH served as a loading control. The representative results from at least two separate experiments are shown. (**b**) RT–PCR analysis for the expression of the AP-1 family in human T-cell lines. Two control human T-cell lines, six ATL cell lines and three CTCL cell lines were examined as indicated. PHA-PBMC served as a positive control. GAPDH served as a loading control. The representative results from two separate experiments are shown. (**c**) Immunocytochemical staining for Fra-2, JunB and JunD in normal CD4+ T cells and primary ATL cells. Normal CD4+ T cells from healthy donors (purity, >96%) and primary ATL cells (leukemic cells, >90%) from two patients were stained with anti-Fra-2, anti-JunB or anti-JunD. Normal rabbit IgG was used as the negative control (control). The representative results from two separate experiments are shown. Original magnification: ×400. (**d**) Immunohistochemical staining of CCR4, Fra-2, JunB and JunD in ATL skin lesions. Tissue sections from ATL skin lesions ($n=6$) were stained with anti-CCR4, anti-Fra-2, anti-JunB or anti-JunD. Mouse IgG_1 and normal rabbit IgG were used as the negative controls (control). Tissue sections were counterstained using Gill's hematoxylin. The representative results from a single donor are shown. Original magnification: ×400.

cells also constitutively expressed JunD and JunB even though JunD expression appeared to be upregulated in primary ATL cells. Other members of the AP-1 family were mostly negative in primary ATL cells, while activated normal CD4$^+$ T cells expressed c-Fos, Fra-1 and c-Jun at high levels. There was no correlation in expression between Fra-2 and the virally encoded HTLV-1 basic leucine zipper factor HBZ or Tax in primary ATL cells. We also confirmed that Fra-2 is not inducible by Tax using JPX-9, a subline of Jurkat carrying the HTLV-1 Tax gene under the control of the metallothionein gene promoter (Nagata *et al.*, 1989; data not shown). Thus, the constitutive expression of Fra-2 is highly unique for primary ATL cells.

We also examined expression of the same set of genes in various human T-cell lines. As shown in Figure 2b, compared to control T-cell lines, ATL cell lines consistently expressed CCR4 and Fra-2 at high levels. ATL cell lines also expressed JunB and JunD at high levels. HTLV-1 Tax has been shown to induce various AP-1 family members (Nagata *et al.*, 1989; Iwai *et al.*, 2001), which may be involved in HTLV-1 gene expression and cell proliferation (Jeang *et al.*, 1991). Consistently, ATL cell lines expressing Tax (H582, HUT102 and MT-1) also expressed other AP-1 family members at low levels. Cutaneous T-cell lymphomas (CTCLs) are a subset of HTLV-1-negative T-cell lymphomas resembling ATL and known to be frequently positive for CCR4 (Kim *et al.*, 2005). CTCL cell lines were also found to strongly express CCR4, Fra-2, JunB and JunD. Thus, the constitutive expressions of Fra-2, JunB and JunD were shared by CCR4-expressing ATL and CTCL cell lines.

We also examined the Fra-2, JunB and JunD protein expression in freshly isolated primary ATL cells and normal resting CD4$^+$ T cells. As shown in Figure 2c, primary ATL cells were indeed stained strongly positive for Fra-2, while normal CD4$^+$ T cells were totally negative for Fra-2. Primary ATL cells were also strongly positive for JunB and JunD, while normal CD4$^+$ T cells were variably positive for JunB and JunD at the single cell level. These results were highly consistent with the results from reverse transcription (RT)–PCR; Figure 2a). We also confirmed the CCR4, Fra-2, JunB and JunD protein expression in skin-infiltrating ATL cells (Figure 2d).

Activation of the CCR4 promoter by Fra-2/JunB and Fra-2/JunD heterodimers

AP-1 is known to function as a heterodimer of a member of the Fos family (c-Fos, FosB, Fra-1 and Fra-2) and a member of the Jun family (c-Jun, JunB and JunD) or a homodimer of the Jun family (Shaulian and Karin, 2002; Eferl and Wagner, 2003). We, therefore, next examined the activation of the CCR4 promoter by individual AP-1 family members singly or in combination. As recipients, we used two T-cell lines, namely, MOLT-4 and Jurkat. The expression levels of AP-1 members, including Fra-2, JunB and JunD, were very low in these cell lines (Figure 2b). As shown in Figure 3a, only Fra-2/JunB or Fra-2/JunD potently activated the CCR4 promoter in both cell lines. We confirmed that other members of the AP-1 family (c-Fos, FosB, Fra-1 and c-Jun) were transcriptionally active by using a synthetic promoter containing two tandem AP-1 consensus-binding sites (pGL3-2xAP-1; Figure 3b). Thus, among the AP-1 family members, only the Fra-2/JunB and Fra-2/JunD heterodimers are uniquely capable of activating the CCR4 promoter. This is highly consistent with their constitutive expression in primary ATL cells (Figure 2a).

Recently, the mRNA of HTLV-1 HBZ has been shown to be expressed in primary ATL cells (Satou *et al.*, 2006). We indeed observed the expression of HBZ in some primary ATL samples (Figure 2a). HBZ has been shown to activate JunB homodimer- or JunD homodimer-dependent transcription (Basbous *et al.*, 2003; Thebault *et al.*, 2004). Therefore, we also examined the effects of HBZ as well as Tax on the CCR4 promoter in MOLT-4 and Jurkat cells. As shown in Figure 3c, HBZ alone or in combination with Fra-2, JunB, JunD, Fra-2/JunB or Fra-2/JunD showed no effect on the activation of the CCR4 promoter. Similarly, Tax had no significant effect on the CCR4 promoter either alone or in combination with Fra-2, JunB, JunD, Fra-2/JunB or Fra-2/JunD. Thus, HTLV-1 encoded HBZ or Tax neither activates the CCR4 promoter nor affects its activation by Fra-2/JunB or Fra-2/JunD.

We have also confirmed that GATA-3 is constitutively expressed in primary ATL cells and activates the CCR4 promoter (data not shown). In normal CD4$^+$ T cells, GATA-3 may be responsible for the selective expression of CCR4 in Th2 cells (Imai *et al.*, 1999; Rengarajan *et al.*, 2000).

Specific binding of Fra-2, JunB and JunD to the AP-1 site in the CCR4 promoter

We next examined the specific binding of AP-1 family members to the AP-1 site in the CCR4 promoter using the NoShift transcription factor assay, an enzyme-linked immunosorbent assay (ELISA)-like colorimetric assay that is an alternative to the electrophoretic mobility shift assay. As shown in Figure 4a, when the nuclear extracts of two control T-cell lines (MOLT-4 and Jurkat) were used, the specific binding of any AP-1 family members to the AP-1 site of the CCR4 promoter was hardly observed. On the other hand, when the nuclear extracts of two ATL cell lines (HUT102 and ST1) were used, we detected a high level of specific binding of Fra-2, JunB and JunD to the AP-1 site. These results are highly consistent with the results from RT–PCR analyses (Figure 2b) and the luciferase reporter assays (Figure 3a).

By using the chromatin immunoprecipitation (ChIP) assay, we further examined the binding of Fra-2, JunB and JunD to the AP-1 site of the CCR4 promoter *in vivo*. As shown in Figure 4b, we detected specific binding of Fra-2, JunB and JunD to the AP-1 site of the endogenous CCR4 promoter in primary ATL cells but not in normal CD4$^+$ T cells. These results further

support the hypothesis that the CCR4 gene is a direct target gene of Fra-2/JunB and Fra-2/JunD heterodimers in primary ATL cells.

Effects of Fra-2, JunB and JunD small interfering RNAs on CCR4 expression and cell proliferation

To examine the role of Fra-2, JunB and JunD in CCR4 expression and cell proliferation in ATL cells, we next employed the small interfering RNA (siRNA) knockdown technique. As shown in Figure 5a, Fra-2 siRNA, JunB siRNA and JunD siRNA specifically reduced Fra-2 mRNA, JunB mRNA and JunD mRNA, respectively, in two ATL cell lines. On the other hand, control siRNA showed no such effect. Under these conditions, we examined the effects of these siRNAs on CCR4 expression and cell growth. As shown in Figure 5b, Fra-2 siRNA and JunD siRNA reduced CCR4 expression by approximately 50% in both cell lines, whereas JunB siRNA had hardly any inhibitory effect and control siRNA showed no inhibitory effect. Furthermore, as shown in Figure 5c, Fra-2 siRNA and JunD siRNA significantly reduced cell proliferation in both cell lines, whereas JunB siRNA or control siRNA did not. None of the siRNAs affected the growth of the control T-cell lines MOLT-4 and Jurkat. We also compared the effects of single and double knockdown of Fra-2 and JunD on cell growth in two ATL cell lines (Figure 5d). Compared to the effect of single knockdown of Fra-2 or JunD, no additive effect was observed by double knockdown of Fra-2 and JunD in both cell lines. These results may be consistent with the notion that Fra-2 and JunD promote growth in ATL cell lines by functioning as a heterodimer.

To further demonstrate the growth-promoting effects of Fra-2 and JunD, we performed stable transfection of Fra-2 and JunD in the control T-cell line Jurkat. As shown in Figure 5e, Jurkat cells overexpressing Fra-2 or JunD (see inset) indeed showed enhanced growth compared to those transfected with the vector alone. We were, however, unable to isolate Fra-2/JunD double transfectants in Jurkat, probably because of some adverse effects on Jurkat cells by the overexpression of both Fra-2 and JunD.

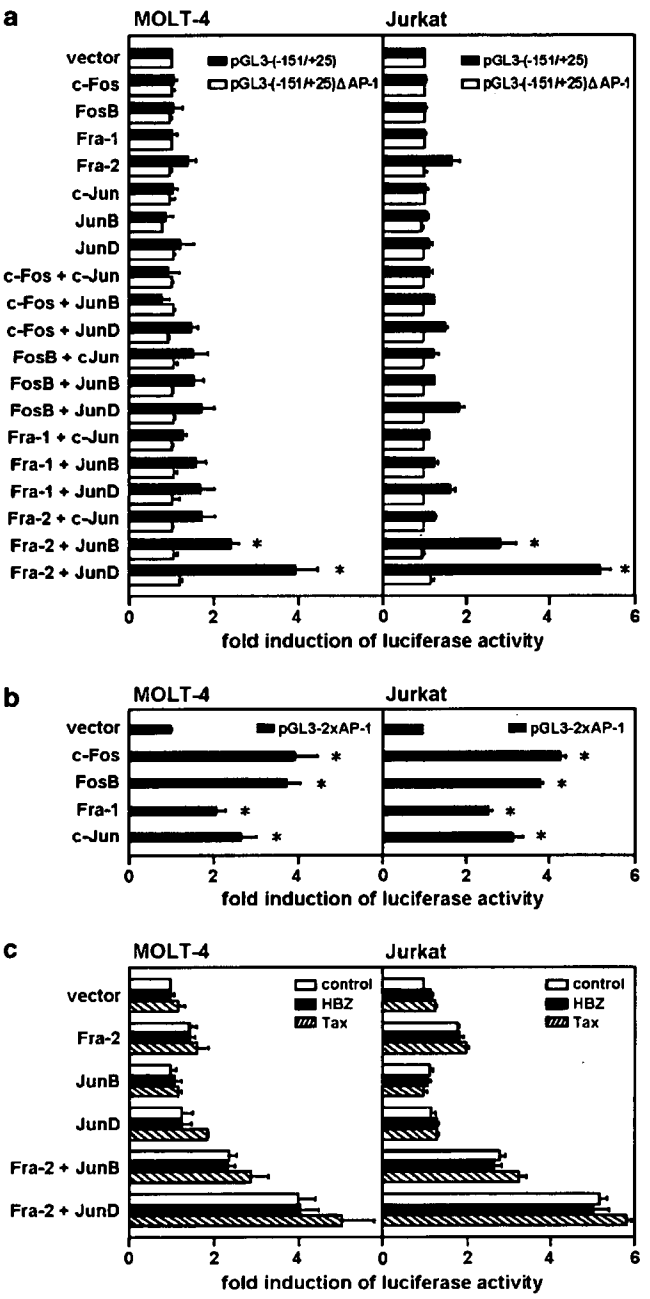


Figure 3 Transactivation of the CCR4 promoter by Fra-2/JunD and Fra-2/JunB. (**a**) Transactivation of the CCR4 promoter with or without the AP-1 site. MOLT-4 and Jurkat cells were cotransfected with pSV-β-galactosidase and pGL3-CCR4 (−151/+25) or pGL3-CCR4 (−151/+25)ΔAP-1 and an expression vector for c-Fos, FosB, Fra-1, Fra-2, c-Jun, JunB, JunD or a control vector as indicated. After 24–27 h, luciferase assays were performed in triplicate. Promoter activation was expressed as the fold induction of luciferase activity in cells transfected with an indicated AP-1 expression vector versus cells transfected with the vector alone. Transfection efficiency was normalized by β-galactosidase activity. Each bar represents the mean ± s.e.m. from three separate experiments. *$P<0.05$. (**b**) Transactivation of a synthetic promoter with two copies of the consensus AP-1 site. MOLT-4 and Jurkat cells were cotransfected with pSV-β-galactosidase and pGL3-2xAP-1 and an expression vector for c-Fos, FosB, Fra-1, c-Jun or the vector alone as indicated. Promoter activation was expressed as the fold induction of luciferase activity in cells transfected with an indicated expression vector versus cells transfected with a control vector. After 24–27 h, luciferase assays were performed in triplicate. Transfection efficiency was normalized by β-galactosidase activity. Each bar represents the mean ± s.e.m. from three separate experiments. *$P<0.05$. (**c**) Effect of HBZ or Tax on the activation of the CCR4 promoter. MOLT-4 and Jurkat cells were cotransfected with pSV-β-galactosidase and the pGL3-basic vector or pGL3-CCR4 (−151/+25) and an expression vector for Fra-2, JunB, JunD or a control vector and an expression vector for HBZ, Tax or a control vector as indicated. After 24–27 h, luciferase assays were performed in triplicate. Promoter activation was expressed as the fold induction of luciferase activity in cells transfected with an indicated expression vector versus cells transfected with a control vector. Transfection efficiency was normalized by β-galactosidase activity. Each bar represents the mean ± s.e.m. from three separate experiments.

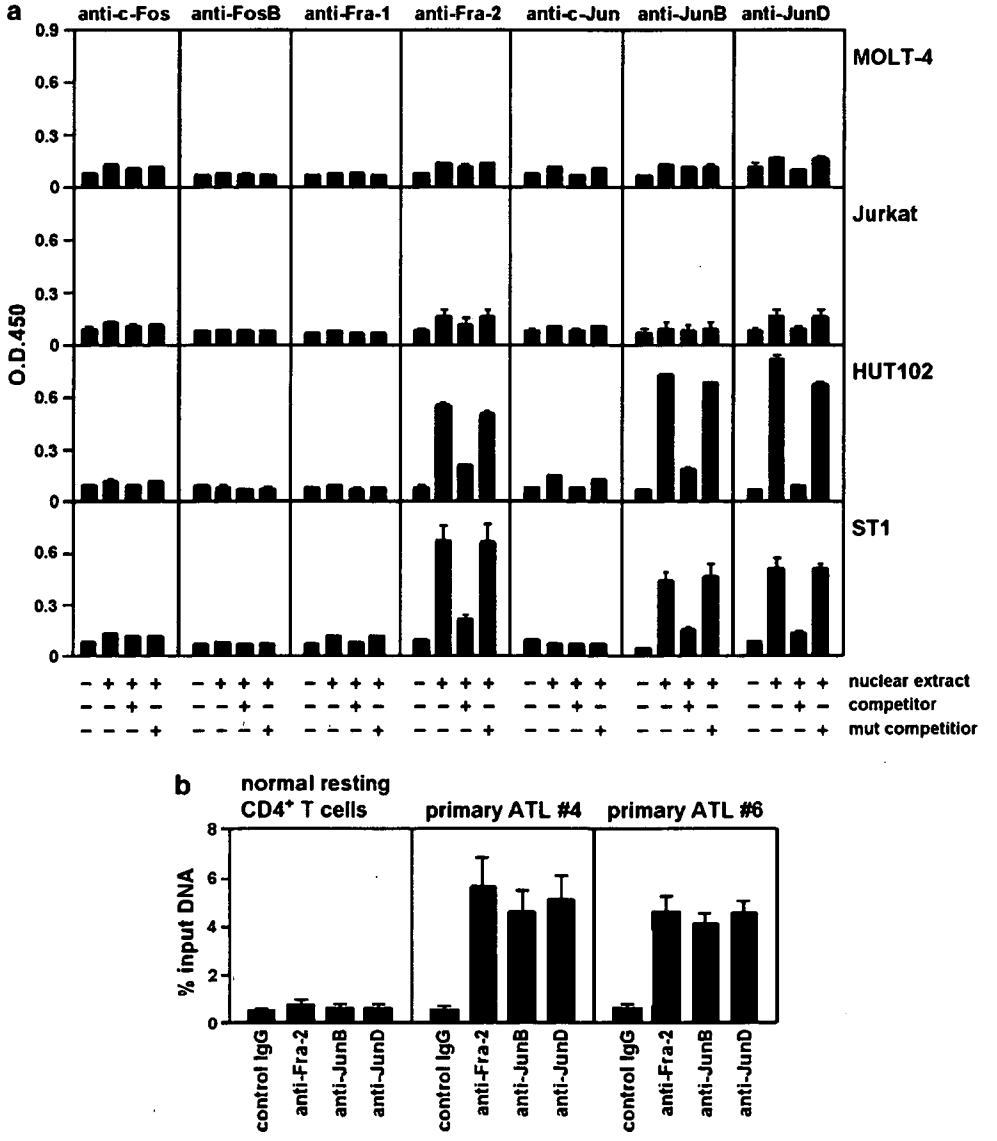


Figure 4 Specific binding of Fra-2, JunB and JunD to the AP-1 site in the CCR4 promoter. (**a**) NoShift assay. Nuclear extracts were prepared from two control T-cell lines (MOLT-4 and Jurkat) and two adult T-cell leukemia (ATL) cell lines (HUT102 and ST1). Nuclear proteins that bound to the biotinylated AP-1 site oligonucleotide (TGGGAAATGACTAAGAATCAT) were captured on an avidin-coated plate and detected by anti-c-Fos, anti-FosB, anti-Fra-1, anti-Fra-2, anti-c-Jun, anti-JunB or anti-JunD, as indicated. Specificity was determined by adding unlabeled probe (competitor; TGGGAAATGACTAAGAATCAT) or mutant probe (mut competitor; TGGGAAATGTCAAAGAATCAT; differences underlined). Each bar represents the mean ± s.e.m. from three separate experiments. (**b**) Chromatin immunoprecipitation (ChIP) assay. Chromatins from normal CD4$^+$ T cells from healthy donors (purity, >96%) and primary ATL cells from two patients (leukemic cells, >90%) were immunoprecipitated with anti-Fra-2, anti-JunD or control IgG. The amounts of precipitated DNA relative to total input DNA were quantified by real-time PCR for the CCR4 promoter region containing the AP-1 site. Each bar represents the mean ± s.e.m. from three separate experiments.

Identification of downstream target genes of the Fra-2/JunD heterodimer in ATL cells

To identify the target genes of Fra-2 in ATL cells, we compared the gene expression profiles of ATL-derived ST1 cells transfected with Fra-2 siRNA or control siRNA using the Affymetrix high-density oligonucleotide microarray. As summarized in Figure 6a, at least 49 genes were downregulated more than threefold by Fra-2 siRNA. The classification of these genes according to their biological functions shows that Fra-2 promotes the expression of genes involved in signal transduction (10 genes), protein biosynthesis and modification (8 genes) and transcription (6 genes); it also stimulates the expression of 10 genes of unknown function. Most notably, the list includes the proto-oncogenes c-Myb, BCL-6 and MDM2 (Oh and Reddy, 1999; Pasqualucci *et al*., 2003; Vargas *et al*., 2003). As shown in Figure 6b, RT–PCR analysis verified that not only Fra-2 siRNA but also JunD siRNA downregulated these proto-oncogenes in two ATL cell lines. Therefore, c-Myb, BCL-6 and MDM2 are the downstream target genes of the Fra-2/JunD heterodimer in both cell lines. This prompted us to examine the expression of c-Myb, BCL-6 and MDM2 in freshly isolated primary ATL cells by

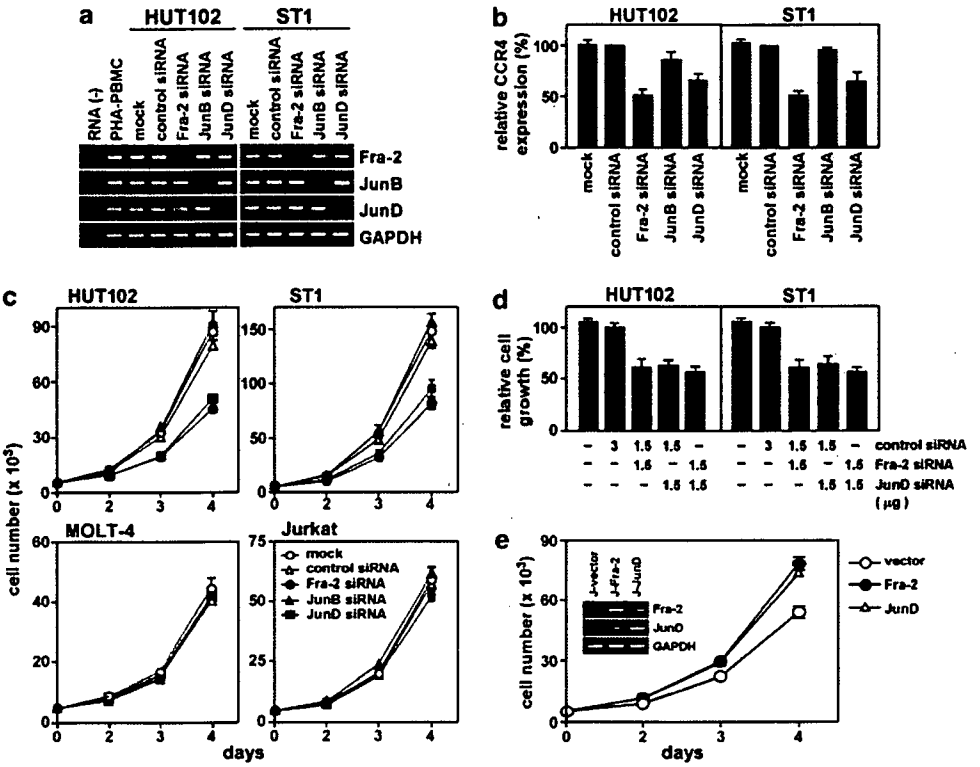


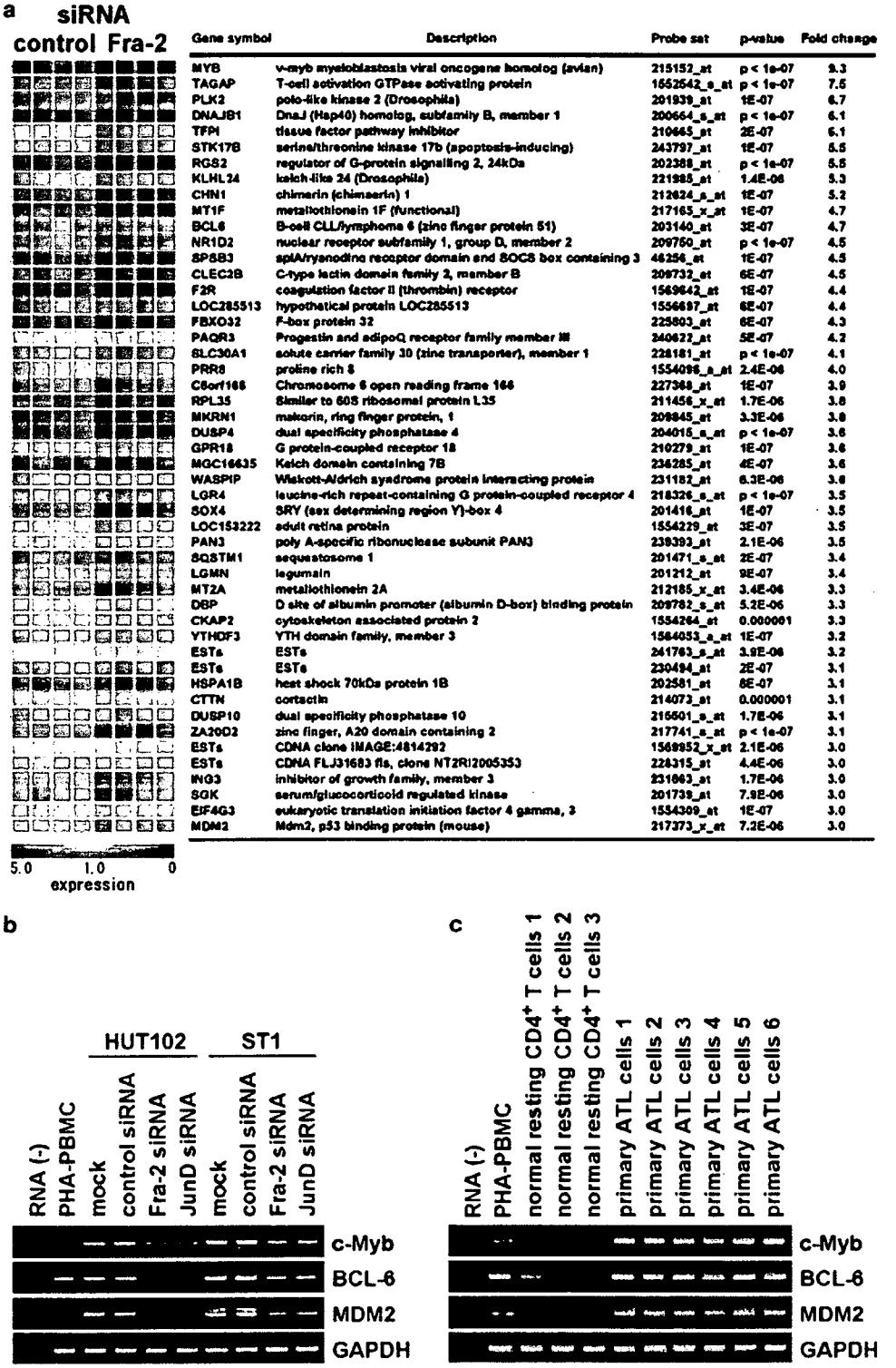
Figure 5 Dominant role of Fra-2/JunD in CCR4 expression and cell proliferation in adult T-cell leukemia (ATL). (**a**) Reverse transcription (RT)–PCR analysis to determine the effect of siRNAs. HUT102 and ST1 were transfected with control siRNA or siRNA for Fra-2, JunB or JunD. After 48 h, total RNA was prepared. The representative results from three separate experiments are shown. (**b**) Real-time RT–PCR analysis for CCR4 expression. HUT102 and ST1 were transfected with control siRNA or siRNA for Fra-2, JunB or JunD. After 48 h, total RNA was prepared and real-time RT–PCR was performed for CCR4 and 18S ribosomal RNA (an internal control). Data are presented as the mean ± s.e.m. of three separate experiments. (**c**) Effect of siRNAs on cell growth. HUT102, ST1, MOLT-4 and Jurkat were transfected with control, Fra-2, JunB and JunD siRNAs and cultured in a 96-well plate at 0.5×10^4 cells per well. At the indicated time points, viable cell numbers were determined using a FACSCalibur by gating out cells stained with propidium iodide. Data are shown as the mean ± s.e.m. of three separate experiments. (**d**) Effect of double knockdown of Fra-2 and JunD on cell growth. HUT102 and ST1 were transfected with control, Fra-2 and JunD siRNAs as indicated and cultured in a 96-well plate at 0.5×10^4 cells per well. At 4 days, viable cell numbers were determined on a FACSCalibur by gating out dead cells stained with propidium iodide. Data are shown as the mean ± s.e.m. of three separate experiments. (**e**) Effect of stable expression of Fra-2 and JunD on cell growth. Jurkat cells were transfected with a control IRES-EGFP expression vector or an IRES-EGFP expression vector for Fra-2 or JunD. Stable transfectants expressing green fluorescence protein were sorted and cultured in a 96-well plate at 0.5×10^4 cells per well. At the indicated time points, viable cell numbers were determined on a FACSCalibur by gating out dead cells stained with propidium iodide. Data are shown as the mean ± s.e.m. of three separate experiments.

RT–PCR. As shown in Figure 6c, we indeed detected the constitutive expression of c-Myb, BCL-6 and MDM2 at high levels in primary ATL cells. In sharp contrast, normal resting $CD4^+$ T cells hardly expressed these proto-oncogenes.

Discussion

The AP-1 transcription factors function as homodimers or heterodimers formed by Jun (c-Jun, JunB and JunD), Fos (c-Fos, FosB, Fra-1 and Fra-2) and the ATF family proteins (Shaulian and Karin, 2002; Eferl and Wagner, 2003). Most of them are rapidly and transiently induced by extracellular stimuli that trigger the activation of the Janus kinase (JNK), extracellular signal regulated protein kinases 1 and 2 (ERK1/2) or p38 mitogen-activated protein (MAP) kinase pathways (Shaulian and Karin, 2002; Eferl and Wagner, 2003). The AP-1 family is known to be involved in cellular proliferation, oncogenesis and even tumor suppression, depending on the combination of AP-1 proteins and the cellular context (Shaulian and Karin, 2002; Eferl and Wagner, 2003). Previously, by using the AP-1 site of the IL-8 promoter, Mori *et al.* demonstrated a strong Tax-independent expression of JunD in primary ATL cells (Mori *et al.*, 2000). In the present study, we have shown that Fra-2 is constitutively expressed at high levels in primary ATL cells (Figure 2a). Furthermore, except for JunB and JunD, other members of the AP-1 family are mostly negative in primary ATL cells (Figure 2a). Therefore, as demonstrated in the present study, the Fra-2/JunD and Fra-2/JunB heterodimers may be the major AP-1 factors constitutively active in primary ATL cells.

It has been shown that HTLV-1 Tax induces the expression of various AP-1 family members such as c-Fos, Fra-1, c-Jun and JunD (Nagata *et al.*, 1989; Iwai *et al.*, 2001). We indeed observed the expression of



Gene symbol	Description	Probe set	p-value	Fold change
MYB	v-myb myeloblastosis viral oncogene homolog (avian)	215152_at	p < 1e-07	9.3
TAGAP	T-cell activation GTPase activating protein	1552542_a_at	p < 1e-07	7.5
PLK2	polo-like kinase 2 (Drosophila)	201939_at	1E-07	6.7
DNAJB1	DnaJ (Hsp40) homolog, subfamily B, member 1	200664_s_at	p < 1e-07	6.1
TFPI	tissue factor pathway inhibitor	210665_at	2E-07	6.1
STK17B	serine/threonine kinase 17b (apoptosis-inducing)	243797_at	1E-07	5.5
RGS2	regulator of G-protein signalling 2, 24kDa	202388_at	p < 1e-07	5.5
KLHL24	kelch-like 24 (Drosophila)	221985_at	1.4E-06	5.3
CHN1	chimerin (chimaerin) 1	212624_s_at	1E-07	5.2
MT1F	metallothionein 1F (functional)	217165_x_at	1E-07	4.7
BCL6	B-cell CLL/lymphoma 6 (zinc finger protein 51)	203140_at	3E-07	4.7
NR1D2	nuclear receptor subfamily 1, group D, member 2	209750_at	p < 1e-07	4.5
SPSB3	splA/ryanodine receptor domain and SOCS box containing 3	46256_at	1E-07	4.5
CLEC2B	C-type lectin domain family 2, member B	209732_at	6E-07	4.5
F2R	coagulation factor II (thrombin) receptor	1569642_at	1E-07	4.4
LOC285513	hypothetical protein LOC285513	1556697_at	6E-07	4.4
FBXO32	F-box protein 32	225803_at	6E-07	4.3
PAQR3	Progestin and adipoQ receptor family member III	240627_at	5E-07	4.2
SLC30A1	solute carrier family 30 (zinc transporter), member 1	228181_at	p < 1e-07	4.1
PRR8	proline rich 8	1554098_a_at	2.4E-06	4.0
C8orf166	Chromosome 8 open reading frame 166	227368_at	1E-07	3.9
RPL35	Similar to 60S ribosomal protein L35	211456_x_at	1.7E-06	3.8
MKRN1	makorin, ring finger protein, 1	208845_at	3.3E-06	3.8
DUSP4	dual specificity phosphatase 4	204015_s_at	p < 1e-07	3.6
GPR18	G protein-coupled receptor 18	210279_at	1E-07	3.6
MGC16635	Kelch domain containing 7B	236285_at	4E-07	3.6
WASPIP	Wiskott-Aldrich syndrome protein interacting protein	231182_at	6.3E-06	3.6
LGR4	leucine-rich repeat-containing G protein-coupled receptor 4	218326_s_at	p < 1e-07	3.5
SOX4	SRY (sex determining region Y)-box 4	201416_at	1E-07	3.5
LOC153222	adult retina protein	1554229_at	3E-07	3.5
PAN3	poly A-specific ribonuclease subunit PAN3	238393_at	2.1E-06	3.5
SQSTM1	sequestosome 1	201471_s_at	2E-07	3.4
LGMN	legumain	201212_at	9E-07	3.4
MT2A	metallothionein 2A	212185_x_at	3.4E-06	3.3
DBP	D site of albumin promoter (albumin D-box) binding protein	209782_s_at	5.2E-06	3.3
CKAP2	cytoskeleton associated protein 2	1554264_at	0.000001	3.3
YTHDF3	YTH domain family, member 3	1564053_a_at	1E-07	3.2
ESTs	ESTs	241763_s_at	3.9E-06	3.2
ESTs	ESTs	230494_at	2E-07	3.1
HSPA1B	heat shock 70kDa protein 1B	202581_at	8E-07	3.1
CTTN	cortactin	214073_at	0.000001	3.1
DUSP10	dual specificity phosphatase 10	215501_s_at	1.7E-06	3.1
ZA20D2	zinc finger, A20 domain containing 2	217741_s_at	p < 1e-07	3.1
ESTs	CDNA clone IMAGE:4814292	1569952_x_at	2.1E-06	3.0
ESTs	CDNA FLJ31683 fis, clone NT2RI2005353	228315_at	4.4E-06	3.0
ING3	inhibitor of growth family, member 3	231663_at	1.7E-06	3.0
SGK	serum/glucocorticoid regulated kinase	201739_at	7.9E-06	3.0
EIF4G3	eukaryotic translation initiation factor 4 gamma, 3	1554309_at	1E-07	3.0
MDM2	Mdm2, p53 binding protein (mouse)	217373_x_at	7.2E-06	3.0

Figure 6 Identification of downstream target genes of Fra-2 in adult T-cell leukemia (ATL). (**a**) Microarray analysis. ST1 cells were transfected with control siRNA or Fra-2 siRNA. After 48 h, microarray analysis was performed using the Affymetrix GeneChip HG-U133 Plus 2.0 array. Four independent transfection samples were analysed for each group. Each column represents the expression level of a given gene in an individual sample. Red represents increased expression and blue represents decreased expression relative to the normalized expression of the gene across all samples. We computed the statistical significance level for each gene between the Fra-2-knockdown group and the control group with a mean fold change of >3 by the *t*-test ($P<10^{-5}$). (**b**) Reverse transcription (RT)–PCR analysis. HUT102 and ST1 cells were transfected with control siRNA or siRNA for Fra-2 or JunD. After 48 h, the expression of c-Myb, BCL-6, MDM2 and GAPDH was determined by RT–PCR. The representative results from three separate experiments are shown. (**c**) RT–PCR analysis. Normal $CD4^+$ T cells from healthy donors ($n=3$; purity, $>96\%$) and PBMC from ATL patients ($n=6$; leukemic cells, $>90\%$) were examined for the expression of c-Myb, BCL-6 and MDM2 by RT–PCR. The representative results from two separate experiments are shown.

various AP-1 family members in primary ATL cells (patient nos. 1 and 5) and in some ATL cell lines expressing Tax (Figures 2a and b). However, the constitutive expression of Fra-2, JunD and JunB in freshly isolated primary ATL cells and ATL cell lines is apparently independent from Tax expression (Figures 2a and b). This is further supported by the finding that CCR4-expressing HTLV-1-negative CTCL cell lines also constitutively express Fra-2, JunB and JunD at high levels (Figure 2). By using JPX-9, which is a subline of Jurkat carrying the HTLV-1 Tax gene under the control of the metallothionein gene promoter (Nagata *et al.*, 1989), we have also confirmed that Fra-2 is not inducible by Tax (data not shown).

The CCR4 promoter was potently activated by the Fra-2/JunB and Fra-2/JunD heterodimers (Figure 3a). Fra-2, JunB and JunD were also shown to bind specifically to the AP-1 site in the CCR4 promoter *in vitro* by the NoShift binding assays and *in vivo* by the ChIP assays (Figures 4a and b). By using the siRNA knockdown technique, however, only Fra-2 siRNA and JunD siRNA efficiently suppressed CCR4 expression and cell growth in ATL cell lines (Figure 5). On the other hand, JunB siRNA showed little such effect (Figure 5). Therefore, it is likely that, at least in terms of CCR4 expression and cell proliferation, the Fra-2/JunD heterodimer plays a more dominant role than the Fra-2/JunB heterodimer in ATL cells. It thus remains to be determined whether the Fra-2/JunB heterodimer has any specific functions in ATL.

The most striking finding in the present study is the aberrant expression of Fra-2 in primary ATL cells. Fra-2 expression is essentially absent in normal $CD4^+$ T cells under various conditions thus far examined (Figures 2a and c). Physiologically, Fra-2 is known to be expressed by various epithelial cells and in cartilaginous structures and has been shown to be required for efficient cartilage development (Karreth *et al.*, 2004). With regard to lymphoid cells, developing murine thymocytes were reported to express Fra-2 (Chen *et al.*, 1999). Previous studies have shown that individual homodimeric and heterodimeric AP-1 proteins have different functional properties and target genes (Shaulian and Karin, 2002; Eferl and Wagner, 2003). However, little is known about the target genes of Fra-2 and even less is known about the oncogenic role of Fra-2 in human malignancies. In this study, we have shown that CCR4 is the direct target gene of Fra-2 in association with JunD in ATL cells. Furthermore, we have shown that at least 49 genes are downregulated more than threefold in the ATL cell line ST-1 by Fra-2 siRNA (Figure 6). Among these genes, the proto-oncogenes c-Myb, BCL-6 and MDM2 (Oh and Reddy, 1999; Pasqualucci *et al.*, 2003; Vargas *et al.*, 2003) are further confirmed to be dependent on the Fra-2/JunD heterodimer and to be expressed at high levels in primary ATL cells (Figure 6). It remains to be seen whether the Fra-2/JunD heterodimer directly induces these proto-oncogenes or indirectly maintains their expression by promoting cell growth.

c-Myb is the genomic homologue of the avian myeloblastosis virus oncogene v-Myb. c-Myb is widely expressed in immature hematopoietic cells and also in various leukemias and carcinomas (Oh and Reddy, 1999; Shetzline *et al.*, 2004; Hess *et al.*, 2006). The target genes of c-Myb include the anti-apoptotic genes BCL-2 and BCL-X_L and also c-Myc (Ramsay *et al.*, 2003). Thus, c-Myb may promote the survival of ATL cells via BCL-2 and BCL-X_L (Galonek and Hardwick, 2006) and also cell cycle progression via c-Myc (Dang, 1999). BCL-6 was originally identified as the target gene of recurrent chromosomal translocations affecting 3q27 in non-Hodgkin's lymphoma. The expression of BCL-6 is frequently upregulated in diffuse large-cell lymphoma and follicular lymphoma through promoter substitution or somatic promoter point mutations (Ye *et al.*, 1993; Migliazza *et al.*, 1995; Chang *et al.*, 1996). Frequent expression of BCL-6 has also been reported in some T-cell lymphomas (Kerl *et al.*, 2001). The BCL-6 protein has been shown to exert cell-immortalizing and anti-senescence activities (Shvarts *et al.*, 2002; Pasqualucci *et al.*, 2003). Thus, BCL-6 may also inhibit apoptosis and promote cell cycle progression in ATL. The MDM2 protein is a negative regulator of p53 and suppresses p53-mediated cell cycle arrest and apoptosis (Vargas *et al.*, 2003). Elevated expression of MDM2 has been demonstrated in various types of human cancer (Rayburn *et al.*, 2005). Given that only a minor fraction of ATL cases have mutations affecting p53 (Cesarman *et al.*, 1992), the elevated expression of MDM2 may contribute to the functional downregulation of p53 in the majority of ATL cases.

CTCLs are a group of T-cell lymphomas derived from skin-homing memory T cells. CTCLs are not associated with HTLV-1 infection but resemble ATL and frequently express CCR4 (Ferenczi *et al.*, 2002; Kim *et al.*, 2005). Furthermore, CCR4 expression has been shown to be a consistent feature of the large-cell transformation of mycosis fungoides (Jones *et al.*, 2000). In the present study, we have shown that CTCL cell lines also express Fra-2, JunB and JunD at high levels (Figure 2b). Therefore, it is likely that aberrantly expressed Fra-2 in association with Jun proteins, particularly JunD, is also involved in CCR4-expression and cell proliferation in CTCLs.

In conclusion, we have shown that aberrantly expressed Fra-2 in association with JunD is responsible for CCR4 expression in ATL and is also likely to play an important role in ATL oncogenesis in part by inducing the expression of the proto-oncogenes c-Myb, BCL-6 and MDM2. Future studies are necessary to elucidate how the Fra-2/JunD heterodimer induces the expression of these proto-oncogenes and their individual roles in ATL oncogenesis. It also remains to be seen how ATL cells aberrantly express Fra-2 at high levels. Furthermore, the expression and function of Fra-2 in CTCLs remain to be determined.

Materials and methods

Cells

All the human T-cell lines used were described previously (Nagata *et al.*, 1989; Yamada *et al.*, 1996; Hata *et al.*, 1999;

Yoshie *et al.*, 2002). Peripheral blood mononuclear cells (PBMC) were isolated from heparinized blood samples obtained from healthy adult donors and acute ATL patients with a high leukemic cell count (>90%) by using Ficoll-Paque (Amersham Biosciences Corp, Piscataway, NJ, USA). Normal $CD4^+$ T cells (purity, >96%) were further prepared from PBMC by negative selection using an IMagnet system (BD Pharmingen, San Diego, CA, USA). Activated $CD4^+$ T cells were prepared by stimulating $CD4^+$ T cells with anti-CD3 (clone HIT3a; BD Pharmingen) and anti-CD28 (clone CD28.2; BD Pharmingen) for 24 h. The preparation of naive $CD4^+CD45RA^+$ T cells and their polarization into Th1 and Th2 cells were performed as described previously (Imai *et al.*, 1999). Primary ATL cells and normal resting $CD4^+$ T cells were used without culture for the experiments. This study was approved by the local ethical committee and written informed consent was obtained from each patient.

Transfection and luciferase assay

The major transcriptional start site (+1) of the human CCR4 gene was determined by the method of rapid amplification of cDNA 5′-ends and was found to be located 1797 bases upstream from the translation start codon (data not shown). To generate a promoter–reporter construct, the 1-kb promoter region of the human CCR4 gene (−983 to +25) was amplified from the genomic DNA by PCR using primers based on a GenBank genomic DNA sequence (accession no. NC_000003) and inserted into the reporter plasmid pGL3-Basic (Promega, Madison, WI, USA). Deletions and site-directed mutations were also performed using PCR. pGL3-2xAP-1 was constructed by introducing a sequence containing two copies of the AP-1 consensus binding site (TGATGACTCAGCCG-GAATGATGACTCAGCC) in front of a minimal CCR4 promoter pGL3 (−96/+25; Figure 1b). The coding regions of human FosB and GATA-3 were amplified from a cDNA library generated from phytohemagglutinin (PHA)-stimulated PBMC by PCR and cloned into the expression vector pSG5 (Stratagene, La Jolla, CA, USA). The coding region of HTLV-1 HBZ was amplified from a cDNA library generated from the HTLV-1^+ T-cell line C8166 by PCR and cloned into the expression vector pEF4/myc-His A (Invitrogen, Carlsbad, CA, USA). The expression vectors for c-Fos, Fra-1, Fra-2, c-Jun, JunB, JunD and Tax were described previously (Iwai *et al.*, 2001). Cells (5×10^5) were transfected with 2 μg of reporter plasmid, 0.5 μg of expression plasmids for various transcription factors and 1 μg of pSV-β-galactosidase using DMRIE-C (Invitrogen). After 24–27 h, luciferase assays were performed using a Luciferase Assay kit (Promega). Luciferase activity was normalized by β-galactosidase activity that served as an internal control for transfection efficiency.

RT–PCR

RT–PCR was carried out as described previously (Yoshie *et al.*, 2002). The primers used were as follows: +5′-AAGAA GAACAAGGCGGTGAAGATG-3′ and −5′-AGGCCCC TGCAGGTTTTGAAG-3′ for CCR4; +5′-TACTACCACTC ACCCGCAGACTC-3′ and −5′-CTTTTCCCTTCGGATTCT CCTTTT-3′ for c-Fos; +5′-TAGCAGCAGCTAAATGC AGGAAC-3′ and −5′-CCAGCTGAAGCCATCTTCCTT AG-3′ for FosB; +5′-CAGTGGATGGTACAGCCTCA TTT-3′ and −5′-GCCCAGATTTCTCATCTTCCAGT-3′ for Fra-1; +5′-CCAGCAGAAATTCCGGGTAGATA-3′ and −5′-TCTCCTCCTCTTCAGGAGACAGC-3′ for Fra-2; +5′-AAACAGAGCATGACCCTGAACCT-3′ and −5′-CTC CTGCTCATCTGTCACGTTCT-3′ for c-Jun; +5′-AAAAT GGAACAGCCCTTCTACCA-3′ and −5′-AGCCCTGACCA GAAAAGTAGCTG-3′ for JunB; +5′-AACACCCTTCT ACGGCGATGAG-3′ and −5′-GGGTAGAGGAACTGTG AGCTCGT-3′ for JunD; +5′-GAATTGGTGGACGGG CTATTATC-3′ and −5′-TAGCACTATGCTGTTTCGCCT TC-3′ for HBZ; +5′-CCGGCGCTGCTCTCATCCCGGT-3′ and −5′-GGCCGAACATAGTCCCCCAGAG-3′ for Tax; +5′-AAGGCATCCAGACCAGAAACCG-3′ and −5′-AGC ATCGAGCAGGGCTCTAACC-3′ for GATA-3; +5′-CAGT GACGAGGATGATGAGGACT-3′ and −5′-AACGTTTCG GACCGTATTTCTGT-3′ for c-Myb; +5′-ATTCCAGCTT CGGAACAAGAGAC-3′ and −5′-GTCCTTTTGATCAC TCCCACCTT-3′ for MDM2; +5′-CAAGAAGTTTCTAGG AAAGGCCGG-3′ and −5′-GATTGATCACACTAA GGTTGCATT-3′ for BCL-6 and +5′-GCCAAGGTCATCC ATGACAACTTTGG-3′ and −5′-GCCTGCTTCACCA CCTTCTTGATGTC-3′ for glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The amplification conditions were denaturation at 94 °C for 30 s (5 min for the first cycle), annealing at 60 °C for 30 s and extension at 72 °C for 30 s (5 min for the last cycle) for 34 cycles for CCR4; 35 cycles for c-Fos, FosB, Fra-1, Fra-2, c-Jun, JunB, JunD, HBZ, Tax, c-Myb, BCL-6 and MDM2; 29 cycles for GATA-3 and 27 cycles for GAPDH. Amplification products were electrophoretically run on a 2% agarose gel and stained with ethidium bromide.

Quantitative real-time PCR was carried out using the TaqMan assay and a 7700 Sequence Detection System (Applied Biosystems, Foster City, CA, USA). The conditions for PCR were 50 °C for 2 min, 95 °C for 10 min and then 50 cycles of 95 °C for 15 s (denaturation) and 60 °C for 1 min (annealing extension). The primers and fluorogenic probes for CCR4 and 18S ribosomal RNA were obtained from a TaqMan kit (Applied Biosystems). Quantification of CCR4 expression was performed using the Sequence Detector System Software (Applied Biosystems).

NoShift transcription factor assay

Anti-c-Fos (sc-52), anti-FosB (sc-7203), anti-Fra-1 (sc-22794), anti-Fra-2 (sc-604), anti-c-Jun (sc-1694), anti-JunB (sc-73) and anti-JunD (sc-74) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Transcription factors bound to specific DNA sequences were identified using the NoShift Transcription Factor Assay Kit (EMD Biosciences, Madison, WI, USA). Nuclear extracts were prepared from human T-cell lines by using the NucBuster Protein Extraction Kit (EMD Biosciences). The oligonucleotides used were as follows (differences underlined): TGGGAAATGACTAAGAATCAT for the biotinylated probe and unlabeled competitor of the AP-1 site and TGGGAAATGTCAAAGAATCAT for the mutated AP-1 site.

ChIP assay

This assay was performed using a ChIP assay kit (Upstate Biotechnology, Lake Placid, NY, USA) following the manufacturer's instructions. In brief, cells (1×10^6) were cross-linked with 1% formaldehyde for 10 min at room temperature. The cell pellets were lysed with sodium dodecyl sulfate (SDS) lysis buffer and sonicated to shear DNA to a size range between 200 and 1000 bp. After centrifugation, the supernatant was diluted 10-fold in ChIP dilution buffer and incubated overnight at 4 °C with 4 μg of anti-Fra-2 (sc-604), anti-JunB (sc-73), anti-JunD (sc-74) or normal rabbit IgG (DAKO, Kyoto, Japan). Immunocomplexes were collected by adding protein A-agarose beads. The immune complexes were incubated at 65 °C for 4 h to reverse the protein/DNA cross-links. DNAs were then purified by phenol/chloroform extraction and used as templates for quantitative real-time PCR. The primers and

the fluorogenic probe for the AP-1 site of the CCR4 promoter were as follows: primers: +5′-GGTCTTGGGAAATGACT AAGAATCA-3′ and −5′-TCTCCCTCACCCAACTGTACT AAGT-3′; probe: 5′-TCTGCTTCCTACTTCTATCAAA AAACCCCACTTG-3′.

Immunological staining
Cells were spotted on a glass slide and fixed with 4% paraformaldehyde. Tissue sections were prepared from formalin-fixed and paraffin-embedded biopsy tissue samples and subjected to microwave irradiation for 5 min three times in Target Retrieval Solution (DAKO). Slides and tissue sections were incubated for 1 h at room temperature with anti-Fra-2 (sc-604), anti-JunB (sc-73), anti-JunD (sc-74) or mouse monoclonal anti-CCR4 (KM-2160; Kyowa Hakko, Tokyo, Japan). Normal rabbit IgG and control mouse IgG_1 (DAKO) were used as negative controls. After washing, the slides and tissue sections were incubated with biotin-labeled goat anti-rabbit IgG or biotin-labeled horse anti-mouse IgG followed by detection using the Vectastain ABC/HRP kit (Vector Laboratories, Burlingame, CA, USA). Finally, cells and sections were counterstained with Gill's hematoxylin (Polysciences, Warrington, PA, USA), dehydrated and mounted.

Transfection of siRNA
siRNAs for Fra-2 (SI00420455), JunB (SI03077445), JunD (SI00075985) and the negative control (1022064) were obtained from Qiagen (Hilden, Germany). Transfection experiments were performed using Amaxa Nucleofector (Amaxa, Cologne, Germany). Cells (1×10^6) were resuspended in 100 µl of Nucleofector solution (T solution for MOLT-4, HUT102 and ST1 and V solution for Jurkat) and transfected with 2.5 µg of siRNA using program O-17 for MOLT-4, HUT102 and ST1 and program S-18 for Jurkat. The transfection efficiency was ~95% as determined using fluorescent siRNA (Qiagen).

Cell proliferation assay
Cells were seeded in a 96-well plate at a density of 0.5×10^4 per well and cultured. The number of viable cells was determined every 24 h on a FACSCalibur system (Becton Dickinson, Mountain View, CA, USA) by gating out cells stained with propidium iodide. To prepare stable transfectants of Fra-2 and JunD, the coding regions of human Fra-2 and JunD were inserted into the pIRES2-EGFP vector (BD Biosciences, San Diego, CA, USA). Jurkat cells were transfected with the plasmids using DMRIE-C (Invitrogen). Stable transfectants expressing green fluorescence protein were sorted by flow cytometry using FACSVantage (Becton Dickinson).

Oligonucleotide microarray
Microarray analysis was performed as described previously (Igarashi *et al.*, 2007) using the Affymetrix GeneChip HG-U133 Plus 2.0 array (Affymetrix, Santa Clara, CA, USA). In brief, the ATL-derived cell line ST1 was transfected with control siRNA or Fra-2 siRNA. Four independent transfections were performed for each group. After 48 h, total RNA samples were prepared and confirmed to be of good quality with the Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). All microarray data have been submitted to the Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo; accession no. GSE6379). The analysis was performed using the BRB Array Tools software version 3.3.0 (http://linus.nci.nih.gov/BRB-ArrayTools.html) developed by Richard Simon and Amy Peng.

Acknowledgements

We thank Namie Sakiyama for her excellent technical assistance. We also thank Dr Rich Simon and Dr Amy Peng for providing the BRB ArrayTools software. This work was supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports and Technology, Japan; by Solution-Oriented Research for Science and Technology (SORST) from Japan Science and Technology Corporation and by High-Tech Research Center Project for Private Universities: matching fund subsidy from the Ministry of Education, Culture, Sports, Science and Technology of Japan, 2002–2009.

References

Basbous J, Arpin C, Gaudray G, Piechaczyk M, Devaux C, Mesnard JM. (2003). The HBZ factor of human T-cell leukemia virus type I dimerizes with transcription factors JunB and c-Jun and modulates their transcriptional activity. *J Biol Chem* **278**: 43620–43627.

Cesarman E, Chadburn A, Inghirami G, Gaidano G, Knowles DM. (1992). Structural and functional analysis of oncogenes and tumor suppressor genes in adult T-cell leukemia/lymphoma shows frequent p53 mutations. *Blood* **80**: 3205–3216.

Chang CC, Ye BH, Chaganti RS, Dalla-Favera R. (1996). BCL-6, a POZ/zinc-finger protein, is a sequence-specific transcriptional repressor. *Proc Natl Acad Sci USA* **93**: 6947–6952.

Chen F, Chen D, Rothenberg EV. (1999). Specific regulation of fos family transcription factors in thymocytes at two developmental checkpoints. *Int Immunol* **11**: 677–688.

Dang CV. (1999). c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol Cell Biol* **19**: 1–11.

Eferl R, Wagner EF. (2003). AP-1: a double-edged sword in tumorigenesis. *Nat Rev Cancer* **3**: 859–868.

Ferenczi K, Fuhlbrigge RC, Pinkus J, Pinkus GS, Kupper TS. (2002). Increased CCR4 expression in cutaneous T cell lymphoma. *J Invest Dermatol* **119**: 1405–1410.

Galonek HL, Hardwick JM. (2006). Upgrading the BCL-2 Network. *Nat Cell Biol* **8**: 1317–1319.

Grassmann R, Aboud M, Jeang KT. (2005). Molecular mechanisms of cellular transformation by HTLV-1 Tax. *Oncogene* **24**: 5976–5985.

Hata T, Fujimoto T, Tsushima H, Murata K, Tsukasaki K, Atogami S *et al.* (1999). Multi-clonal expansion of unique human T-lymphotropic virus type-I-infected T cells with high growth potential in response to interleukin-2 in prodromal phase of adult T cell leukemia. *Leukemia* **13**: 215–221.

Hess JL, Bittner CB, Zeisig DT, Bach C, Fuchs U, Borkhardt A *et al.* (2006). c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells. *Blood* **108**: 297–304.

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

Iellem A, Mariani M, Lang R, Recalde H, Panina-Bordignon P, Sinigaglia F *et al.* (2001). Unique chemotactic response profile and specific expression of chemokine receptors CCR4 and CCR8 by CD4(+)CD25(+) regulatory T cells. *J Exp Med* **194**: 847–853.

Igarashi T, Izumi H, Uchiumi T, Nishio K, Arao T, Tanabe M *et al.* (2007). Clock and ATF4 transcription system regulates drug resistance in human cancer cell lines. *Oncogene* **26**: 4749–4760.

Imai T, Nagira M, Takagi S, Kakizaki M, Nishimura M, Wang J *et al.* (1999). Selective recruitment of CCR4-bearing Th2 cells toward antigen-presenting cells by the CC chemokines thymus and

activation-regulated chemokine and macrophage-derived chemokine. *Int Immunol* **11**: 81–88.

Ishida T, Utsunomiya A, Iida S, Inagaki H, Takatsuka Y, Kusumoto S *et al.* (2003). Clinical significance of CCR4 expression in adult T-cell leukemia/lymphoma: its close association with skin involvement and unfavorable outcome. *Clin Cancer Res* **9**: 3625–3634.

Iwai K, Mori N, Oie M, Yamamoto N, Fujii M. (2001). Human T-cell leukemia virus type I tax protein activates transcription through AP-1 site by inducing DNA binding activity in T cells. *Virology* **279**: 38–46.

Jeang KT, Chiu R, Santos E, Kim SJ. (1991). Induction of the HTLV-I LTR by Jun occurs through the Tax-responsive 21-bp elements. *Virology* **181**: 218–227.

Jones D, O C, Kraus MD, Perez-Atayde AR, Shahsafaei A, Wu L *et al.* (2000). Expression pattern of T-cell-associated chemokine receptors and their chemokines correlates with specific subtypes of T-cell non-Hodgkin lymphoma. *Blood* **96**: 685–690.

Karreth F, Hoebertz A, Scheuch H, Eferl R, Wagner EF. (2004). The AP1 transcription factor Fra2 is required for efficient cartilage development. *Development* **131**: 5717–5725.

Karube K, Ohshima K, Tsuchiya T, Yamaguchi T, Kawano R, Suzumiya J *et al.* (2004). Expression of FoxP3, a key molecule in CD4CD25 regulatory T cells, in adult T-cell leukaemia/lymphoma cells. *Br J Haematol* **126**: 81–84.

Kerl K, Vonlanthen R, Nagy M, Bolzonello NJ, Gindre P, Hurwitz N *et al.* (2001). Alterations on the 5′ noncoding region of the BCL-6 gene are not correlated with BCL-6 protein expression in T cell non-Hodgkin lymphomas. *Lab Invest* **81**: 1693–1702.

Kim EJ, Hess S, Richardson SK, Newton S, Showe LC, Benoit BM *et al.* (2005). Immunopathogenesis and therapy of cutaneous T cell lymphoma. *J Clin Invest* **115**: 798–812.

Matsubara Y, Hori T, Morita R, Sakaguchi S, Uchiyama T. (2005). Phenotypic and functional relationship between adult T-cell leukemia cells and regulatory T cells. *Leukemia* **19**: 482–483.

Matsuoka M. (2003). Human T-cell leukemia virus type I and adult T-cell leukemia. *Oncogene* **22**: 5131–5140.

Migliazza A, Martinotti S, Chen W, Fusco C, Ye BH, Knowles DM *et al.* (1995). Frequent somatic hypermutation of the 5′ noncoding region of the BCL6 gene in B-cell lymphoma. *Proc Natl Acad Sci USA* **92**: 12520–12524.

Mori N, Fujii M, Ikeda S, Yamada Y, Tomonaga M, Ballard DW *et al.* (1999). Constitutive activation of NF-kappaB in primary adult T-cell leukemia cells. *Blood* **93**: 2360–2368.

Mori N, Fujii M, Iwai K, Ikeda S, Yamasaki Y, Hata T *et al.* (2000). Constitutive activation of transcription factor AP-1 in primary adult T-cell leukemia cells. *Blood* **95**: 3915–3921.

Nagakubo D, Jin Z, Hieshima K, Nakayama T, Shirakawa AK, Tanaka Y *et al.* (2007). Expression of CCR9 in HTLV-1$^+$ T cells and ATL cells expressing Tax. *Int J Cancer* **120**: 1591–1597.

Nagata K, Ohtani K, Nakamura M, Sugamura K. (1989). Activation of endogenous c-fos proto-oncogene expression by human T-cell leukemia virus type I-encoded p40tax protein in the human T-cell line, Jurkat. *J Virol* **63**: 3220–3226.

Oh IH, Reddy EP. (1999). The myb gene family in cell growth, differentiation and apoptosis. *Oncogene* **18**: 3017–3033.

Pasqualucci L, Bereschenko O, Niu H, Klein U, Basso K, Guglielmino R *et al.* (2003). Molecular pathogenesis of non-Hodgkin's lymphoma: the role of Bcl-6. *Leuk Lymphoma* **44**(Suppl 3): S5–S12.

Ramsay RG, Barton AL, Gonda TJ. (2003). Targeting c-Myb expression in human disease. *Expert Opin Ther Targets* **7**: 235–248.

Rayburn E, Zhang R, He J, Wang H. (2005). MDM2 and human malignancies: expression, clinical pathology, prognostic markers, and implications for chemotherapy. *Curr Cancer Drug Targets* **5**: 27–41.

Rengarajan J, Szabo SJ, Glimcher LH. (2000). Transcriptional regulation of Th1/Th2 polarization. *Immunol Today* **21**: 479–483.

Satou Y, Yasunaga J, Yoshida M, Matsuoka M. (2006). HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. *Proc Natl Acad Sci USA* **103**: 720–725.

Shaulian E, Karin M. (2002). AP-1 as a regulator of cell life and death. *Nat Cell Biol* **4**: E131–E136.

Shetzline SE, Rallapalli R, Dowd KJ, Zou S, Nakata Y, Swider CR *et al.* (2004). Neuromedin U: a Myb-regulated autocrine growth factor for human myeloid leukemias. *Blood* **104**: 1833–1840.

Shvarts A, Brummelkamp TR, Scheeren F, Koh E, Daley GQ, Spits H *et al.* (2002). A senescence rescue screen identifies BCL6 as an inhibitor of anti-proliferative p19(ARF)-p53 signaling. *Genes Dev* **16**: 681–686.

Thebault S, Basbous J, Hivin P, Devaux C, Mesnard JM. (2004). HBZ interacts with JunD and stimulates its transcriptional activity. *FEBS Lett* **562**: 165–170.

Vargas DA, Takahashi S, Ronai Z. (2003). Mdm2: a regulator of cell growth and death. *Adv Cancer Res* **89**: 1–34.

Yamada Y, Ohmoto Y, Hata T, Yamamura M, Murata K, Tsukasaki K *et al.* (1996). Features of the cytokines secreted by adult T cell leukemia (ATL) cells. *Leuk Lymphoma* **21**: 443–447.

Yamamoto N, Hinuma Y. (1985). Viral aetiology of adult T-cell leukaemia. *J Gen Virol* **66**: 1641–1660.

Ye BH, Lista F, Lo Coco F, Knowles DM, Offit K, Chaganti RS *et al.* (1993). Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. *Science* **262**: 747–750.

Yoshida M. (2001). Multiple viral strategies of HTLV-1 for dysregulation of cell growth control. *Annu Rev Immunol* **19**: 475–496.

Yoshie O, Fujisawa R, Nakayama T, Harasawa H, Tago H, Izawa D *et al.* (2002). Frequent expression of CCR4 in adult T-cell leukemia and human T-cell leukemia virus type 1-transformed T cells. *Blood* **99**: 1505–1511.

Yoshie O, Imai T, Nomiyama H. (2001). Chemokines in immunity. *Adv Immunol* **78**: 57–110.

Respirology (2007) **12**, 856–862 doi: 10.1111/j.1440-1843.2007.01158.x

ORIGINAL ARTICLE

Development and validation of diagnostic prediction model for solitary pulmonary nodules

KAN YONEMORI,[1] UKIHIDE TATEISHI,[1] HAJIME UNO,[2] YOKO YONEMORI,[3] KOJI TSUTA,[4] MASAHIRO TAKEUCHI,[2] YOSHIHIRO MATSUNO,[4] YASUHIRO FUJIWARA,[5] HISAO ASAMURA[6] AND MASAHIKO KUSUMOTO[1]

Divisions of [1]Diagnostic Radiology and [4]Clinical Laboratory, [5]Medical Oncology Division and [6]Thoracic Surgery Division, National Cancer Center Hospital, [2]Division of Biostatistics, School of Pharmaceutical Sciences, Kitasato University, Tokyo, and [3]Department of Basic Pathology, Graduate School of Medicine, Chiba University, Chiba, Japan

Development and validation of diagnostic prediction model for solitary pulmonary nodules
YONEMORI K, TATEISHI U, UNO H, YONEMORI Y, TSUTA K, TAKEUCHI M, MATSUNO Y, FUJIWARA Y, ASAMURA H, KUSUMOTO M. *Respirology* 2007; **12**: 856–862

Background and objective: The aim of this study was to develop a simple prediction model for the underlying diagnosis of solitary pulmonary nodules (SPN) based on clinical characteristics and thin-section CT findings.

Methods: Retrospective analysis was carried out on 452 patients with SPN (113 benign and 339 malignant) smaller than 30 mm, who underwent thin-section CT followed by surgical resection and histological diagnosis. The clinical characteristics were collected from medical records, and radiographic characteristics from thin-section CT findings. The prediction model was determined using multivariate logistic analysis. The prediction model was validated in 148 consecutive patients with undiagnosed SPN, and the diagnostic accuracy of the model was compared with that of an experienced chest radiologist.

Results: The prediction model comprised the level of serum CRP, the level of carcinoembryonic antigen, the presence or absence of calcification, spiculation and CT bronchus sign. The areas under the receiver-operating characteristic curve in training and validation sets were 0.966 and 0.840, respectively. The diagnostic accuracies of the prediction model and the experienced chest radiologist for the validation set were 0.858 and 0.905, respectively.

Conclusion: The simple prediction model consisted of two biochemical and three radiographic characteristics. The diagnostic accuracy of an experienced chest radiologist was higher compared with the prediction model.

Key words: benign, CT, malignant, prediction model, solitary pulmonary nodule.

INTRODUCTION

The prevalence of a solitary pulmonary nodule (SPN), discovered on a CXR or CT scan, is reportedly 0.09 to 0.20% for all chest radiographs.[1,2] Most SPNs are benign, but malignancy accounts for approximately 20% of SPN, with a range of 3–80%.[3–5]

Although the diagnosis of an SPN may require resection, physicians should, as far as possible, minimize the risk of unnecessary surgery, especially in patients with benign diseases. Most physicians attempt to distinguish benign from malignant nodules using CXR, CT scanning, PET scans and fine-needle aspiration. Recent advances in radiographic

Correspondence: Kan Yonemori, Division of Diagnostic Radiology, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. Email: kyonemor@ncc.go.jp

Presented in part at the 100th American Thoracic Society International Conference, San Diego, USA, 20–25 May 2005.

Received 24 November 2006; invited to revise 20 December 2006 and 28 January 2007; revised 23 December 2006 and 15 February 2007; accepted 5 March 2007 (Associate Editor: Yoichi Nakanishi).

modalities and technology have improved the physicians' ability to assess the SPN, using information such as margin patterns, size, growth rate and location. Diagnostic and treatment decisions depend on the physician's assessment of the probability of malignancy of each SPN.

Improving the diagnostic accuracy of SPN in individual patients is a challenge. Previous methods for estimating the likelihood of each diagnosis used either Bayesian statistics or multivariate logistic regression analysis.[6–8] A Bayesian statistical model estimates the likelihood ratios by measuring the probability of finding a given feature in a population with malignant SPNs and dividing that by the probability of finding the same feature in a population with benign SPNs.[6,7] A certain scepticism regarding the Bayesian approach exists, and these models are not often applied in Western countries.[9] After the advent of CT scanning, Swensen *et al.* reported a multivariate logistic model that incorporated patient age, smoking history, a history of extra-thoracic malignancy, diameter, upper lobe location of the SPN, and the presence or absence of spiculation.[8]

Serum tumour markers are associated with malignancy.[10] No study has analysed the role of biochemical markers in differentiating between malignant and benign SPNs, and whether it may improve the performance of a prediction model. The aim of this study was to develop a simple prediction model for the diagnosis of SPN, based on clinical characteristics and thin-section CT findings, and to evaluate its diagnostic accuracy compared with that of an experienced chest radiologist.

METHODS

Records of 452 patients with SPNs who underwent surgery because of suspicion of primary lung cancer between July 1998 and March 2004 were retrieved (the training set: 113 benign SPNs, 339 malignant SPNs). All resected SPNs were evaluated histologically and a diagnosis made. Inclusion criteria were an SPN which was a solitary, round or oval lesion ≤30 mm in diameter and without associated adenopathy, atelectasis or effusion.[11] Any SPN diagnosed as metastatic extrapulmonary cancer was excluded. Benign SPN was defined as a benign pathological diagnosis. Malignant SPN was defined as that caused by primary lung cancer, such as non-small-cell or small-cell lung cancer. This study was approved by the institutional review board after confirmation of informed consent by the patient to have records and images reviewed.

The CT scans were acquired on single helical or multidetector scanners (X-Vigor, or Aquilion, Toshiba Medical Systems, Tokyo, Japan). The helical technique in all patients consisted of 10.0-mm collimation for individual scans of the entire lung (120 kVp, 150–200 mA) and reconstruction by a standard algorithm prior to surgery. In all patients, additional thin-section CT images were obtained with 2.0-mm collimation, a 20-cm field of view, 120 kVp and 200 mA per rotation, 0.5- or 1.0-s gantry rotation, and a high-spatial-frequency reconstruction algorithm. Hard-copy images were photographed at window settings for the lung (centre: –600 HU and width: 2000 HU) and mediastinum (centre: 50 HU and width: 550 HU).

Development of the prediction model

The clinical characteristics of patients comprising the training set ($n = 452$) were collected from the medical records: gender, age, smoking status (current or former smoker), history of extra-thoracic malignancies (>5 years previously), blood count and blood chemistry, including WCC, CRP, and serum tumour markers: carcinoembryonic antigen (CEA), sialyl-specific embryonic antigen, cytokeratin 19 fragment, squamous-cell carcinoma antigen, carbohydrate antigen 19–9, neurone-specific enolase and progastrin-releasing peptide. Two observers (K.Y., S.T.), blind to the clinical or histopathological diagnoses, reviewed the thin-section CT images, and the final decisions on diagnosis were reached by consensus. Thin-section CT findings of SPN in the training set were reviewed as follows: lung side (right vs left), location (lobe), diameter (mm), the presence or absence of calcification, cavity, lobulation, spiculation and CT bronchus sign. The CT bronchus sign was defined as a visible bronchus or bronchiolus leading to the SPN.[12]

Categorical variables of clinical characteristics and thin-section CT findings were compared using the chi-square test. Continuous variables were compared using the *t*-test. The prediction model was identified from multivariate logistic regression analysis. A stepwise procedure was used to select independent variables from the statistically significant variables in univariate analysis. Continuous variables in serum tumour markers without complete data sets were excluded from the logistic regression analysis. Some functional forms for continuous variables were also explored in the logistic regression analysis. The Hosemer-Lemeshow test was performed to evaluate goodness-of-fit of the prediction model. The performance of the prediction model was evaluated by calculating the prediction accuracy, the receiver-operating characteristic (ROC) curve and area under the curve. The statistical analysis was performed with SAS version 8.2 (SAS Institute, Cary, NC, USA), and the significance level was set at 0.05 (two-sided).

Validation of the prediction model

The prediction model was validated by comparing its accuracy with that of an experienced chest radiologist in the validation set. The validation set comprised a consecutive series of 148 patients with undiagnosed SPNs, who underwent thin-section CT imaging prior to surgery, between May 2004 and April 2005. The resected SPN was diagnosed pathologically as benign or malignant.

The reviews of thin-section CT images were performed independently before surgery. An observer (U.T.) reviewed the radiological predictive factors for

Table 1 Comparison of the clinical and radiological characteristics of the patients diagnosed with a benign or malignant solitary pulmonary nodule in the training set

Classification	Benign (n = 113)	Malignant (n = 339)	Total (N = 452)	P-value*
Clinical characteristics				
Male (%)	54	56	55	0.70
Mean age (years)	58	64	62	0.002
Current or past smoker (%)	47	54	49	0.86
Mean pack-years (n)	19	24	23	0.59
Other cancer >5 years ago (%)	11	10	10	0.79
Mean WCC ($\times 10^9$/L)	6.27	6.33	6.31	0.41
Mean serum CRP (mg/L)	1.0	3.0	2.0	<0.001
Mean serum CEA (ng/mL)	2.7	5.6	4.9	<0.001
Radiological characteristics				
Mean diameter (mm)	17	20	19	0.02
Side (right) (%)	65	63	64	0.65
Location (%)				
LUL	11	17	15	0.14
Lingular segment	5	5	5	0.80
LLL	18	14	15	0.37
RUL	30	31	31	0.91
RML	13	9	10	0.21
RLL	22	22	22	0.99
Calcification (%)	12	2	5	<0.001
Cavitation (%)	10	8	8	0.49
Lobulation (%)	16	12	13	0.26
Spiculation (%)	2	84	63	<0.001
CT bronchus sign (%)	32	82	69	<0.001

*Univariate analysis: chi-squared test and t-tests were performed for proportional differences or mean differences in variables between benign and malignant SPNs.

CEA, carcinoembryonic antigen; LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

the prediction models in the thin-section CT images. An experienced chest radiologist (M.K.) reviewed the thin-section CT images with no knowledge of either the clinical data or the present prediction model. This experienced chest radiologist evaluated prediction of an SPN on a 5-point scale, scored as follows: 1, benign; 2, probably benign; 3, indeterminate; 4, probably malignant; and 5, malignant. He then determined a diagnosis (malignant or benign SPNs). Based on the pathological diagnosis, all SPNs were divided into benign or malignant for the purpose of evaluation of the validation set.

RESULTS

The clinical and the radiological characteristics of the 452 patients in the training set (113 benign SPNs, 339 malignant SPNs) are presented in Table 1.

Sixty-three per cent of the benign SPNs were granulomas, and the remainder included sclerosing haemangioma and hamartoma. The majority of the malignant SPNs were adenocarcinoma of the lung (80.2%). Other malignancies included squamous-cell carcinoma (14.5%), large-cell carcinoma (3.5%), adenosquamous carcinoma (0.3%), adenoid cystic carcinoma (0.3%), neuroepidermoid carcinoma (0.3%) and small-cell lung carcinoma (0.9%).

The SPN prediction model defined the probability of benign SPN as follows: Probability of benign SPN = $e^x/(1 + e^x)$, x = 3.7009 + (3.0705 × calcification) + (−1.3243 × CT bronchus sign) + (−5.3399 × spiculation) + (−1.16 × $\sqrt{}$CEA) + (−1.4987 × $\sqrt{}$CRP), e is the base of natural logarithms. Calcification = 1 if calcification is present in the SPN (otherwise = 0); CT bronchus sign = 1 if CT bronchus sign is present (otherwise = 0); spiculation = 1 if spiculated appearance is present (otherwise = 0); CEA = serum CEA level (ng/mL); CRP = serum CRP level (mg/L). The area under ROC in the training set was 0.966 (Fig. 1). The goodness-of-fit statistic, as described by Hosmer-Lemeshow, for the derivation dataset was χ^2 = 11.608, P-value = 0.170. If the biological parameters were not available, the alternative prediction model was as follows: Probability of benign SPN = $e^x/(1 + e^x)$, x = 1.084 + (2.7851 × calcification) + (−1.1795 × CT bronchus sign) + (−5.4481 × spiculation). The area under ROC of the alternative prediction model in the training set was 0.94.

Following derivation of the prediction model, a further 148 patients with a newly discovered indeterminate SPN were identified between May 2004 and

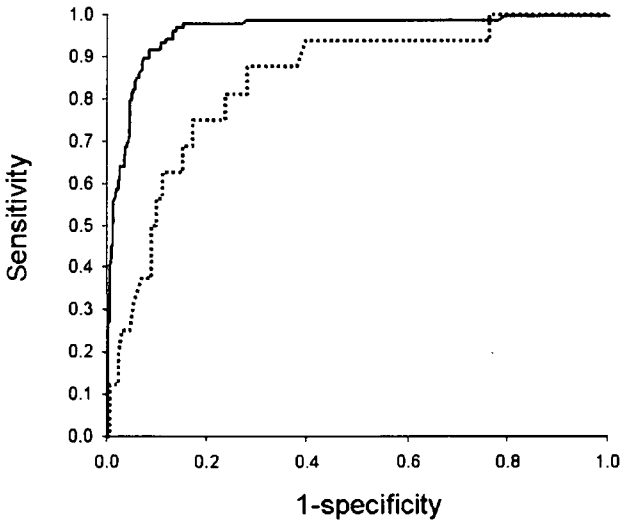


Figure 1 Receiver-operating characteristic curve for the prediction model for the diagnosis of a solitary pulmonary nodule (SPN) (solid line). The prediction model was developed based on data from 452 patients with a solitary pulmonary nodule. Receiver-operating characteristic curve for the prediction model when validated in 148 consecutive patients with SPNs (dotted line).

April 2005 (the validation set). The differences in the clinical and radiological characteristics of both sets are presented in Table 2.

Eighty-nine per cent of the validation set had malignant SPNs, and the reminder benign SPNs. Forty-seven per cent of the benign SPNs were granulomas; other diagnoses included sclerosing haemangioma, hamartoma, fibrosis, organizing pneumonia and inflammation. The majority of the malignant SPNs were adenocarcinoma of the lung (88%); other malignancies included squamous-cell carcinoma (7.4%), large-cell carcinoma (2.7%), pleomorphic carcinoma (0.7%), small-cell lung carcinoma (0.7%), and mucosa-associated lymphoid tissue (MALT) lymphoma of the lung (0.7%). The respective mean serum CEA and CRP levels in benign SPNs were 3.1 ng/mL (range 0.8–16.7) and 1.0 mg/L (range 1.0–3.0). The corresponding mean serum CEA and CRP levels in malignant SPNs were 5.1 ng/mL (range 0.5–58.4) and 3.0 mg/L (range 1.0–165). The mean diameters of the benign and malignant SPNs were 16 mm (range 10–23) and 17 mm (range 8–30), respectively.

In the validation set (n = 148), the area under ROC of the prediction model was 0.84 (Fig. 1), and the accuracies of the prediction model and the experienced chest radiologist were 0.858 and 0.905, respectively. ROC and accuracy of the alternative prediction model were 0.81 and 0.75. The experienced chest radiologist's prediction was inconsistent with 59% of benign SPNs, and 4% of malignant SPNs (3: adenocarcinoma; 1: pleomorphic carcinoma; 1: MALT of the lung).

DISCUSSION

This study has demonstrated a simple prediction model for the diagnosis of SPN based on two clinical variables and three radiological variables using thin-section CT findings. The accuracy of the prediction model was demonstrated to be high in a validation based on consecutive patients with undiagnosed SPNs, although not as high as the accuracy of the experienced chest radiologist's diagnostic prediction.

Statistical analysis indicated two biochemical markers for inclusion in the prediction model, but did not include the patient characteristics of age, cigarette-smoking status and past history of extra-pulmonary malignancy.[8] The frequency of a malignant SPN rises with increasing patient age. In the present study, age was significantly associated with a malignant SPN on univariate analysis. Several studies have reported the frequency of malignant SPN to be more than 50% higher in patients aged 50 years and over than in those aged less than 50 years.[13] Similarly, cigarette-smoking and past history of extra-pulmonary malignancy reportedly have strong associations with lung cancer or malignant SPNs. The training and validation sets included 9% and 6% of patients less than 50 years of age, respectively. Although the mean age and frequency of past history of extra-pulmonary malignancy in this study were similar to those in the previously reported prediction model, the frequency of cigarette smoking in this study was slightly lower than the 68% of patients with SPNs in the previous study.[9] According to data from Japan Tobacco Industry, Inc., the overall smoking prevalences in men and women were 57.5% and 14% in 1996 respectively, which are significantly different from the prevalence in the 1960s.[14] These patients' demographic characteristics varied from the frequencies in other SPN reports, possibly influenced by the role of the institution involved, the period or the patient population.

Serum markers based on tumour biology would not have been strongly influenced by patient population or institution. CEA is a marker for a wide range of malignancies, including lung cancer. Elevation of serum CEA was found in 16% to 27% of patients with clinical stage IA non-small-cell lung cancer.[15,16] Although the serum CEA level was associated with cigarette smoking and age,[17] the multivariate analysis used to develop the prediction model indicated the significance of serum CEA rather than these known confounding factors included in the previous prediction model.

CRP, an acute-phase protein synthesized in hepatocytes, is up-regulated by cytokines, such as IL-6 and tumour necrosis factor-α.[18,19] Increased serum CRP levels have also been recognized as part of a paraneoplastic syndrome for several malignant tumours, including lung cancer.[20] In addition, serum CRP was significantly higher in patients with lung cancer than in those with benign lung disease or the healthy population.[21–23] The laboratory variables in the prediction model were supported by these previous studies and our results; that is, the serum markers in patients with malignant SPNs were both higher than in those with benign SPNs. CT has been shown to be an effective means of morphologically differentiating benign from malignant SPNs. Numerous studies have described the radiographic characteristics of SPN,

Table 2 Comparison of the clinical and radiological characteristics in both training and validation patient sets

	Training set (n = 452)	Validation set (n = 148)	P-value*
Clinical characteristics			
Male (%)	56	57	0.69
Mean age (years)	62	63	0.82
Current or past smoker (%)	49	50	0.82
Other cancer >5 years ago (%)	10	11	0.60
Mean serum CEA (ng/mL)	4.9	4.9	0.99
Mean serum CRP (mg/L)	2.0	3.0	0.56
Radiological characteristics			
Mean diameter (mm)	19	17	<0.01
Side (right) (%)	64	61	0.63
Location (%)			
LUL	15	20	0.15
Lingular segment	5	2	0.13
LLL	15	16	0.73
RUL	31	32	0.74
RML	10	7	0.32
RLL	22	22	0.85
Calcification (%)	5	0	<0.01
Spiculation (%)	63	68	0.27
CT bronchus sign (%)	69	74	0.31

*Univariate analysis: chi-squared test and t-tests were performed for proportional differences or mean differences in variables between training and validation sets (two-sided).

CEA, carcinoembryonic antigen; LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

and each suggests, but does not guarantee, a diagnosis of benign versus malignant SPN. Spiculation, lobulation and vascular convergence are typically associated with malignancy.[23,24] Calcification, a well-defined margin, cavitation and CT bronchus signs including bronchus involvement show considerable overlap between benign and malignant SPNs.[24–28] The reported frequencies of these radiological findings have varied among studies. In the present study, the prediction model identified calcification, spiculation and CT bronchus sign as being useful for distinguishing benign from malignant SPNs. The validity of the present prediction model is supported by the previous prediction model and the typical radiographic features of malignant versus benign SPNs.

There were no differences in the laboratory or radiological results or in the pattern of care for SPN between the training set and the validation set. However, a limitation of the present study is the potential for selection bias, just as in previous studies. Selection bias has arisen because the model is based on a population with clinically suspicious lesions that required resection. It is unclear how this would affect the performance of the prediction model in a more general population. There were significant differences in two radiological variables (mean diameter and calcification) between the training set and the validation set. The prediction model would also need to be validated in other patient populations outside that of our single institution.

The best clinical management of a patient with an SPN requires evaluation of the probability of malignancy, which determines the most cost-effective diagnosis and treatment strategies. In several decision analyses, the probability of malignancy is divided into four categories: low (<10%), intermediate (10–60%), high (>60 to 90%) and very high (>90%).[29,30] The management pathway should take into consideration patient preference and the potential complications after assessment of the probability of malignancy. CT or CXR follow up is preferred in patients with a low probability of malignancy. When the probability is intermediate, additional testing would be required, including contrast-enhanced CT, fine-needle aspiration, bronchoscopy and PET/CT. Surgery is recommended in patients with a high or very high probability of malignancy. PET/CT has become important in differentiating benign from malignant nodules, and its estimated sensitivity for identifying a malignant process is 96%, and its specificity is 88%.[31] Currently, investigators have made efforts to develop a computer-aided diagnosis system for SPN. The computer-aided diagnosis system has the potential to improve diagnostic accuracy of SPN in future.[32]

Radiological examinations such as CT will detect a large number of SPNs. The physician can use the prediction model for differentiating between benign and malignant SPNs. Even if serum CEA and CRP cannot be measured, the alternative prediction model might assist in the evaluation of the probability of malignancy. The prediction model and particular observation of CT imaging for each patient would assist in determining the optimal management and in identifying the indications for invasive and