Table 2
Association analysis of adjusted BMD in Groups-A and -B

SNP	Genotype	Group-A (n = 387) (spinal BMD Z-score)		P value[a]	Group-B (n = 384) (radial-adjusted BMD)		P value[a]
		Mean ± SD	n		Mean ± SD	n	
Q89R	Q	−0.15 ± 1.50	332	0.85	0.401 ± 0.055	333	0.052
	QR	−0.14 ± 1.54	35		0.387 ± 0.049	48	
	R	−0.57 ± 1.10	3		0.360 ± 0.059	2	
A1330V	A	−0.04 ± 1.61	178	0.034*	0.405 ± 0.053	195	0.019*
	AV	−0.35 ± 1.38	174		0.395 ± 0.059	155	
	VV	−0.47 ± 1.47	35		0.385 ± 0.041	34	
MC17-1677C > A	CC	−0.01 ± 1.42	166	0.005**	0.405 ± 0.053	194	0.055
(before expansion)	CA	−0.35 ± 1.44	118		0.395 ± 0.059	152	
	A	−0.78 ± 1.36	24		0.390 ± 0.040	29	

±Data were from the previous study before subject expansion.
[a] P values were calculated for the linear regression.
* $P < 0.05$.
** $P < 0.01$.

and four had remarkably low BMD (Z-scores < −3.0). For distribution analysis, we selected 91 subjects with Z-scores > +1.0 and 139 with Z-scores < −1.0. Among those women, the Z-score was greater than 2.0 in 33 and smaller than −2.0 in 34 of them. Biochemical markers of bone turnover including serum intact osteocalcin, urinary pyridinoline, and deoxypyridinoline were measured in most of the 387 subjects ascertained from the Institute; each had given informed consent prior to the study.

DNA samples for a population-based analysis were obtained from peripheral blood of 384 adult Japanese women [20,21]. In this group (Group-B), mean ages and body mass indices (BMI) with standard deviations (SD) were 58.4 ± 8.6 (range 32–69) years and 23.7 ± 3.61 (range 14.7–38.5) kg/m^2, respectively. The BMD of radial bone (expressed in g/cm^2) of each participant was measured by dual energy X-ray absorptiometry (DXA) using DTX-200 (Osteometer Meditech Inc., Hawthorne, CA, USA) and was normalized for differences in age, height, and weight by multiple linear regression analysis [21]. Forearm BMD in the distal radius was measured according to the Guidelines for Osteoporosis Screening in a health check-up program in Japan [22]; the instruments (DTX-200) were calibrated at every measurement, and the coefficients of variation were kept within the 1.0 ± 0.5% (CV ± SD). No participant had medical complications or was undergoing treatment for conditions known to affect bone metabolism [21]. All were non-related volunteers, and written informed consent was obtained from each of them. The ethical committee of the Institutional Review Board approved the entire project.

Mutation search, SNP selection, and genotyping

Mutation analysis of all 23 exons and flanking regions was carried out in DNA from the 12 subjects with the highest (>3.0, n = 8) and lowest (<−3.0, n = 4) Z-scores by direct sequencing of PCR products in the ABI Prism 377 DNA system (Applied Biosystems).

As potentially functional SNPs, two previously reported missense coding-SNPs, A1330V (c.3989C > T) and Q89R (c.266A > G), were selected for testing association of BMD levels, but a third, V667M, was eliminated because it was undetectable among our subjects (Table 1). Another intronic variation previously reported to associate with low BMD also was tested for association. For LD analysis, we first tested more than 40 variations within the *LRP5* locus, including all 38 polymorphic variations archived in the JSNP-database (http://snp.ims.u-tokyo.ac.jp/index.html), dbSNP, and from the literature [11,23]. However, since some of the variations showed minor allele frequencies of <0.05 among our Group-B 384 subjects, we chose to use only 29 SNPs for LD analysis.

Genotypes for these 29 selected SNPs were determined either by the Sd-PCR method [20], Invader assay (Third Wave Technologies, Madison, WI) [24], or TaqMan Assay (Applied Biosystems). The Sd-PCR method was used for 16 SNPs, according to a protocol described previously [20]. In brief, the Sd-PCR reaction was carried out using two allele-specific primers (AS-primers) and one reverse primer in a standard reaction mixture containing fluorescence-labeled dCTP. Discrimination of alleles, on the basis of by five-base differences between the AS-primers, was achieved using an ABI Prism 377 DNA system [21]. The Invader assay was applied for 10 other SNPs, according to the manufacturer's protocol. In brief, 1 μl of the diluted PCR product (1:333 in distilled water) of the region flanking each SNP was used as template in a 6-μl reaction mixture in 384-well plates, and fluorescent signals for FAM and Redmond Red were detected by a plate-reader after standard 1-h incubation [25]. The other three SNPs were genotyped by TaqMan Assay according to the manufacturer's protocol.

Table 3
Physical and clinical characteristics of the subjects in Group-A (healthy subjects from a clinic)

Group-A subjects (n = 387)	A1330V			P value*
	AA	AV	VV	
Number	178	174	35	–
Age (Years)	65.6 ± 10.1 (41–86)	64.1 ± 10.4 (33–89)	59.3 ± 15.0 (25–87)	0.012
Weight (kg)	51.1 ± 7.9 (34–74)	50.0 ± 7.9 (33–76)	50.2 ± 8.1 (34–70)	NS
BMI (kg/m^2)	22.5 ± 3.0 (16.4–32.9)	21.9 ± 2.7 (14.3–31.8)	21.9 ± 3.2 (15.3–29.5)	NS
Height (cm)	150.8 ± 6.2 (135–167)	150.8 ± 6.7 (134–172)	151.1 ± 5.4 (140–162)	NS
Spine BMD (g/cm^2)	0.914 ± 0.215 (0.387–1.732)	0.878 ± 0.188 (0.400–1.506)	0.898 ± 0.196 (0.518–1.310)	NS
BMD Z-score	−0.03 ± 1.62 (−3.1–7.1)	−0.35 ± 1.38 (−3.5–5.3)	−0.45 ± 1.47 (−2.9–3.1)	0.034
Intact-OC (ng/ml)[†]	7.18 ± 3.16 (n = 117)	8.40 ± 3.82 (n = 118)	8.87 ± 3.81 (n = 18)	0.004
Pyridinoline (pmol/μmol crea.)	34.2 ± 11.1 (n = 138)	33.6 ± 11.2 (n = 133)	34.5 ± 11.4 (n = 20)	NS
Deoxypyridinoline (pmol/μmol crea.)	7.20 ± 2.50 (n = 139)	7.29 ± 2.42 (n = 133)	7.85 ± 3.01 (n = 20)	NS

* P values are calculated for regression analysis with ANOVA F-test.
[†] Serum intact osteocalcin level.

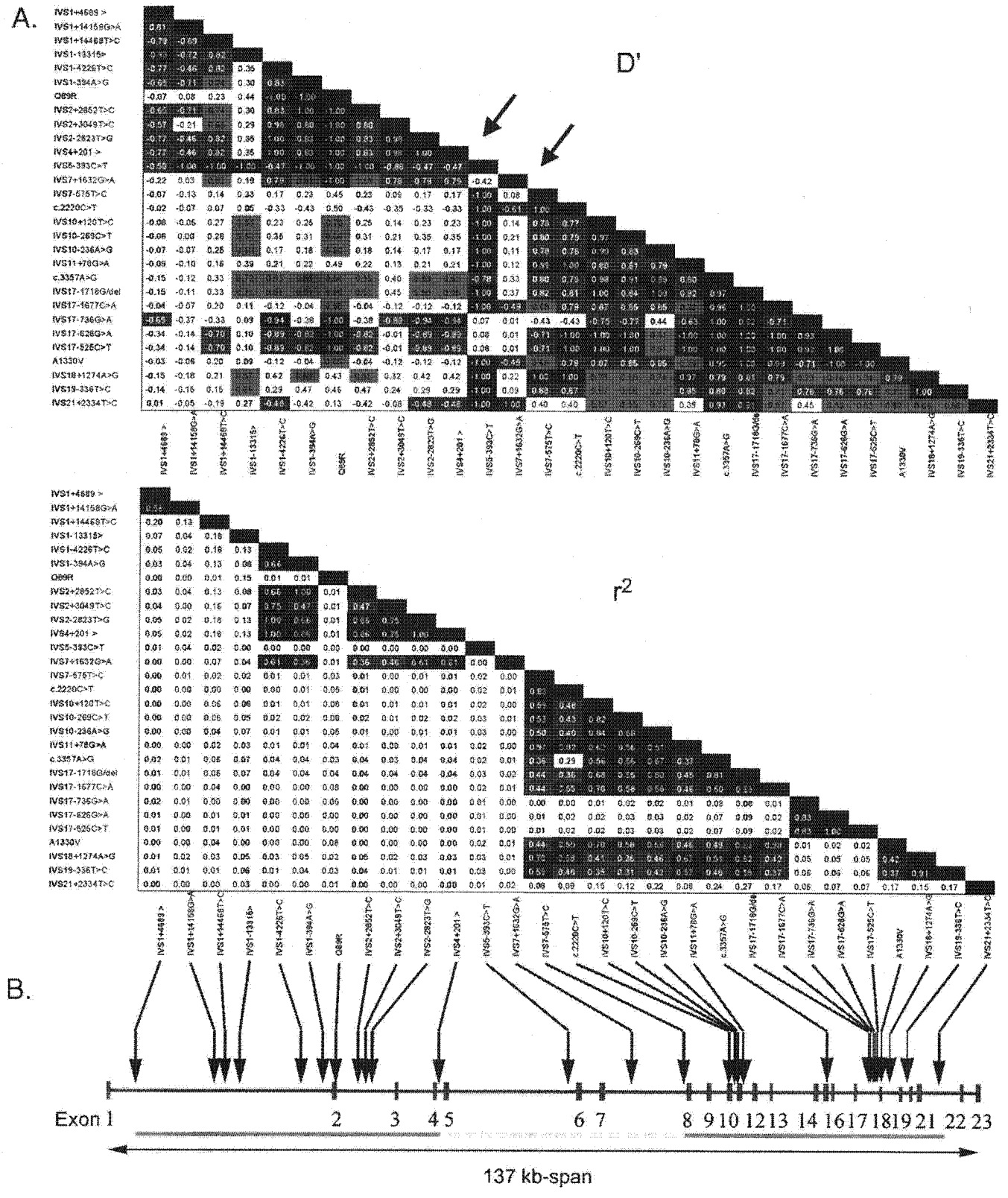

Fig. 1. Linkage disequilibrium analysis among 29 SNPs within the *LRP5* locus. (A) Indices of linkage disequilibrium D' (upper half) and r^2 (lower half), calculated from the predicted haplotype frequencies and presented as separate tables. D' values greater than 0.5 and r^2 values greater than 0.1 are highlighted by gray-to-black gradients. Arrows indicate a possible hot spot for recombination. (B) Schematic diagram of the *LRP5* gene showing locations of the 23 exons. Locations of the 29 tested SNPs are indicated by downward arrows.

Table 4
Contingency table analysis on Group-A subjects

Group-A	n	A1330V			P value[a]
		AA	AV	VV	
Z-score > 3.0	8	5	2	1	NA
Z-score < −3.0	4	1	3	0	
Z-score > 2.0	33	22	8	3	0.025*
Z-score < −2.0	34	13	14	7	
Z-score > 1.0	91	52	33	6	0.046*
Z-score < −1.0	139	61	63	15	

NA: not applicable.
[a] P values are calculated in chi-square test for trend.
* $P < 0.05$.

Haplotyping, LD analysis, and statistical analysis

Maximum likelihood haplotype frequencies among the 582 chromosomes of 291 subjects from Group-B were estimated by an EM algorithm using SNPAlyze v3.1 (DYNA-COM, Chiba, Japan). The LD for all possible two-way combinations of SNPs was tested with D, D', and r^2 [26,27]. After defining LD blocks and tag-SNPs in each block, each individual diplotype was estimated by SNPAlyze v3.1 software.

Quantitative associations between genotypes and adjusted BMD values (g/cm^2) were examined by analysis of variance (ANOVA), with regression analysis as a post hoc test using Instat 3 software (GraphPad Software, San Diego, CA). The three genotypic categories of each SNP were converted into incremental values (0, 1, and 2), which represent the number of chromosomes carrying the major allele. Significant association was defined when the given P value of the ANOVA F-test was less than 5% ($P < 0.05$). Similarly, quantitative association was tested for major 5- and 4-haplotypes (frequency >5%) defined for each LD block in the *LRP5* gene locus. Chi-square tests were used to ascertain Hardy–Weinberg equilibrium among genotypes ($P > 0.05$). Multiple linear regression analysis was applied for examining potential combined effects of Q89R and A1330V alleles, using Instat 3 software. Distribution analysis of the genotype frequencies among phenotypically divided groups, i.e., BMD Z-scores > 1.0 ($n = 91$) and Z-scores < −1.0 ($n = 139$) was conducted by chi-square tests as in our previous experiments [21]. Predictive analysis of protein function for the two missense coding SNPs was conducted using the SIFT (Sorting Intolerant From Tolerant) program (http://blocks.fhcrc.org/sift/SIFT.html) [28]. When the given score was less than 0.05, the alteration was deemed to be intolerant (i.e., deleterious).

Results

To identify the genetic ground involved in a previously detected association of an intronic *LRP5* polymorphism with low BMD in Group-A subjects [15], we first searched for a causative mutation anywhere among the 23 exons of the gene. However, among 12 subjects who had extremely high or low BMD Z-scores (>3.0 or <−3.0), no mutations were detected apart from two already known missense variations (Q89R and A1330V). We analyzed the effect of these amino acid alterations, using predictive computer program SIFT [28]; this program estimated a deleterious effect of 1330-V (score = 0.03) and possible mild effect of 89-R (score = 0.19). Although the lowest score was predicted for another missense variation V667M (score = 0.01), this variant was not detected in our test population. Based on the notion that strong candidate for the causative polymorphism would be missense variation, we investigated association of these SNPs among the entire Group-A subjects.

To analyze the effects on BMD, all the extended 387 subjects were genotyped for these two SNPs as well as previously tested IVS17-1677C > A (Table 2). No association of the Q89R genotypes was detected for BMD Z-score ($r = 0.03$, NS). However, we detected significant correlation of the A1330V genotype with spinal BMD Z-score ($r = 0.11$, $P = 0.034$); homozygous carriers of the minor T-allele (V/V) had the lowest BMD Z-scores (−0.47 ± 1.47, $n = 35$), heterozygous individuals (A/V) were intermediate (−0.35 ± 1.38; $n = 174$), and homozygous carriers of the C-allele (A/A) had the highest BMD Z-scores (−0.03 ± 1.62; $n = 178$) (Tables 2 and 3). The result was consistent with previously detected results for the intron SNP IVS17-1677C > A [15], as indicated (Table 2). We noticed that these SNPs were in strong linkage disequilibrium (see Fig. 1). Consistent results were achieved in a distribution analysis of phenotypically divided subjects (high BMD Z-scores: $n = 91$ and low BMD Z-scores: $n = 139$), using 2 × 3 chi-square tests for detecting trends ($P = 0.046$) (Table 4). A similar result was obtained by analyzing subjects with more prominent phenotypes (Z-scores > +2.0: $n = 33$, and Z-scores < −2.0: $n = 34$; $P = 0.025$) (Table 4).

No significant differences were detected in physical characteristics (body weight, height, or BMI) between groups genotypically classified according to A1330V alleles. Although we detected a significant correlation of A1330V with levels of intact osteocalcin in serum ($P = 0.004$), this might be affected by a correlation detected for age distribution ($P = 0.01$) (Table 3).

To examine the reproducibility of that correlation, 384 adult Japanese women recruited from general Japanese population were analyzed (Group-B) on adjusted values of radial BMD by ANOVA with linear regression. As a result, significant correlation of the A1330V genotype with adjusted BMD was replicated (Tables 2 and 5). Homozygous carriers of the minor

Table 5
Physical and clinical characteristics of the subjects in Group-B (cohort subjects)

Group-B subjects (n = 384)	A1330V			P value*
	AA	AV	VV	
Number	195	155	34	–
Age (Years)	58.4 ± 8.72 (34–69)	58.3 ± 8.14 (32–69)	58.9 ± 9.93 (32–69)	NS
Weight (kg)	53.2 ± 9.01 (34.8 ± 84.8)	55.1 ± 8.73 (28.8–89.0)	55.9 ± 8.7 (44.6–80.0)	0.027
BMI (kg/m^2)	23.3 ± 3.63 (16.2–37.7)	23.9 ± 3.61 (14.7–38.5)	24.6 ± 3.29 (17.5–31.3)	0.018
Height (cm)	151.2 ± 6.04 (130.5–169.0)	151.7 ± 5.41 (139.0–166.0)	150.6 ± 6.65 (140.0–163.4)	NS
Radial BMD (g/cm^2)	0.402 ± 0.082 (0.222–0.635)	0.398 ± 0.084 (0.178–0.613)	0.390 ± 0.079 (0.270–0.566)	NS
Adjusted BMD (g/cm^2)	0.405 ± 0.053 (0.275–0.551)	0.395 ± 0.059 (0.225–0.554)	0.385 ± 0.041 (0.298–0.448)	0.019

* P values are calculated for regression analysis with ANOVA F-test.

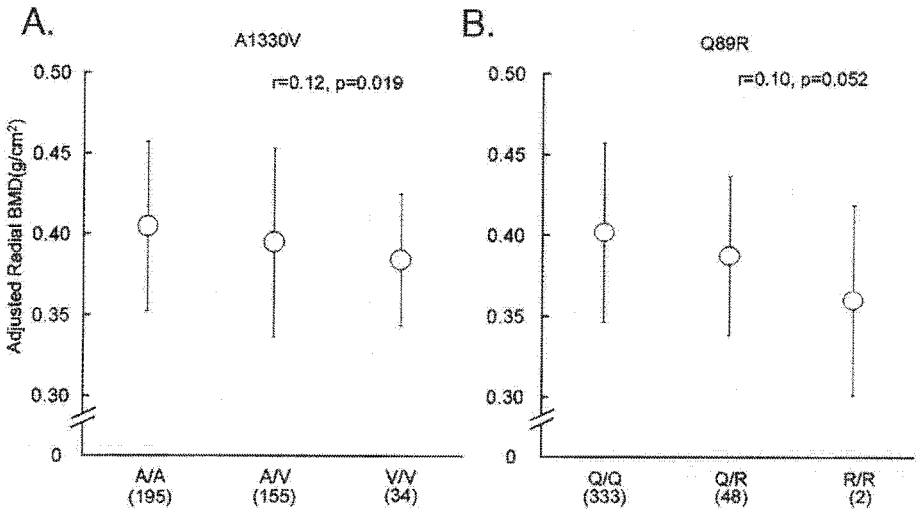

Fig. 2. Reproducible association of two missense *LRP5* variations (Q89R and A1330V) with adjusted BMD. (A) Plots of adjusted BMD values for three genotypically classified subgroups by A1330V among 384 subjects from the general population. (B) Plots of adjusted BMD values for three genotypically classified subgroups by Q89R. Open circles indicate mean values and error bars indicate standard deviations. Correlations between the number of minor allele possessed and the adjusted BMD were tested by linear regression analysis.

T-allele (V/V) had the lowest adjusted BMDs (0.385 ± 0.041 g/cm^2, $n = 34$); heterozygous individuals (A/V) were intermediate (0.395 ± 0.059 g/cm^2; $n = 155$); and homozygous C-allele carriers (A/A) had the highest adjusted BMDs (0.405 ± 0.053 g/cm^2; $n = 195$), indicating a possible codominant effect of the minor T-allele (corresponding to the second nucleotide for a valine codon, gTg) on lowering adjusted BMD levels (Fig. 2A; $r = 0.12$, $P = 0.019$; $n = 384$). Interestingly in these subjects, we observed a tendency for association of Q89R minor G-allele (corresponds to arginine; cGg) to low adjusted BMD, although this tendency did not quite reach statistical significance (Fig. 2B; $r = 0.10$, $P = 0.052$). We also examined if any patient characteristics correlated with the genotype; however, no significant differences were detected except in body weight ($r = 0.11$, $P = 0.03$) or BMI ($r = 0.12$, $P = 0.02$). A contribution of A1330V to control of body mass appeared to be a possibility (Table 5). For analysis of possible combined effects of Q89R and A1330V on BMD determination, multivariate linear regression analysis estimated a fitting equation explained about 2% of BMD variances ($r^2 = 0.023$, $P = 0.01$). However only the A1330V genotype contributed significantly to the equation ($P = 0.004$).

We then analyzed LD in the LRP5 locus by genotyping these subjects for informative 29 SNPs (minor allele frequencies >0.05) (Table 1). Average distance between neighboring SNPs was 4.6 kb (median, 3.0 kb; range, 33–21,779 bp). Genotype, allelic frequency, and heterozygosity were clarified successfully for each SNP, and no deviation was detected from Hardy–Weinberg equilibrium. Indices of pairwise LD (D' and r^2) were calculated by estimating the maximum-likelihood haplotypes and their frequencies from 293 Group-G subjects genotyped for all 29 SNPs. These procedures detected two LD blocks separated by a region of about 11 kb, from intron 5 to intron 7 (Fig. 1). In the second block, the indices of pairwise LD between the A1330V and IVS17-1677C > A variants were remarkably high ($D' = 0.99$, $r^2 = 0.98$). More important, the two missense SNPs, Q89R (c.266A > G) and A1330V (c.3989C > T), appeared to localize in different LD blocks, although the indices of pairwise LD between the two SNPs were slightly higher than the neighboring intronic SNPs ($D' = 0.64$, $r^2 = 0.07$). By analyzing haplotype frequency estimated by all the belonging SNPs for each LD block (11 and 18, respectively), representative 4 and 5 SNPs were selected (Tables 6 and 7). By estimating diplotype in each individual,

Table 6
Representative haplotypes for LD block-1

Hap-No	IVS1 + 4689C > G	IVS1 + 14158G > A	IVS1 + 14468T > C	IVS2 + 2852T > C	Frequency (%)	SUM (%)
#1–1	0	0	0	0	37.6	
#1–2	1	1	0	0	26.4	
#1–3	1	1	1	0	15.0	
#1–4	1	1	1	1	6.5	
#1–5	0	1	0	0	6.3	
#1–6	1	0	1	0	2.1	
#1–7	1	0	0	0	2.1	
#1–8	0	0	1	0	1.0	
#1–9	0	0	0	1	0.9	98.0

0 and 1 represent major and minor alleles, respectively.

Table 7
Representative haplotypes for LD block-2

Hap-No	IVS7-575T > C	c.3357A > G	IVS 17-626G > A	A1330V	IVS21 + 2334T > C	Frequency (%)	SUM (%)
#2–1	0	0	0	0	0	53.0	
#2–2	1	1	0	1	0	13.9	
#2–3	1	1	0	1	1	9.2	
#2–4	0	1	0	0	0	3.8	
#2–5	1	1	0	0	0	3.5	
#2–6	0	1	0	1	0	3.4	
#2–7	1	0	0	0	0	3.2	
#2–8	0	1	1	0	1	2.9	
#2–9	0	1	0	1	1	1.8	
#2–10	0	1	1	0	0	1.6	
#2–11	0	1	0	0	1	1.4	97.6

0 and 1 represent major and minor alleles, respectively.

association was examined for 9 haplotypes (5 and 4 haplotypes for each block). Significant association was represented by haplotype-#2–3 (C-G-G-V-C) for the second LD block (r = 0.15, P = 0.0041); homozygous and heterozygous haplotype#3 carriers had lower adjusted BMD (0.384 ± 0.049, n = 6 and 0.379 ± 0.057, n = 57) compared to non-carriers (0.402 ± 0.053, n = 299).

Discussion

In the study reported here, we investigated a possible contribution of *LRP5* polymorphisms to determination of bone-mineral density in Japanese women. A search for mutations, an LD analysis, and a study for association between two candidate SNPs and adjusted BMD values in two independently collected groups revealed reproducible association of the *LRP5* variations with BMD levels. During the preparation and submitting this manuscript, five independent groups reported similar results. One reported consistent correlation of A1330V with radial BMD in adult Japanese women [16], and the other reported a similar tendency of Q89R to be associated with hip BMD in young Korean males [17]. Although another study, involving Caucasians, reported a c.2047G > A variation (V667M: nucleotide numbering is from the cap site of mRNA in that report [18]), that variant appeared to be so rare in our test subjects that we excluded it from our investigations. Recent two articles further supported consistent association of A1330V, although both reports emphasized association of other intron SNPs in strong LD with A1330V [29,30]. We propose that *LRP5* variations may be important determinants of BMD in the general population, and believe that the A1330V polymorphism is the most likely LRP5 determinant of bone mass as well, at least in adult Japanese women.

Although the molecular basis for the function of LRP5 in determining BMD is not fully clarified, evidence is accumulating. Several mutations in the *LRP5* gene have been identified as causing inherited diseases with bone phenotypes, including OPPG and autosomal-dominant types of osteopetrosis [8–11]. Functional molecular analyses, including gene targeting experiments and analysis of mutated LRP5 products, have revealed a mechanistic basis for bi-directional phenotypic expression, where a key is the Wnt pathway and binding of its antagonists to the YWTD motif in the extracellular domains of this co-receptor [4,10,31]. So far, however, the cellular and in vivo mechanisms causing reduction of bone mass are not well understood. Therefore, functional differences focusing on the alanine residue at 1330 and the LDLR class-A domain should be investigated, and binding affinity for Wnt ligands or other co-binding regulators like dickkopfs (Dkks) [6] should be examined. In our experiments, a predictive informatics program (SIFT [28]) for functional alterations caused by the two missense variations in *LRP5* estimated an intolerable alteration of alanine to valine at A1330V (score = 0.03) but not for glutamine to arginine in Q89R (score = 0.19). Those estimates were consistent with our assumption from association analysis that the effect of Q89R might be weaker than that of A1330V. However, a cross-sectional study conducted elsewhere had indicated a selective influence of a V667M variant on vertebral bone mass among Caucasians where the SIFT program predicted an intolerant effect of 667M (score = 0.01). Longitudinal human studies, or an in vivo animal study, may clarify the developmental period(s) during which *LRP5* variants affect BMD most strongly.

The feasibility of a candidate-gene association study generally depends on control of confounding factors resulting from population bias, unreliability of diagnosis or phenotypic evaluation, interaction with environmental factors, or deficiencies in statistical power [32–34]. In our study, a reliable measurement of BMD, a quantitative trait with significant heredity, was evaluated with adjustment for age and BMI among unbiased population with no evidence of the stratification. Moreover, the reproducibly demonstrated association of the 1330 V variant with low BMD among independent groups of subjects, in other studies and ours [16,17], seems to provide comprehensive evidence for a functional contribution of this variant, even though in all of the studies cited, the significance levels were moderate without adjustment for multiple testing [16–18]. Supportive data from linkage analysis includes a quantitative trait locus (QTL) for spinal BMD at 11q12–13 that has been identified in general populations [35]. A possible combined effect of the multiple coding SNP(s) tested in our

multiple regression analysis should be re-evaluated in a larger cohort study.

The existence of overlapping mechanisms for lipid metabolism, body mass regulation, and bone metabolism is a classically discussed issue [36–38]. These processes are under systemic control of leptin and neuro-endocrinological systems [37–39], but are influenced by apolipoprotein E polymorphisms [40], subject to developmental controls of mesenchymal-cell differentiation by transcription factors like PPAR-gamma [41], and affected by statins [42]. In addition, the Wnt-LRP system may affect regulation of both bone and body mass [43]. Our data support that notion because body weight and BMI, as well as adjusted BMD, correlated with A1330V genotypes among 384 adult women from the general Japanese population. Other effector molecules in the Wnt signaling system might contribute as well.

In summary, our data suggest a functional effect of polymorphic variants in a candidate osteoporosis-susceptibility gene, *LRP5*, whose common polymorphisms significantly correlated with bone mineral density of women recruited from two independent populations. Because osteoporosis is a multifactorial disease, other genes, especially genes acting through the Wnt pathway, may have to be examined for potential contributions to disease susceptibility. That information will help to clarify the complex mechanism of BMD determination in vivo, and may explain, at least in part, the pathogenesis of postmenopausal osteoporosis. Such studies should provide a novel viewpoint for establishing suitable treatment designs and preventive strategies for the disease.

Acknowledgments

We thank Dr. Tatsuhiko Tsunoda at RIKEN SNP Research Center for critical reading of our manuscript, and Mina Kodaira, Miho Kawagoe, and Tamami Uchida for technical assistance. This work was supported by a grant for Strategic Research from the Ministry of Education, Science, Sports and Culture of Japan; by a Research Grant for Research from the Ministry of Health and Welfare of Japan; and by a Research for the Future Program Grant of The Japan Society for the Promotion of Science.

References

[1] Riggs BL, Melton III LJ. Involutional osteoporosis. N Engl J Med 1986;314:1676–86.
[2] Hartmann C, Tabin DJ. Wnt-14 plays a pivotal role in inducing synovial joint formation in the developing appendicular skeleton. Cell 2001;140: 341–51.
[3] Yang Y, Topol L, Lee H, Wu J. Wnt5a and Wnt5b exhibit distinct activities in coordinating chondrocyte proliferation and differentiation. Development 2003;130:1003–15.
[4] Kato M, Patel MS, Levasseur R, Lobov I, Chang BHJ, Glass II DA, et al. Cbfa1-independent decrease in osteoblast proliferation, osteopenia, and persistent embryonic eye vascularization in mice deficient in Lrp5, a Wnt coreceptor. J Cell Biol 2002;157:303–14.
[5] Mao J, Wang J, Liu B, Pan W, Farr III GH, Flynn C, et al. Low-density lipoprotein receptor-related protein-5 binds to axin and regulates the canonical Wnt signaling pathway. Mol Cell 2001;7:801–9.
[6] Zorn AM. Wnt signaling: antagonistic Dickkopfs. Curr Biol 2001;11: R592–5.
[7] Gong Y, Vikkula M, Boon L, Liu J, Beighton P, Ramesar R, et al. Osteoporosis-pseudoglioma syndrome, a disorder affecting skeletal strength and vision, is assigned to chromosome region 11q12–13. Am J Hum Genet 1996;59:146–51.
[8] Gong Y, Slee RB, Fukai N, Rawadi G, Roman-Roman S, Reginato AM, et al. LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development. Cell 2001;107:513–23.
[9] Little RD, Carulli JP, Del Mastro RG, Dupuis J, Osborne M, Folz C, et al. A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait. Am J Hum Genet 2002;70: 11–9.
[10] Boyden LM, Mao J, Belsky J, Mitzner L, Farhi A, Mitnick MA, et al. High bone density due to a mutation in LDL-receptor-related protein 5. N Engl J Med 2002;346:1513–21.
[11] Van Wesenbeeck L, Cleiren E, Gram J, Beals RK, Benichou O, Scopelliti D, et al. Six novel missense mutations in the LDL receptor-related protein 5 (LRP5) gene in different conditions with an increased bone density. Am J Hum Genet 2003;72:763–71.
[12] Grant SFA, Reid DM, Blake G, Herd R, Fogelman I, Ralston SH. Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I α 1 gene. Nat Genet 1996; 14:203–5.
[13] Tavtigian SV, Simard J, Teng DH, Abtin V, Baumgard M, Beck A, et al. A candidate prostate cancer susceptibility gene at chromosome 17p. Nat Genet 2001;27:172–224.
[14] Rebbeck TR, Walker AH, Zeigler-Johnson C, Weisburg S, Martin AM, Nathanson KL, et al. Association of HPC2/ELAC2 genotypes and prostate cancer. Am J Hum Genet 2000;67:1014–9.
[15] Urano T, Shiraki M, Ezura Y, Fujita M, Sekine E, Hoshino S, et al. Association of a single-nucleotide polymorphism in low-density lipoprotein receptor-related protein 5 gene with bone mineral density. J Bone Miner Metab 2004;22:341–5.
[16] Mizuguchi T, Furuta I, Watanabe Y, Tsukamoto K, Tomita H, Tsujihata M, et al. LRP5, low-density-lipoprotein-receptor-related protein 5, is a determinant for bone mineral density. J Hum Genet 2004;49:80–6.
[17] Koh JM, Jung MH, Hong JS, Park HJ, Chang JS, Shin HD, et al. Association between bone mineral density and LDL receptor-related protein 5 gene polymorphisms in young Korean men. J Korean Med Sci 2004;19:407–12.
[18] Ferrari SL, Deutsh S, Choudhury U, Chevalley T, Bonjour JP, Dermitzakis ET, et al. Polymorphisms in the low-density lipoprotein receptor-related protein 5 (LRP5) gene are associated with variation in vertebral bone mass, vertebral bone size, and stature in whites. Am J Hum Genet 2004; 74:866–75.
[19] Orimo H, Hayashi Y, Fukunaga M, Sone T, Fujiwara M, Shiraki M, et al. Diagnostic criteria for primary osteoporosis: year 2000 revision. J Bone Miner Metab 2001;19:331–7.
[20] Ezura Y, Nakajima T, Kajita M, Ishida R, Inoue S, Yoshida H, et al. Association of molecular variants, haplotypes, and linkage disequilibrium within the human vitamin D-binding protein (DBP) gene with postmenopausal bone mineral density. J Bone Miner Res 2003;18:1642–9.
[21] Iwasaki H, Emi M, Ezura Y, Ishida R, Kajita M, Kodaira M, et al. Association of a Trp16Ser variation in the gonadotropin releasing hormone (GnRH) signal peptide with bone mineral density, revealed by SNP-dependent PCR (Sd-PCR) typing. Bone 2002;32:185–90.
[22] Ogawa S, Hosoi T, Shiraki M, Orimo H, Emi M, Muramatsu M, et al. Association of estrogen receptor b gene polymorphism with bone mineral density. Biochem Biophys Res Commun 2000;269:537–41.
[23] Okubo M, Horinishi A, Kim DH, Yamamoto TT, Murase T. Seven novel sequence variations in the human low density lipoprotein receptor related protein 5 (LRP5) gene. Hum Mutat 2002;19:186.
[24] Mein CA, Barratt BJ, Dunn MG, Siegmund T, Smith AN, Esposito L, et al. Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. Genome Res 2000;10:330–43.
[25] Haga H, Yamada R, Ohnishi Y, Nakamura Y, Tanaka T. Gene-based SNP discovery as part of the Japanese Millennium Genome Project:

identification of 190 562 genetic variations in the human genome. J Hum Genet 2002;47:605–10.
[26] Miller PT, Sardina IB, Saccone NL, Putzel J, Laitinen T, Cao A, et al. Juxtaposed regions of extensive minimal linkage disequilibrium in human Xq25 and Xq28. Nat Genet 2000;25:324–8.
[27] Thompson EA, Deeb S, Walker D, Motulsky AG. The detection of linkage disequilibrium between closely linked markers: RFLPs at the AI-CIII apolipoprotein genes. Am J Hum Genet 1988;42:113–24.
[28] Ng PC, Henikoff S. Accounting for human polymorphisms predicted to affect protein function. Genome Res 2002;12:436–46.
[29] Koay MA, Woon PY, Zhang Y, Miles LJ, Duncan EL, Ralston SH, et al. Influence of LRP5 polymorphisms on normal variation in BMD. J Bone Miner Res 2004;19:1619–27.
[30] Koller DL, Ichikawa S, Johnson ML, Lai D, Xuei X, Edenberg HJ, et al. Contribution of the LRP5 gene to normal variation in peak BMD in women. J Bone Miner Res 2005;20:75–80.
[31] Babij P, Zhao W, Small C, Kharode Y, Yaworsky PJ, Bouxsein ML, et al. High bone mass in mice expressing a mutant LRP5 gene. J Bone Miner Res 2003;18:960–74.
[32] Eisman JA. Genetics of osteoporosis. Endocr Rev 1999;20:788–804.
[33] Peacock M, Turner CH, Econs MJ, Foroud T. Genetics of osteoporosis. Endocr Rev 2002;23:303–26.
[34] Hobson EE, Ralston SH. Role of genetic factors in the pathophysiology and management of osteoporosis. Clin Endocrinol 2001;54:1–9.
[35] Koller DL, Econs MJ, Morin PA, Christian JC, Hui SL, Parry P, et al. Genome screen for QTLs contributing to normal variation in bone mineral density and osteoporosis. J Clin Endocrinol Metab 2000; 85:3116–20.
[36] Reid IR. Relationships among body mass, its components, and bone. Bone 2002;31:547–55.
[37] Amling M, Takeda S, Karsenty G. A neuro(endo)crine regulation of bone remodeling. BioEssays 2000;22:970–5.
[38] Whipple T, Sharkey N, Demers L, Williams N. Leptin and the skeleton. Clin Endocrinol 2002;57:701–11.
[39] Takeda S, Elefteriou F, Levasseur R, Liu X, Zhao L, Parker KL, et al. Leptin regulates bone formation via the sympathetic nervous system. Cell 2002;111:305–3017.
[40] Shiraki M, Shiraki Y, Aoki C, Hosoi T, Inoue S, Kaneki M, et al. Association of bone mineral density with apolipoprotein E phenotype. J Bone Miner Res 1997;12:1438–45.
[41] Lecka-Czernik B, Moerman EJ, Grant DF, Lehmann JM, Manolagas SC, Jilka RL. Divergent effects of selective peroxisome proliferator-activated receptor-gamma 2 ligands on adipocyte versus osteoblast differentiation. Endocrinology 2002;143:2376–84.
[42] McFarlane SI, Muniyappa R, Francisco R, Sowers JR. Pleiotropic effects of statins: lipid reduction and beyond. J Clin Endocr Metab 2002;87: 1451–8.
[43] Johnson ML, Harnish K, Nusse R, Van Hul W. LRP5 and Wnt signaling: a union made for bone. J Bone Miner Res 2005;19:1749–57.

Oncogene (2007), 1–11

www.nature.com/onc

ONCOGENOMICS

Identification of novel androgen response genes in prostate cancer cells by coupling chromatin immunoprecipitation and genomic microarray analysis

K Takayama[1,3], K Kaneshiro[2], S Tsutsumi[2], K Horie-Inoue[3], K Ikeda[3], T Urano[1,3], N Ijichi[3], Y Ouchi[1], K Shirahige[4], H Aburatani[2,5] and S Inoue[1,3]

[1]*Department of Geriatric Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan;* [2]*Genome Science Division, Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan;* [3]*Division of Gene Regulation and Signal Transduction, Research Center for Genomic Medicine, Saitama Medical University, Saitama, Japan;* [4]*Division for Gene Research, Center for Biological Resources and Informatics, Tokyo Institute of Technology, Tokyo, Japan and* [5]*CREST, Japan Science and Technology Agency (JST), Tokyo, Japan*

The androgen receptor (AR) plays a key role as a transcriptional factor in prostate development and carcinogenesis. Identification of androgen-regulated genes is essential to elucidate the AR pathophysiology in prostate cancer. Here, we identified androgen target genes that are directly regulated by AR in LNCaP cells, by combining chromatin immunoprecipitation (ChIP) with tiling microarrays (ChIP-chip). ChIP-enriched or control DNAs from the cells treated with R1881 were hybridized with the ENCODE array, in which a set of regions representing approximately 1% of the whole genome. We chose 10 bona fide AR-binding sites (ARBSs) ($P<1e-5$) and validated their significant AR recruitment ligand dependently. Eight upregulated genes by R1881 were identified in the vicinity of the ARBSs. Among the upregulated genes, we focused on UGT1A and CDH2 as AR target genes, because the ARBSs close to these genes (in UGT1A distal promoter and CDH2 intron 1) were most significantly associated with acetylated histone H3/H4, RNA polymerase II and p160 family co-activators. Luciferase reporter constructs including those two ARBSs exhibited ligand-dependent transcriptional regulator/enhancer activities. The present study would be powerful to extend our knowledge of the diversity of androgen genetic network and steroid action in prostate cancer cells.
Oncogene advance online publication, 5 February 2007; doi:10.1038/sj.onc.1210229

Keywords: androgen receptor; androgen response element; chromatin immunoprecipitation; prostate cancer; UGT1A

Correspondence: Dr S Inoue, Department of Geriatric Medicine, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.
E-mail: INOUE-GER@h.u-tokyo.ac.jp
Received 18 October 2006; revised 20 November 2006; accepted 21 November 2006

Introduction

Androgen is a key regulator of male sexual differentiation as well as prostate development and carcinogenesis. Androgen-regulated gene expression is mediated by the action of androgen receptor (AR), which is a member of nuclear receptor superfamily that functions as a ligand-dependent transcription factor. Prostate cancer is originally an androgen-dependent tumor, whose growth and survival are under the control of AR signaling. Thus, androgen deprivation is the most common option of the cancer treatment. The therapy, however, eventually fails and most patients will relapse owing to adaptive progression of the surviving prostate cancer cells. The recurrent cancer is usually referred to as 'androgen independent' (Grossmann *et al.*, 2001). Nevertheless, advanced prostate cancer often continues to express AR and androgen-regulated genes, suggesting a functional role of AR in the recurrent stage. Alterations of the AR gene including mutation and amplification are also shown in some recurrent tumors, but these mechanisms will not explain the hormone-refractory responses in the majority of patients after androgen deprivation. Indeed, a modest increase in AR mRNA has been shown to be associated with the resistance to anti-androgen therapy in isogenic prostate cancer xenograft models (Chen *et al.*, 2004). Therefore, understanding the global aspects of AR signaling network and the distinct roles of AR target genes are essential for the development of new diagnostic procedures and therapeutic options for prostate cancer in various disease states.

AR regulates the expression of target genes by binding to androgen response elements (AREs) in the genome, or by interacting with other transcription factors bound to their specific recognition sites. AR-mediated gene transcription has been studied using prostate-specific antigen (PSA) as a prototypic model, and AREs in PSA promoter and enhancer have been shown to recruit various co-activators and general transcription factors including histone acetyltransferases,

p160 family, mediator and RNA polymerase II (PolII) (Wang *et al.*, 2005). Efforts have been paid to search various androgen target genes by using microarray techniques since last decade, identifying hundreds of genes with altered expression by hormone stimulation in cells. The gene expression profiling is powerful to depict the global function of androgen in a specified model; however, the technique will not be suitable to determine whether the alteration of gene expression is owing to direct or indirect action of AR transcription. Recent advance of human genome project enables to search putative AREs bioinformatically in the transcription regulatory regions of androgen target genes; yet, few AREs are identified as physiological elements in AR signaling (Horie-Inoue *et al.*, 2004, 2006). Thus, the development of a new high-throughput method that identifies bona fide AR-binding sites (ARBSs) in the genome is a prerequisite for the elucidation of AR gene network.

Recently, a combined technique of chromatin immunoprecipitation (ChIP) analysis with DNA microarray has been established to identify chromatin-interacting domains of transcription factors in a genome-wide manner (Cawley *et al.*, 2004; Bernstein *et al.*, 2005). Regarding nuclear receptors, ligand-dependent estrogen receptor (ER)-binding sites have been recently shown by this ChIP-chip technique using Affymetrix-tiling oligonucleotide microarrays of chromosomes 21 and 22 (Carroll *et al.*, 2005), or a custom-made promoter microarrays (Laganière *et al.*, 2005). In this study, we have performed ChIP-chip using a sampler DNA microarray of the human genome, the so-called ENCODE chip. In this microarray, a set of regions representing approximately 1% (30 Mb) of the whole genome are included as the target for the pilot project that has been selected by the research consortium of the ENCyclopedia of DNA Elements (ENCODE Project Consortium, 2004). Fifty percent of the 30-Mb genomic regions, consisting of 14 regions (ENm001–ENm014), were manually selected, and the remaining 50% were composed of 30, 500 kb regions (ENr111–ENr334) selected according to a stratified random-sampling strategy based on gene density and level of non-exonic conservation.

Here, we find a discrete number of ARBSs in the selected regions of the ENCODE regions. Intriguingly, most of the AR-interacting regions have been shown to locate in non-promoter proximal regions; yet, they contained ARE sequences and were validated to recruit AR ligand dependently. In the vicinity of the functional ARBSs, we found several genes with upregulated transcript levels by hormone stimulation. Some of the AR target genes that have been identified in this study are previously known to be associated with AR expression, whereas some are novel targets. Our ChIP-chip analysis and transcriptional study indicate that non-promoter ARBSs play roles in the AR-dependent transcriptional regulation, potentially dissecting a series of AR-regulated mechanisms in a genome-wide manner.

Results

Screen of ARBSs on ENCODE DNA microarray

To perform a screen of ARBSs in AR-positive cells on tiling oligonucleotide microarrays, we first investigated the time course of ligand-dependent AR recruitment in human prostate cancer LNCaP cells. After 3-day hormone depletion, cells were stimulated with vehicle or a synthetic androgen R1881 (10 nM) for 2, 6 or 24 h. Cross-linked protein–DNA complexes extracted from the cells were immunoprecipitated with anti-AR antibody, and quantitative polymerase chain reaction (qPCR) for ARE regions in the proximal promoter and enhancer of PSA was performed using the purified precipitated DNAs as templates. AR binding in response to ligand stimulation exhibited maximal levels at 24 h (data not shown).

We next performed ChIP-chip analyses using the ENCODE tiling microarrays comprised of the total 30-Mb human genomic DNA, which corresponds to 1% of the genome. The chromatin DNAs immunoprecipitated by anti-AR or without ChIP (input control) were amplified unbiasedly by *in vitro* transcription (IVT), and the amplified DNAs were fragmented and biotin-labeled, then hybridized with the ENCODE chips for duplication. Using the Affymetrix Tiling Analysis Software, raw intensity data of duplicate arrays for each experimental group were transformed and signal and *P*-values for each genomic position interrogated were determined after quantile normalization. The results were mapped to genomic positions that could be visualized in the Affymetrix Integrated Genome Browser or the UCSC Genome Browser (NCBI Build 35). Applying a *P*-value cutoff of 1e-5 for a significant AR binding, we identified 10 ARBSs (Table 1) in the ENCODE genomic regions. Among them, five ARBSs were involved in the manually defined regions of the 30-Mb ENCODE regions (ARBSs no. 3–no. 7), whereas the remaining five binding sites were derived from the randomly selected regions.

Notably, most of the ARBSs were located within intronic regions or gene upstream regions at least 10 kb apart from the transcriptional start sites (TSSs) of their closest genes. One of the ARBSs included in the ENCODE chips was ARBS no. 1, which was located adjacent to UGT1A locus, in the 5′ upstream region > 17 kb upstream of UGT1A1 gene TSS or in intron 1 of UGT1A3 on chromosome 2q37. As another example, ARBS no. 10 was situated in intron 1 of CDH2 on chromosome 18q11.2 (Figure 1).

We next investigated whether the 10 ARBSs included sequences highly similar to the previously established consensus AREs. Using a weighted matrix-based finder TRANSFAC (Matys *et al.*, 2003) with the matrix conservation > 75% or a sequence analysis utility of JASPER with the relative profile score threshold > 70% (Sandelin *et al.*, 2004), we identified canonical ARE sequences in all of the ARBSs (Table 2).

To verify whether the identified ARBSs in ChIP-chip were authentic ARBSs in the genome, we performed new independent ChIP experiments in LNCaP cells. We

Table 1 ARBSs identified in the encyclopedia of DNA element (ENCODE) regions by ChIP-chip experiments

ARBS no.	*ENCODE region*	*Chromosome*	*Start*	*Stop*	*Closest gene*	*Distance from TSS (bp)*[a]	*Position*
1	ENr131	2	234 433 433	234 433 623	UGT1A1	−17 390	5′ upstream
2	ENr334	6	41 823 411	41 823 433	PGC	−323	5′ upstream
3	ENm010	7	26 807 344	26 807 566	SCAP2	−129 874	5′ upstream
4	ENm013	7	89 501 530	89 501 614	STEAP2	+15 922	intron 3
5	ENm013	7	89 980 335	89 980 369	PFTK1	−3010	5′ upstream
6	ENm001	7	115 551 316	115 551 473	TES	+106 864	3′ downstream
7	ENm001	7	116 022 567	116 022 922	MET	+116 336	intron 17
8	ENr233	15	41 640 443	41 640 980	KIAA0377	+23 671	intron 27
9	ENr233	15	41 739 974	41 740 495	CATSPER2	−11 904	5′ upstream
10	ENr213	18	23 990 311	23 990 793	CDH2	+20 637	intron1

Abbreviations: ARBs, androgen receptor-binding sites; TSS, transcriptional start site. [a]Distance from the TSS of the closet RefSeq gene to the corresponding ARBS based on the genomic position published in NCBI Build 35.

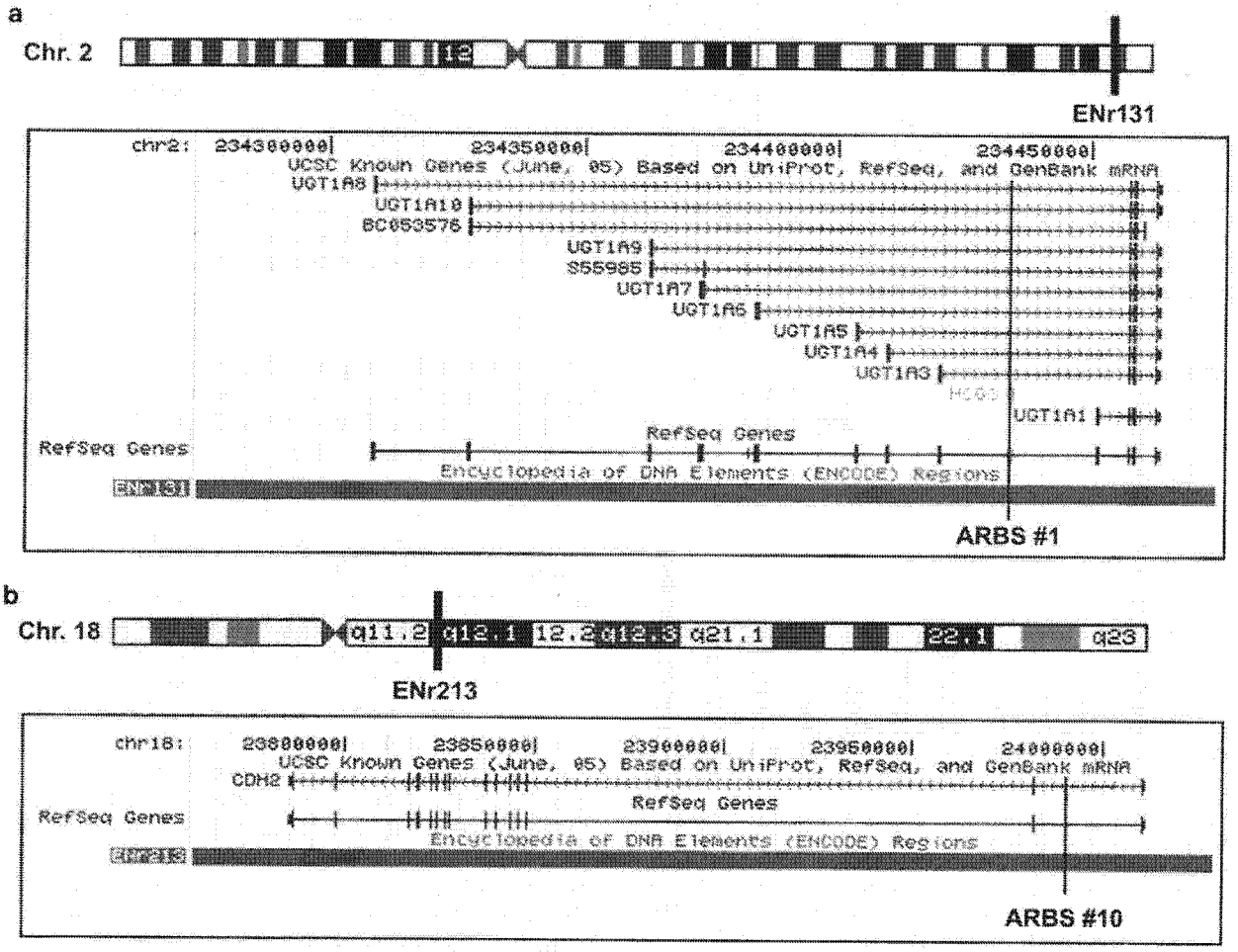

Figure 1 Identification of *in vivo* ARBSs in LNCaP cells on the encyclopedia of DNA elements (ENCODE) array by ChIP-chip analysis. (**a**) An expanded view of the UGT1A locus on the ENCODE region ENr131 from chromosome 2q37 is shown in its genuine 5′–3′ orientation. ARBS no. 1 is located on the 5′ upstream region of UGT1A1, or on intron 1 of other UGT1A isoforms. (**b**) An expanded view of the CDH2 on the ENCODE region ENr213 from chromosome 18q11.2 is shown in its genuine 3′–5′ orientation. ARBS no. 10 is located on intron 1 of CDH2.

confirmed that >10-fold enrichment of R1881-dependent AR binding was shown in the all regions involved in the defined 10 ARBSs, targeting the identified ARE sequences (Figure 2). Thus, with the cutoff value 1e-5, we validated that our ChIP-chip results did not include false positives.

Identification of androgen target genes adjacent to AR-binding sites

To identify novel androgen target genes by using ChIP-chip data, we examined the alteration of gene expression closest to the ARBSs in LNCaP cells in response to R1881 (Figure 3). Eight of 10 genes adjacent to the

Table 2 ARE sequences identified in the ARBSs detected by ENCODE chip

ARBS no.	*Gene*	*Distance from TSS (bp)*	*Position*	*Chromosome*	*Genomic location*	*Strand*	*ARE sequence*
1	UGT1A1	−17422	5′	2	234433576	+1	TGAACAttcTGTCCT
2	PGC	−399	5′	6	41823498	+1	GGAACAaatAGTTCT
3	SCAP2	−130006	5′	7	26807587	−1	AGAACCccaGGACCC
4	STEAP2	15941	INTRON 3	7	89501592	+1	GGAAAGaatTGTTCT
5	PFTK1	−2978	5′	7	89980382	−1	AGTAAGaagAGTTGC*
		−2947	5′	7	89980415	−1	ACAGCActcAGTACT*
6	TES	+106870	3′	7	115551402	+1	AAAACActcAGTTGC*
7	MET	+116298	INTRON 17	7	116022708	+1	TGCACAgtgTTTTAC*
8	KIAA0377	+27774	INTRON 27	15	41640739	−1	TAACCAtccTGTACC
9	CATSPER2	−11955	5′	15	41740290	−1	TAACCAtccTGTACC
10	CDH2	+20827	INTRON1	18	23990362	−1	GGTACAgaaTGTCAC
		+20805	INTRON1	18	23990384	−1	GGTACAgaaTGTCAC
		+20761	INTRON1	18	23990428	−1	GGTACAgcaTGTCAC
		+20739	INTRON1	18	23990450	−1	GGTACAgaaTGTCAC
		+20695	INTRON1	18	23990494	−1	GGTACAgcaTGTCAC
		+20673	INTRON1	18	23990516	−1	GGTACAgcaTGTCAC
		+20651	INTRON1	18	23990538	−1	GGTACAgaaTGTCAC
		+20629	INTRON1	18	23990560	−1	GGTACAgcaTGTCAC
		+20607	INTRON1	18	23990582	−1	GGTACAgcaTGTCAC
		+20541	INTRON1	18	23990648	−1	GGTACAgcaTGTCAC
		+20519	INTRON1	18	23990670	−1	GGTACAgcaTGTCAC
		+20651	INTRON1	18	23990538	−1	GGTACAgaaTGTCAC
		+20453	INTRON1	18	23990736	−1	GGTACAgaaTGTCAC

ARE sequences were primarily determined by a position-weighted matrix method TRANSFAC with the matrix conservation >75% (Matys *et al.*, 2003). If no ARE was predicted by the first criteria, alternative ARE sequences (indicated as *) were determined by a sequence analysis utility of JASPER with the relative profile score threshold >70% (Sandelin *et al.*, 2004). Genomic location of ARE sequence indicates the position of the center base.

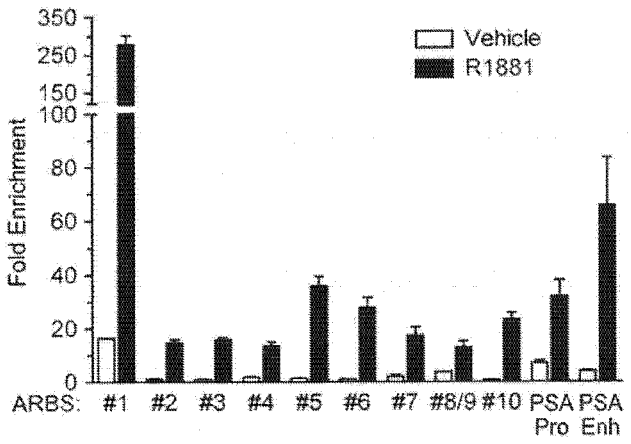

Figure 2 Validation of androgen-dependent AR enrichment by quantitative ChIP analysis on the identified ChIP-chip ARBSs in LNCaP cells. Hormone-deprivated cells were stimulated with R1881 (10 nM) or vehicle (0.1% ethanol) for 24 h. Cross-linked samples were immunoprecipitated with anti-AR antibody. The precipitated DNA fragments were subjected to qPCR. PCR primer sets were designed to include ARE sequences on individual ARBSs no. 1–no. 10. PCR products including ARBSs no. 8 and no. 9 were not distinguishable, as ARBSs no. 8 and no. 9 are located in genome duplication regions from the same origin. PSA promoter (PSA Pro) and enhancer (PSA Enh) regions including ARE sequences were used as positive controls. Data are fold enrichment compared with individual input non-enriched DNA (mean±s.d., $n=2$).

ARBSs exhibited a ligand-dependent increase in expression levels by >2-fold compared with a vehicle-treated control by 48 h after treatment. Among them, UGT1A1 and Pepsinogen C (PGC) levels elevated by >30- and >1000-fold, respectively, with 48-h R1881 treatment (Figure 3a and b). Residual two of 10 genes adjacent to the ARBSs, SCAP2 and MET, exhibited a ligand-dependent decrease in expression levels (Figure 3c).

UGT1A gene locus encodes nine distinct isoforms with unique exon 1 based on the difference of TSSs (Gong *et al.*, 2001). Individual first exons are determinants for the structure of N-terminal UGT1A isoforms, which are important for substrate specificity. We examined whether androgen regulated the transcription of distinct UGT1A isoforms in LNCaP cells. Interestingly, only the mRNA levels of UGT1A1 (Figure 3a) and UGT1A3 isoforms, which have TSSs most adjacent to ARBS no. 1, were significantly increased up to 48 h after R1881 (10 nM) stimulation (UGT1A3 mRNA levels: 1.2±0.1 fold at 12 h, 5.8±0.2 fold at 24 h and 11.2±0.2 fold at 48 h after treatment). On the contrary, the mRNA levels of other UGT1A isoforms were not basically altered or rather decreased during the time course (data not shown).

Taken together, our data suggest that AR could regulate transcription of genes adjacent to ARBSs.

ARBSs are associated with histone acetylation and facilitate recruitment of RNA PolII

We next examined whether these ARBS regions recruited components indicative of transcriptional activation. ChIP analyses for acetylated histone H3/H4 (AcH3/H4) and RNA PolII were performed on the 10 ARBS regions in LNCaP cells (Figure 4). The AREs in PSA promoter and enhancer were used as positive

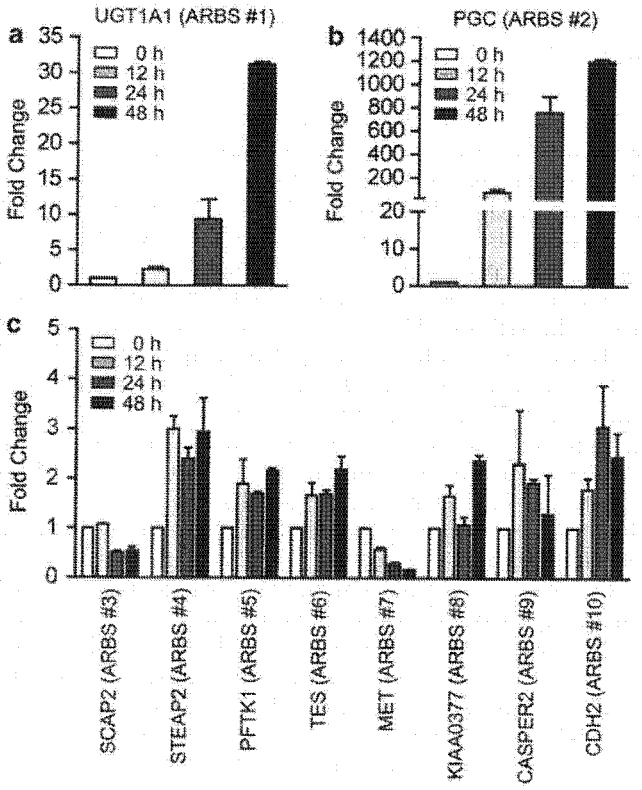

Figure 3 Androgen-dependent changes in expression of genes adjacent to the ChIP-chip ARBSs in LNCaP cells. Hormone-deprivated cells were stimulated with R1881 (10 nM) or vehicle for 12, 24 and 48 h. Quantitative RT–PCR analysis regarding the expression of 10 proximal genes close to ARBSs no. 1–no. 10 was performed using the reverse-transcribed cDNAs from the cells. Data are fold change compared with vehicle-treated cells at individual time point (mean ± s.d., $n = 2$). (**a**) UGT1A1 mRNA levels, (**b**) PGC mRNA levels and (**c**) mRNA levels of gene adjacent to ARBSs no. 3–no. 10.

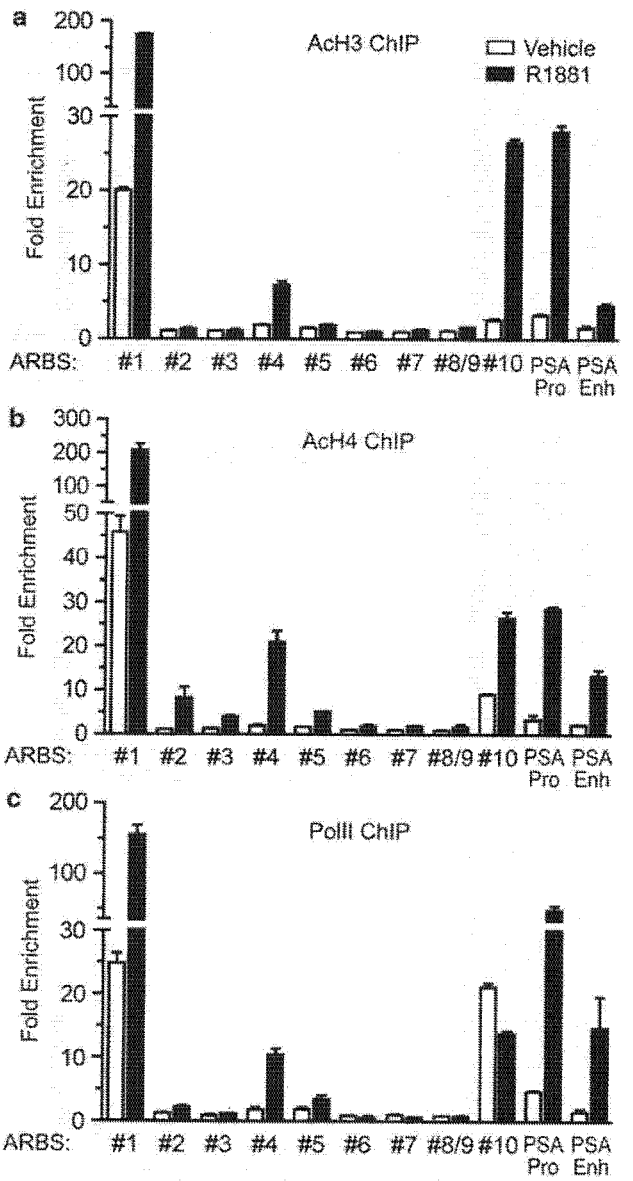

Figure 4 Some of the ChIP-chip ARBSs associate with histone acetylation and RNA PolII recruitment. Hormone-deprivated LNCaP cells were stimulated with R1881 (10 nM) or vehicle for 24 h. Cross-linked samples were immunoprecipitated with anti-acetylated H3/H4 (AcH3/AcH4) or anti-RNA PolII antibodies. The precipitated DNA fragments were subjected to qPCR. Identical primer sets were used as described in Figure 2. Data are fold enrichment compared to individual input non-enriched DNA (mean ± s.d., $n = 2$).

controls. Histone acetylation was remarkable by R1881 treatment in ARBSs no. 1 and no. 10, in the vicinity of UGT1A1 and CDH2, respectively. Ligand-dependent RNA PolII recruitment was also significant in ARBS no. 1. In ARBS no. 10, PolII binding was enriched at basal levels and further enhancement of PolII binding was not observed by ligand stimulation (Figure 4c). Moderate histone acetylation and PolII recruitment in response to R1881 were also observed in ARBS no. 4, which was located in intron 3 of STEAP2. In the rest of seven ARBSs, all associated with ligand-dependent AcH3/H4 binding by >2-fold and three of seven recruited PolII ligand dependently by >2-fold. Although most of the identified ARBSs were located rather distal from known genes, our data suggest that a significant number of the ARBSs in the genome physically function as distal transcriptional regulatory domains during transcription of the adjacent genes.

Functional recruitment of p160 co-activators at ARBSs

The p160 SRC family co-activators play scaffold roles in forming co-activator complex involved in nuclear receptor-mediated transcription (Shang and Brown, 2002). In AR-mediated transcription, the p160 co-activators are shown to be recruited to AR complex and to facilitate AR transactivation by their histone acetylase activity (Shang *et al*., 2002). To delineate the functional roles of endogenous p160 co-activators in AR-mediated transcription from the identified ARBSs no. 1–no. 10, we performed ChIP analysis using antibodies against SRC1, GRIP1 and AIB1 (Figure 5). It is notable that all of the p160 co-activators were recruited by >10-fold in the ARE from PSA promoter.

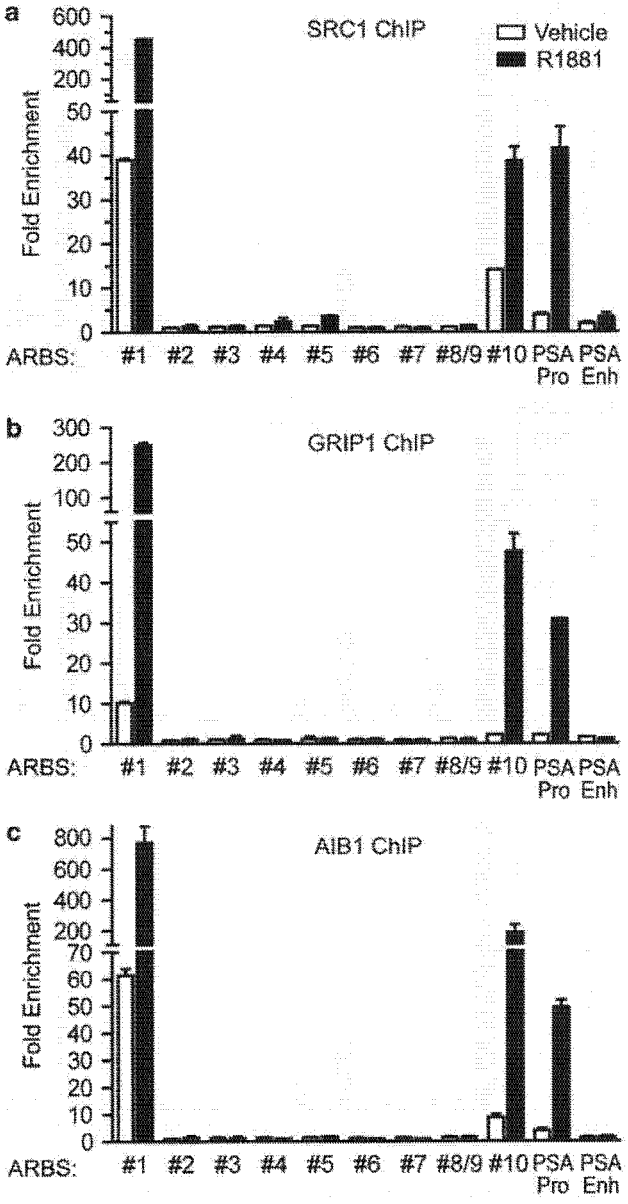

Figure 5 ChIP-chip ARBSs and p160 coactivator recruitment. Hormone-deprivated LNCaP cells were stimulated with R1881 (10 nM) or vehicle for 24 h. Cross-linked samples were immunoprecipitated with anti-SRC1, anti-GRIP1 or anti-AIB1 antibodies. The precipitated DNA fragments were subjected to qPCR. Identical primer sets were used as described in Figure 2. Data are fold enrichment compared to individual input non-enriched DNA (mean ± s.d., $n=2$).

Consistent with the result of histone acetylation and RNA PolII recruitment, all of the p160 co-activators were recruited by >10-fold upon R1881 stimulation compared with vehicle in ARBS no. 1, adjacent to UGT1A1. The second potent binding site for AcH3/H4 and RNA PolII, ARBS no. 10 adjacent to CDH2, recruited SRC1 by >2-fold and GRIP1 and AIB1 by >20-fold upon ligand stimulation. Among other ARBSs, ARBS no. 5 close to PFTK1 recruited SRC1 by >2-fold in response to ligand stimulation. The data show that some of the authentic ARBSs may play roles as enhancers that recruit various transcriptional regulators and co-activators.

Distal and intronic ARBSs function as transcriptional regulators in androgen-dependent transcription

To further assess the possibility that the distal or intronic ARBSs function as bona fide transcriptional regulators in androgen-dependent transcription, we performed promoter activity assay using luciferase reporter constructs including ARE sequences derived from the ARBSs. Using the genomic DNA of LNCaP cells as a template, we amplified fragments included ARE sequences in ARBSs no. 1 and no. 10, corresponding to the 5′ upstream region of UGT1A1 (~−17 kb) and intron 1 of CDH2 (Figure 6a and b). Note that ARBS no. 1 is also located in intron 1 of other UGT1A isoforms. The amplified fragments were ligated to a luciferase reporter plasmid pGL3-vector containing SV40 promoter. Regarding the 5′ upstream region of UGT1A1, we also generated a mutated construct including two substitutions at the positions −2C and +2G from the 3-bp spacer (UGT1A1 5′ Mut-Luc). Using LNCaP cells transfected with reporter constructs, the luciferase activities of UGT1A1 5′-Luc and CDH2 Int 1-Luc were increased ~5- and ~8-fold by R1881 treatment, respectively, whereas MMTV luciferase construct exhibited >100-fold activation in response to ligand stimulation (Figure 6c). UGT1A1 5′ Mut-Luc did not exhibit androgen-dependent transcriptional activation. These results suggest that a significant number of non-promoter ARBSs also play essential roles in AR-mediated gene transcription.

Moreover, UGT1A protein expression could be regulated by androgen (Figure 6d). LNCaP cells after 72-h hormone deprivation were stimulated with R1881 or vehicle and cell lysates were prepared after 24 or 48 h. Although the isoform specificity of UGT1A was not shown by the antibody that we used, overall amounts of UGT1A protein were increased in response to androgen.

Discussion

This study aimed to identify novel androgen target genes in prostate cancer LNCaP cells by performing ChIP-chip analysis, identifying *in vivo* ARBSs in the selected ENCODE genomic regions. This scanning successfully identified 10 bona fide *in vivo* ARBSs with a $P<1e\text{-}5$. Notably, all of the 10 ARBSs included ARE sequences as determined by the sequence analysis utilities based on TRANSFAC or JASPER transcription factor-binding profiles (Matys *et al*., 2003; Sandelin *et al*., 2004), and ChIP-PCR validation confirmed that those ARBSs had abilities to recruit AR ligand dependently. Our ChIP-chip approach is a powerful high-throughput method that can be applied to the whole genome-wide screen of ARBSs.

Efforts have been paid to identify transcription factor-binding motifs for years by searching a consensus-like sequence through *in silico* or *in vitro* studies in

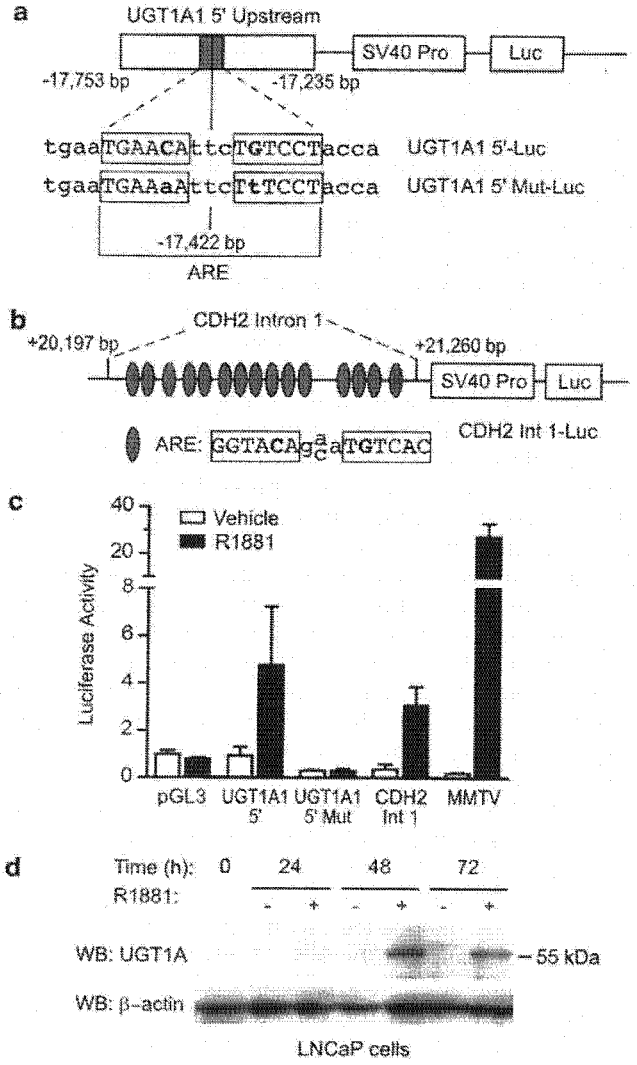

Figure 6 Transcriptional activity of UGT1A1 5′ upstream and CDH2 intronic ARBSs. (**a**) Construction of luciferase reporter plasmids containing UGT1A1 5′ upstream regions. A 519-bp fragment of the 5′ upstream of UGT1A1 (−17753/−17235 bp) amplified from LNCaP cells, and the 5′ fragment of UGT1A1 with mutation at the positions −2C and +2G from the 3-bp spacer in ARE sequences (ARBS no. 1) were cloned into pGL3 vector containing SV40 promoter (SV40 Pro), designated as UGT1A1 5′-Luc and UGT1A1 5′ Mut-Luc, respectively. (**b**) Construction of luciferase reporter plasmid containing CDH2 intron 1. A genomic fragment of CDH2 intron 1 (ARBS no. 10) derived from LNCaP cells (+20 197/+21 260 bp), containing 15 repeats of palindromic ARE sequences with the half-site GGTACA, was cloned into pGL3 vector (CDH2 Int 1-Luc). Note that the number of ARE sequences in LNCaP cells was larger than that in the genome database. (**c**) Androgen-stimulated luciferase activities of ARE sequences involved in UGT1A1 5′ upstream and CDH2 intron 1. LNCaP cells were stimulated with R1881 (10 nM) or vehicle 12 h after transfection with indicated plasmids together with a Renilla luciferase reporter gene, and incubated for another 24 h. Firefly luciferase activity was normalized to Renilla luciferase activity for each data set. MMTV luciferase reporter gene containing ARE sequences was used as a positive control. Data represent the mean ± s.d., $n = 3$. (**d**) Androgen-stimulated expression of UGT1A protein in LNCaP cells. Cells were stimulated with R1881 (10 nM) or vehicle for indicated times and whole cell lysates were separated by 8% SDS–PAGE. The PVDF membrane blotted with proteins were probed with anti-UGT1A or anti-β-actin antibodies.

the vicinity of transcription factor targets. In the days after the completion of the Human Genome Project, genomic DNA microarray has been developed and identification of *in vivo* binding targets of nuclear proteins is enabled by ChIP-chip analysis. For instance, in the ChIP-chip study for ERα on chromosomes 21 and 22, most of the binding sites were found at significant distances including several > 100 kb removed from TSSs (Carroll *et al.*, 2005). It has been suggested that these distal ERα-binding sites play an important role in estrogen-mediated regulation, as they could be physically associated with promoter-proximal regions. Similarly, in the present study, nine of 10 ARBSs that we identified by ChIP-chip were situated in the distal 5′ regions or intronic regions of known genes. Among those distal ARBSs, there were several sites that significantly recruited AcH3/H4 and RNA PolII (Figure 4). ARBSs no. 1 and no. 10, which are located at > 17 kb upstream of UGT1A1 and in intron 1 of CDH2, respectively, could also associate with the p160 co-activators in a ligand-dependent manner (Figure 5). Based on our findings and previous evidence, a number of ARBSs may be located in non-promoter regions of the genome, and often associated with histone acethylation and co-activator recruitment.

In the vicinity of ChIP-chip identified ARBSs, we found several genes upregulated or downregulated by androgen stimulation. PGC, whose location is close to ARBS no. 2, is an androgen-upregulated gene that has been reported previously as a prognostic factor in prostate cancer (Diaz *et al.*, 2002). It is an aspartyl protease and known as a protein involved in the digestion of proteins in the stomach. We identified a novel ARE sequence in ARBS no. 2, at −323 bp upstream of the TSS of PGC, indicating that ChIP-chip is particularly a powerful method to find out a novel transcription factor target regardless of the expression level of the target gene. As the ligand-dependent RNA PolII recruitment was not significant at ARBS no. 2 at the time we investigated (Figure 4c), other PGC regulatory region might be more important in the PolII activation; yet, ARBS no. 2 might play a role in the transcriptional regulation of PGC as the histone acetylation was at least promoted by ligand stimulation (Figure 4b). STEAP2, which includes ARBS no. 4 in intron 3, has been originally cloned as a STEAP homolog gene that encodes a transmembrane protein expressed in prostate cancer (Porkka *et al.*, 2002). Although androgen responsiveness of STEAP2 was not reported previously, our data showed that it was a novel androgen target gene with a genuine ARBS in intron 3, which was also associated with histone acetylation.

ENCODE region ENm001 corresponds to chromosome 7q31, which is known as a fragile site with frequent loss of heterozygosity in advanced prostate cancer (Kawana *et al.*, 2002). Among several genes at 7q31, TES (testis-derived transcript) has been shown as a candidate tumor suppressor gene in prostate cancer (Chene *et al.*, 2004). In the present studies, we showed that TES was an androgen-upregulated gene with a

genuine ARBS (ARBS no. 6) in its 3′ downstream region (−58 kb from the 3′-end). TES protein contains three conserved cysteine-rich zinc-binding motifs called LIM domains, suggesting that TES may play a role in protein-protein interaction and focal adhesion (Coutts *et al.*, 2003). Methylation of the CpG island at the 5′ end of TES is frequently occurred in ovarian cancer cells, and overexpression of TES in culture cells was shown to be growth-inhibitory (Tobias *et al.*, 2001). In contrast, we showed that MET was an androgen-downregulated gene with a novel ARBS (ARBS no. 7) in intron 17. MET encodes a receptor-like tyrosine kinase, c-met proto-oncogene product, which can be activated by hepatocyte growth factor as a receptor (Cooper *et al.*, 1984). It has been reported that MET expression is upregulated by androgen deprivation and MET appears to be preferentially expressed in androgen-insensitive, high-grade prostate cancer cells (Pisters *et al.*, 1995; Humphrey *et al.*, 1995). It has been also shown that overexpression of AR in prostate cancer PC3 cells leads to MET downregulation (Maeda *et al.*, 2006). Based on our bindings, we could propose that both TES and MET at 7q31 are regulated by AR in a way to exhibit negative feedback for prostate cancer progression, while the former is a tumor suppressor gene and the latter is a proto-oncogene.

CDH2 encodes one of the calcium-dependent cell adhesion molecules, N-cadherin. Whereas another calcium-dependent cell adhesion molecules, E-cadherin, is expressed in epithelial cells, N-cadherin is expressed in nerve system, skeletal muscle and mesenchymal cells (Jaggi *et al.*, 2006). Recent evidence suggests that changes in cadherin expression or cadherin switching play a critical role during progression of various tumors including breast cancer (Hazan *et al.*, 1997) and prostate cancer (Tomita *et al.*, 2000). Loss of E-cadherin expression was seen in high-grade breast and prostate cancers, whereas high levels of N-cadherin expression was shown in invasive tumors. Androgen responsiveness of N-cadherin has been shown in neurons, in spinal motoneurons (Monks and Watson, 2001). Our finding demonstrates that CDH2 is an androgen target gene with a novel cluster of ARE repeats in intron 1.

Interestingly, we found a polymorphism in terms of the number of ARE sequences in intron 1 of CDH2. In regard to the 15-bp complete palindromes consisting of the half-site GGTACA motif, LNCaP cells had 15 ARE sequences as shown in Figure 6b, whereas the genomic data published in NCBI Genome Browser contained 13 ARE sequences. We also found that 15 and 14 ARE repeats were contained in the CDH2 intron 1 derived from prostate cancer DU145 cells and benign prostate hyperplasia BPH1 cells, respectively (data not shown). By generating luciferase constructs including the ARE repeats of the CDH2 intron 1 derived from DU145 and BPH1, we showed that the luciferase activities of those constructs were also induced by R1881 stimulation, although the response of BPH1 was smaller than that of LNCaP or DU145. Thus, the polymorphism of ARE repeats in intron 1 may be related to the intensity of androgen responsiveness.

In this study, we demonstrated that UGT1A was a novel androgen-regulated gene with a functional ARE sequence in the 5′ upstream region of UGT1A1 or intron 1 in other UGT1A isoforms. Among several UGT1A isoforms, only UGT1A1 and UGT1A3 have been shown as androgen-upregulated genes in LNCaP cells. Considering out results, it is possible that the isoform-specific androgen responsiveness is linked with the closeness of the functional ARE in ARBS no. 1 to each isoform TSS. Thus, ChIP-chip would be also useful to dissect the isoform specificity of transcription factor target genes that encode a number of isoforms.

The human UGT1A locus spans ~200 kb on chromosome 2q37 and encodes nine UGT1A enzymes that play a crucial role in glucuronidation of xenobiotics and endobiotic substrates such as bilirubin (Chen *et al.*, 2005). UGT1A proteins are expressed in liver, whereas also expressed in extrahaptic tissues like urinary bladder and large intestine (Giuliani *et al.*, 2005). UGT1A gene products are generated by a strategy of exon sharing, resulting in divergent isoforms with a unique N-terminal domain and commonly shared C-terminal 245 amino acids. As UGT proteins are detoxifying enzymes, it is natural that this gene expression is regulated by xenobiotic receptor including pregnenolone X receptor and constitutive androstane receptor (Sugatani *et al.*, 2001; Xie *et al.*, 2003). Reduction of UGT1A expression is involved in the early phase of neoplastic transformation, such as in liver and biliary cancer, bladder cancer and colon cancer (Strassburg *et al.*, 1997; Giuliani *et al.*, 2005). In contrast, decrease in UGT1A1 expression seems to be associated with the reduced risk of endometrial cancer (Duguay *et al.*, 2004). UGT1A1 promoter polymorphism with an $A(TA)_7TAA$ element instead of a normal $A(TA)_6TAA$ element is known to decrease the level of gene expression, and it has been shown that there was a significant inverse association with the seven dinucleotide repeat allele and endometrial cancer risk (Duguay *et al.*, 2004). UGT proteins also glucuronidate steroid hormones, as the UGT1A enzymes showing specificity for estrogens, whereas androgens are substrates for another type of UGT family, UGT2B proteins (Lepine *et al.*, 2004). It has been recently shown that UGT2B15 isoform is an estrogen-regulated gene that is involved in the glucuronidation of androgens as well as estrogens (Harrington *et al.*, 2006). Similarly, there is a possibility that UGT1A1 and UGT1A3 play a role in glucuronidation of androgens as well as estrogens.

In summary, we performed ChIP-chip analysis for *in vivo* ARBSs in prostate cancer LNCaP cells, on the ENCODE regions in the human genome. A number of novel androgen target genes were identified adjacent to the ChIP-chip-based ARBSs. The present results show that ChIP-chip has an advantage over transcript-based microarray analysis, identifying a number of bona fide AR target genes regardless of their expression levels based on the data of functional ARBSs. The androgen target genes identified by the present study would play various important roles in the maintenance of prostate cancer, including detoxification, protein degradation,

cell motility/migration and tumor suppression/progression. Our study could be extended to the whole genome search of ARBSs in different cell systems using various ligands for the receptor. Identification of novel androgen target genes by ChIP-chip will reveal the whole entity of androgen signaling network, and will be applied to develop new clinical methods of prevention, diagnosis and treatment for prostate cancer.

Materials and methods

Reagents

Methyltrienolone 17β-hydroxy-17α-methyl-estra-4,9,11-trien-3-one (R1881) was purchased from NEN Life Science Products (Boston, MA, USA). Anti-AR (H-280), anti-SRC1 (M341), anti-GRIP1 (M343), anti-UGT1A (H-300) antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Anti-AcH3 and anti-AcH4 were from Upstate Biotechnology (Lake Placid, NY, USA). Anti-RNA PolII (8WG16) was from Covance (Berkeley, CA, USA). Anti-β-actin monoclonal antibody was from Sigma (St Louis, MO, USA). Anti-AIB1 antibody was generated from rabbit serum using a glutathione *S*-transferase fusion protein with amino acids 1320–1420 of human AIB1 protein as an epitope.

Cell culture

Human prostate cancer LNCaP cells were purchased from American Type Culture Collection (Rockville, MD, USA). Cells were maintained in RPMI 1640 supplemented with 4.5 g/dl glucose, 1 mM sodium pyruvate, 10 mM HEPES and 10% fetal bovine serum (FBS). Before hormone addition, cells were cultured for 2 days in phenol red-free RPMI 1640 with 5% dextran-charcoal stripped FBS (dcc-FBS) and 1 day in phenol red-free medium supplemented with 2.5% dcc-FBS.

Chromatin immunoprecipitation

ChIP assay and qPCR were performed as previously described (Horie-Inoue *et al*., 2004, 2006). LNCaP cells after 72-h hormone depletion were treated with 10 nM R1881 or vehicle (0.1% ethanol) for the indicated times. Cells were fixed in 1% formaldehyde for 5 min at room temperature. Chromatin was sheared to an average size of 500 bp by sonication using a Bioruptor ultrasonicator (Cosmo-Bio, Tokyo, Japan). Lysates were rotated at 4°C for overnight with specific antibodies. Salmon sperm DNA/protein A–agarose (Upstate Biotechnology, Lake Placid, NY, USA) was added and incubated for 2 h. Precipitated DNA was used as templates for qPCR using Applied Biosystems 7000 sequence detector (Foster City, CA, USA) based on SYBR Green I fluorescence. Genomic fragments containing ARE in the promoter and enhancer regions of PSA (−250/−39 bp and −4170/−3978 bp from the TSS, respectively) were used as positive controls for AR binding (Horie-Inoue *et al*., 2004). Sequences of PCR primers are described in Supplementary Table 1.

DNA amplification and microarray preparation

ChIP-enriched DNA was amplified by two-step IVT as described previously (Katou *et al*., 2006). Briefly, alkali phosphatase-treated ChIP DNA was incubated with terminal transferase for poly-dT tailing, annealed with T7-poly A primer (5′-GCATTAGCGGCCGCGAAATTAATACGACTCACTATAGGGAGAAAAAAAAAAAAAAAAA[C/T/G]-3′), and used as a template for second-strand cDNA synthesis. Using this template DNA, first IVT amplification was performed by T7 RNA polymerase (Ambion Inc., Austin, TX, USA). The first-strand cDNA was synthesized using the amplified cRNA as a template. Second-strand cDNA synthesis and IVT amplification were carried out again. Second amplified RNA was converted into double-strand cDNA with random primers, fragmented with DNase I and end labeled with biotin. Hybridization was performed on the Affymetrix GeneChIP ENCODE01 1.0 Arrays (Santa Clara, CA, USA) using 2 μg of ChIP-enriched and non-enriched input control DNA.

Analysis of microarray data

Array intensity data were analysed by the Affymetrix Tiling Analysis Software based on the algorithm by Cawley *et al*. (2004), and the results were mapped to genomic positions in human genome assembly hg 17 (NCBI Build 35) or in Affymetrix Integrated Genome Browser. In ENCODE01 1.0 Arrays, sets of one probe pair, a perfect matched (PM) probe and a mismatch probe (MM) both 25 bases long are tiled at an average resolution of 22 bp as measured from the central position of adjacent 25-mer oligos, creating an overlap of approximately 3 bp. The (PM-MM) intensity value was recorded for each probe pair as a new probe value, and the distribution of probe value was adjusted to equal across all samples by conducting quantile normalization on each duplicate arrays for two groups, including non-enriched genomic input DNA or ChIP-enriched DNAs by AR antibody. To determine whether a probe x is ChIP-enriched, Wilcoxon rank sum test was applied to rank all the probe pairs within a 550-bp sliding window from x by their $\log_2$(max(PM-MM),1) values for checking whether the sum of ranks of all probe pairs in the ChIP samples were significantly higher than that in the controls (a P-value cutoff of 1e-5). For each window, a signal ratio was also estimated by the Hedges–Lehmann method computing the median of folds enrichment among the probe sets within the window.

Reverse transcription-qPCR

Total RNA was extracted from hormone-treated or 0.1% ethanol-treated cells for indicated times using ISOGEN reagent (Nippon Gene, Tokyo, Japan). First strand cDNA was generated from RNase-free DNase I-treated total RNA by using SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) and oligo-dT_{20} primer. Androgen responsiveness was analysed by quantitative reverse transcription-PCR (RT-qPCR) using Applied Biosystems 7000 sequence detector based on SYBR Green I fluorescence. Primer design and PCR protocol were as previously described (Horie-Inoue *et al*., 2004, 2006). Sequences of PCR primers are described in Supplementary Table 2.

Sequence analysis

The sequences of human RefSeq transcripts (hg 17, NCBI build 35) were retrieved from UCSC genome browser (http://www.genome.ucsc.edu/) (Kent *et al*., 2002). The presence of ARE sequences in the genomic DNA of every ChIP-enriched region were determined by a position weighted matrix method TRANSFAC (Matys *et al*., 2003) with the matrix conservation >75%. If no ARE sequence was predicted by this criteria, the search was performed by a sequence analysis utility of JASPER, an open-access database for eukaryotic transcription factor binding profiles, with the relative profile score threshold >70% (Sandelin *et al*., 2004).

Luciferase assay

Luciferase reporter genes containing ARE sequences in ARBSs no. 1 and no. 10 identified by ChIP-chip were constructed by ligating the fragments derived from UDP-glucuronosyltransferase (UGT) 1A1 5′ upstream region (−17 753/−17 235 bp from the TSS) and cadherin-2 (CDH2) intron 1 region (+20 197/+21 260 bp from the TSS) into pGL3 vector (Promega, Madison, WI, USA) at the sites between *Mlu*I and *Xho*I, designated as UGT1A1 5′-Luc and CDH2 Int 1-Luc, respectively. A mutated UGT1A1 5′ region construct (UGT1A1 5′ Mut-Luc) was also generated, including the identical region of UGT1A1 5′-Luc except two substitutions of conserved C and G for A and T, respectively, at the 2-bp apart positions from the 3-bp spacer of ARE sequence. Mouse mammary tumor virus luciferase construct (MMTV-Luc) was used as a positive control for AR transcription activity (Ogawa *et al*., 1995). LNCaP cells were plated at a density of 10 000 cells/well in a 24-well culture plate and cultured for 3 days in phenol red-free RPMI 1640 with 5% dcc-FBS. Cells were transfected with plasmids using the transfection reagent FuGENE6 (Roche Applied Science, Indianapolis, IN, USA), then 12 h later treated with R1881 (10 nM) or vehicle (0.1% ethanol) for 24 h. Luciferase activity of cell lysate was determined by the Dual Luciferase Assay Kit (Promega, Madison, WI, USA). A renilla luciferase reporter Tk-PRL was co-transfected as a control for evaluating transfection efficiency. Data represent means±s.d. from triplicate sets.

Western blotting

Whole cell lysates were prepared using lysis buffer (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, 1% TritonX-100, 1.5 mM $MgCl_2$, 10 μg/ml aprotinin, 10 μg/ml leupeptin, 1 mM PMSF). Protein concentrations were analyzed using the BCA protein assay kit (Pierce Biotechnology, Rockford, IL, USA). Fifty microgram of proteins were resolved by 8% SDS–polyacrylamide gel electrophoresis and electroblotted onto Immobilon-P Transfer Membrane (Millipore, Billerica, MA, USA). Membranes were incubated with primary antibodies followed by incubation with secondary antibodies. Antibody–antigen complexes were detected using the Western Blotting Chemiluminescence Luminol Reagent (Santa Cruz Biotechnology, Santa Cruz, CA, USA).

Acknowledgements

We thank T Suzuki and R Nozawa for their technical assistance. This work was supported in part by grants-in-aid from the Ministry of Health, Labor and Welfare; from the Japan Society for the Promotion of Science; from The Promotion and Mutual Aid Corporation for Private Schools of Japan. This work was supported in part by a grant of the Genome Network Project from the Ministry of Education, Culture, Sports, Science and Technology.

References

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ *et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**: 169–181.

Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ *et al.* (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**: 33–43.

Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D *et al.* (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**: 499–509.

Chen CD, Welsbie DS, Tran C, Baek SH, Chen R, Vessella R *et al.* (2004). Molecular determinants of resistance to antiandrogen therapy. *Nat Med* **10**: 33–39.

Chen S, Beaton D, Nguyen N, Senekeo-Effenberger K, Brace-Sinnokrak E, Argikar U *et al.* (2005). Tissue-specific, inducible, and hormonal control of the human UDP-glucuronosyltransferase-1 (UGT1) locus. *J Biol Chem* **280**: 37547–37557.

Chene L, Giroud C, Desgrandchamps F, Boccon-Gibod L, Cussenot O, Berthon P *et al.* (2004). Extensive analysis of the 7q31 region in human prostate tumors supports TES as the best candidate tumor suppressor gene. *Int J Cancer* **111**: 798–804.

Cooper CS, Park M, Blair DG, Tainsky MA, Huebner K, Croce CM *et al.* (1984). Molecular cloning of a new transforming gene from a chemically transformed human cell line. *Nature* **311**: 29–33.

Coutts AS, MacKenzie E, Griffith E, Black DM. (2003). TES is a novel focal adhesion protein with a role in cell spreading. *J Cell Sci* **116**: 897–906.

Diaz M, Rodriguez JC, Sanchez J, Sanchez MT, Martin A, Merino AM *et al.* (2002). Clinical significance of pepsinogen C tumor expression in patients with stage D2 prostate carcinoma. *Int J Biol Markers* **17**: 125–129.

Duguay Y, McGrath M, Lepine J, Gagne JF, Hankinson SE, Colditz GA *et al.* (2004). The functional UGT1A1 promoter polymorphism decreases endometrial cancer risk. *Cancer Res* **64**: 1202–1207.

ENCODE Project Consortium, EP (2004). The ENCODE (ENCyclopedia of DNA Elements) Project. *Science* **306**: 636–640.

Giuliani L, Ciotti M, Stoppacciaro A, Pasquini A, Silvestri I, De Matteis A *et al.* (2005). UDP-glucuronosyltransferases 1A expression in human urinary bladder and colon cancer by immunohistochemistry. *Oncol Rep* **13**: 185–191.

Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK *et al.* (2001). Thirteen UDP glucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics* **11**: 357–368.

Grossmann ME, Huang H, Tindall DJ. (2001). Androgen receptor signaling in androgen-refractory prostate cancer. *J Natl Cancer Inst* **93**: 1687–1697.

Harrington WR, Sengupta S, Katzenellenbogen BS. (2006). Estrogen regulation of the glucuronidation enzyme UGT2B15 in estrogen receptor-positive breast cancer cells. *Endocrinology* **147**: 3843–3850.

Hazan RB, Kang L, Whooley BP, Borgen PI. (1997). N-cadherin promotes adhesion between invasive breast cancer cells and the stroma. *Cell Adhes Commun* **4**: 399–411.

Horie-Inoue K, Bono H, Okazaki Y, Inoue S. (2004). Identification and functional analysis of consensus androgen response elements in human prostate cancer cells. *Biochem Biophys Res Commun* **325**: 1312–1317.

Horie-Inoue K, Takayama K, Bono HU, Ouchi Y, Okazaki Y, Inoue S. (2006). Identification of novel steroid target genes through the combination of bioinformatics and functional analysis of hormone response elements. *Biochem Biophys Res Commun* **339**: 99–106.

Humphrey PA, Zhu X, Zarnegar R, Swanson PE, Ratliff TL, Vollmer RT *et al.* (1995). Hepatocyte growth factor and its

receptor (c-MET) in prostatic carcinoma. *Am J Pathol* **147**: 386–396.

Jaggi M, Nazemi T, Abrahams NA, Baker JJ, Galich A, Smith LM *et al.* (2006). N-cadherin switching occurs in high Gleason grade prostate cancer. *Prostate* **66**: 193–199.

Katou Y, Kaneshiro K, Aburatani H, Shirahige K. (2006). Genomic approach for the understanding of dynamic aspect of chromosome behavior. *Methods Enzymol* **409**: 389–410.

Kawana Y, Ichikawa T, Suzuki H, Ueda T, Komiya A, Ichikawa Y *et al.* (2002). Loss of heterozygosity at 7q31.1 and 12p13-12 in advanced prostate cancer. *Prostate* **53**: 60–64.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM *et al.* (2002). The human genome browser at UCSC. *Genome Res* **12**: 996–1006.

Laganière J, Deblois G, Lefebvre C, Bataille AR, Robert F, Giguère V. (2005). From the cover: location analysis of estrogen receptor alpha target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc Natl Acad Sci USA* **102**: 11651–11656.

Lepine J, Bernard O, Plante M, Tetu B, Pelletier G, Labrie F *et al.* (2004). Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and methoxyestrogen metabolites by human uridine diphosphoglucuronosyltransferases expressed in endometrium. *J Clin Endocrinol Metab* **89**: 5222–5232.

Maeda A, Nakashiro K, Hara S, Sasaki T, Miwa Y, Tanji N *et al.* (2006). Inactivation of AR activates HGF/c-Met system in human prostatic carcinoma cells. *Biochem Biophys Res Commun* **347**: 1158–1165.

Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R *et al.* (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* **31**: 374–378.

Monks DA, Watson NV. (2001). N-cadherin expression in motoneurons is directly regulated by androgens: a genetic mosaic analysis in rats. *Brain Res* **895**: 73–79.

Ogawa H, Inouye S, Tsuji FI, Yasuda K, Umesono K. (1995). Localization, trafficking, and temperature-dependence of the Aequorea green fluorescent protein in cultured vertebrate cells. *Proc Natl Acad Sci USA* **92**: 11899–11903.

Pisters LL, Troncoso P, Zhau HE, Li W, von Eschenbach AC, Chung LW. (1995). c-met Proto-oncogene expression in benign and malignant human prostate tissues. *J Urol* **154**: 293–298.

Porkka KP, Helenius MA, Visakorpi T. (2002). Cloning and characterization of a novel six-transmembrane protein STEAP2, expressed in normal and malignant prostate. *Lab Invest* **82**: 1573–1582.

Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res* **32**: D91–94.

Shang Y, Brown M. (2002). Molecular determinants for the tissue specificity of SERMs. *Science* **295**: 2465–2468.

Shang Y, Myers M, Brown M. (2002). Formation of the androgen receptor transcription complex. *Mol Cell* **9**: 601–610.

Strassburg CP, Manns MP, Tukey RH. (1997). Differential down-regulation of the UDP-glucuronosyltransferase 1A locus is an early event in human liver and biliary cancer. *Cancer Res* **57**: 2979–2985.

Sugatani J, Kojima H, Ueda A, Kakizaki S, Yoshinari K, Gong QH *et al.* (2001). The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1A1 gene and regulation by the nuclear receptor CAR. *Hepatology* **33**: 1232–1238.

Tobias ES, Hurlstone AF, MacKenzie E, McFarlane R, Black DM. (2001). The TES gene at 7q31.1 is methylated in tumours and encodes a novel growth-suppressing LIM domain protein. *Oncogene* **20**: 2844–2853.

Tomita K, van Bokhoven A, van Leenders GJ, Ruijter ET, Jansen CF, Bussemakers MJ *et al.* (2000). Cadherin switching in human prostate cancer progression. *Cancer Res* **60**: 3650–3654.

Wang Q, Carroll JS, Brown M. (2005). Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol Cell* **19**: 631–642.

Xie W, Yeuh MF, Radominska-Pandya A, Saini SP, Negishi Y, Bottroff BS *et al.* (2003). Control of steroid, heme, and carcinogen metabolism by nuclear pregnane X receptor and constitutive androstane receptor. *Proc Natl Acad Sci USA* **100**: 4150–4155.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc).

Nuclear cyclin B1 in human breast carcinoma as a potent prognostic factor

Takashi Suzuki,[1,5] Tomohiro Urano,[3] Yasuhiro Miki,[1] Takuya Moriya,[1] Jun-ichi Akahira,[1] Takanori Ishida,[2] Kuniko Horie,[4] Satoshi Inoue[3,4] and Hironobu Sasano[1]

Departments of [1]Pathology and [2]Surgery, Tohoku University School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi-ken, 980-8575; [3]Department of Geriatric Medicine, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655; [4]Research Center for Genomic Medicine, Saitama Medical University, 1397-1 Yamane, Hidaka, Saitama, 350-1241, Japan

(Received November 30, 2006/Revised December 14, 2006/Accepted December 24, 2006/Online publication March 16, 2007)

Cyclin B1 is translocated to the nucleus from the cytoplasm, and plays an essential role in cell proliferation through promotion of mitosis. Although overexpression of cyclin B1 was previously reported in breast carcinomas, the biological significance of the intracellular localization of cyclin B1 remains unclear. Therefore, in this study, we examined cyclin B1 immunoreactivity in 109 breast carcinomas, according to the intracellular localization, that is, nucleus, cytoplasm or total (nucleus or cytoplasm). Total cyclin B1 was detected in carcinoma cells in 42% of breast carcinomas examined, whereas nuclear and cytoplasmic cyclin B1 were positive in 17 and 35% of the cases, respectively. Total or cytoplasmic cyclin B1 were positively associated with histological grade, mitosis, Ki-67, p53, c-myc or 14-3-3σ, and inversely correlated with estrogen or progesterone receptor. Nuclear cyclin B1 was significantly associated with tumor size, lymph node metastasis, histological grade, mitosis, Ki-67 or polo-like kinase 1. Only nuclear cyclin B1 was significantly associated with adverse clinical outcome of the patients, and multivariate analyses of disease-free and overall survival demonstrated nuclear cyclin B1 as the independent marker. A similar tendency was detected in the patients receiving adjuvant therapy after surgery. These results suggest that an onocogenic role of overexpressed cyclin B1 is mainly mediated in nuclei of breast carcinoma cells, and the nuclear translocation is regulated by polo-like kinase 1 and 14-3-3σ. Nuclear cyclin B1-positive breast carcinoma is resistant to adjuvant therapy, and nuclear cyclin B1 immunoreactivity is a potent prognostic factor in breast carcinoma patients. (*Cancer Sci* 2007)

Breast cancer is one of the most common malignancies in women worldwide. Invasive breast cancer has been generally regarded as a disease that metastasizes in an early phase, and clinical outcome of breast carcinoma patients is markedly influenced not only by metastasis of the tumor but also by proliferation activity of the tumor.[1] In fact, a multitude of prognostic factors identified for breast cancer have been demonstrated to be directly or indirectly related to proliferation of breast carcinoma cells.

It is well-known that proliferation of carcinoma cells is closely associated with altered regulation of the cell cycle.[2] Cell cycle progression is mediated by activation of a highly conserved family of cyclin-dependent kinases (Cdk),[3] and activation of a Cdk requires binding to a specific regulatory subunit, named a cyclin. Among the cyclins, cyclin B1 plays an essential role as a mitotic cyclin in the entry of mitosis from G_2 phase.[4] Overexpression of cyclin B1 has been reported in various human tumors, and some of these studies demonstrated the clinical significance of cyclin B1 as a poor prognostic factor for some cancers,[5–7] including lymph node-negative breast carcinoma.[8]

Cyclin B1 is initially localized in the cytoplasm, and is translocated to the nucleus at the beginning of mitosis.[9] Nuclear translocation of cyclin B1 is considered very important to facilitate access of the cyclin B–Cdc2 (also named Cdk1) complex to its nuclear substrate and promote mitosis.[4] Therefore, it becomes very important to examine the intracellular localization of cyclin B1 in tumor tissues, in order to obtain a better understanding of the biological roles of cyclin B1.[10] Previously, Winters *et al.* reported that nuclear cyclin B1 immunoreactivity was significantly associated with reduced disease-free survival of breast carcinoma patients in a log-rank analysis.[11] However, no other information is available regarding the intracellular localization of cyclin B1 in breast carcinoma tissue, and the biological significance of cyclin B1 remains unclear at this juncture. Therefore, in the present study, we examined the intracellular immunolocalization of cyclin B1, and correlated these findings with various clinicopathological parameters of the patients, including their clinical outcome.

Materials and Methods

Patients and tissues. One hundred and nine specimens of invasive ductal carcinoma of the breast were obtained from female patients who underwent mastectomy from 1984 to 1987 at the Department of Surgery, Tohoku University Hospital, Sendai, Japan. Breast tissue specimens were obtained from patients with a mean age of 53.1 years (range 23–82 years). The patients did not receive chemotherapy, irradiation or hormonal therapy prior to surgery. Review of the charts revealed that 85 patients received adjuvant chemotherapy (mitomycin C, methotrexate and fluorouracil, $n = 80$; cyclophosphamide, doxorubicin and fluorouracil, $n = 3$; and cyclophosphamide, mitomycin C and fluorouracil, $n = 2$). Seventeen patients received radiation therapy, and 12 patients received tamoxifen therapy after the surgery. The mean follow-up time was 106 months (range 4–157 months). The histological grade and tubule formation of each specimen was evaluated according to the method of Elston and Ellis.[12] All specimens were fixed with 10% formalin and embedded in paraffin wax. Research protocols for this study were approved by the Ethics Committee at both Tohoku University School of Medicine.

Antibodies. A rabbit polyclonal antibody for cyclin B1 (H-433 [sc-752]) was purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA). This antibody was raised against a recombinant peptide corresponding to amino acids 1–433 representing full-length human cyclin B1. Monoclonal antibodies for estrogen receptor α (ER; ER1D5), progesterone receptor (PR; MAB429), Ki-67 (MIB1), p53 (DO7) and c-myc (1-6E10) were purchased from Immunotech (Marseille, France), Chemicon (Temecula, CA, USA), DAKO (Carpinteria, CA, USA), Novocastra Laboratories (Newcastle, UK) and Cambridge Research Biochemical (Cambridge, UK), respectively. Rabbit polyclonal antibodies for HER2 (A0485) and polo-like kinase 1 (PLK1; 06-813) were obtained

[5]To whom correspondence should be addressed.
E-mail: t-suzuki@patholo2.med.tohoku.ac.jp

doi: 10.1111/j.1349-7006.2007.00444.x

from DAKO and Upstate Biotechnology (Lake Placid, NY, USA), respectively. Goat polyclonal antibody for 14-3-3σ (C-14 [sc-7683]) was purchased from Santa Cruz Biotechnology.

Immunohistochemistry. A Histofine Kit (Nichirei, Tokyo, Japan), which uses the streptavidin–biotin amplification method was used in this study. Antigen retrieval was carried out by heating the slides in an autoclave at 120°C for 5 min in citric acid buffer (2 mM citric acid and 9 mM trisodium citrate dehydrate, pH 6.0) for cyclin B1, ER, PR, HER2, Ki-67 and p53 immunostaining, and antigen retrieval for PLK1 and 14-3-3σ immunostaining was done by heating the slides in a microwave oven for 15 min in the citric acid buffer. Dilutions of primary antibodies used in this study were as follows: cyclin B1, 1/500; ER, 1/50; PR, 1/30; HER2, 1/200; Ki-67, 1/50; p53, 1/200; c-myc, 1/600; PLK1, 1/1500; and 14-3-3σ, 1/1000. The antigen–antibody complex was visualized with 3,3′-diaminobenzidine (DAB) solution (1 mM DAB, 50 mM Tris-HCl buffer [pH 7.6], and 0.006% H_2O_2), and counterstained with hematoxylin. As a negative control, normal mouse, rabbit or goat IgG was used instead of the primary antibodies, and no immunoreactivity was detected in these sections (data not shown).

Scoring of immunoreactivity and statistical analysis. Immunoreactivity of cyclin B1 was detected in the nucleus and cytoplasm, and was evaluated according to a report by Winters *et al.* with some modifications.[11] Briefly, cyclin B1 immunoreactivity was evaluated in the nucleus, cytoplasm or total (nucleus or cytoplasm) in more than 1000 carcinoma cells for each case, and subsequently the percentage of immunoreactivity (i.e. the labeling index [LI]) was determined. ER, PR, Ki-67 and p53 immunoreactivity was detected in the nucleus, and the immunoreactivity was evaluated as a LI. Cases with cyclin B1, ER, PR or p53 LI of more than 10% were considered positive in this study, according to a report for ER.[13] Immunoreactivity for c-myc, PLK1 and 14-3-3σ was detected in the cytoplasm, and cases that had more than 10% of positive carcinoma cells were considered positive. HER2 immunoreactivity was evaluated according to a grading system proposed in HercepTest (DAKO), and moderately or strongly circumscribed membrane staining of HER2 in more than 10% of carcinoma cells was considered positive.

An association between cyclin B1 immunoreactivity and clinicopathological factors was evaluated using a correlation coefficient (r) and regression equation, Student's t-test, or a one-way ANOVA and Bonferroni test. Overall and disease-free survival curves were generated according to the Kaplan–Meier method and the statistical significance was calculated using the log-rank test. Univariate and multivariate analyses were evaluated by a proportional hazard model (COX) using PROC PHREG in SAS software.

Results

Immunolocalization of cyclin B1 in breast carcinoma tissues. Immunoreactivity for cyclin B1 was detected in the nucleus or cytoplasm of breast carcinoma cells (Fig. 1a,b), and the mean values of cyclin B1 LI in the 109 breast carcinoma tissues examined were 12.8% (range 0–56%) in total, 5.4% (range 0–18%) in the nucleus, and 10.1% (range 0–52%) in the cytoplasm. The number of cyclin B1-positive breast carcinomas (i.e. cyclin B1 LI of more than 10%) was 46 cases (42%) in total, 19 cases (17%) in the nucleus, and 38 cases (35%) in the cytoplasm, respectively. Immunoreactivity of cyclin B1 was also detected in some epithelial cells of morphologically normal mammary glands (Fig. 1c), but its LI was less than 1% in all of the intracellular components examined in this study.

Significant associations ($P < 0.0001$) were detected among cyclin B1 LI of the intracellular components, and their correlation coefficients were as follows: $r = 0.95$ (total vs cytoplasm), $r = 0.64$ (total vs nucleus), and $r = 0.51$ (nucleus vs cytoplasm).

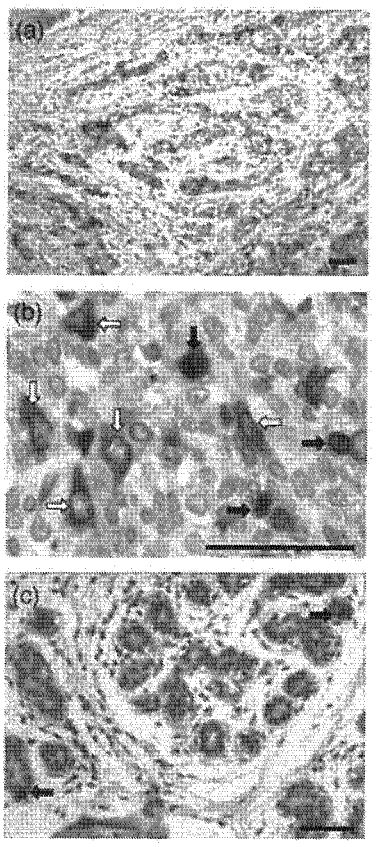

Fig. 1. Immunohistochemistry for cyclin B1 in the invasive ductal carcinoma. Cyclin B1 immunoreactivity was detected in the nucleus and/or cytoplasm of carcinoma cells: (a) lower magnification, (b) higher magnification. (b) Closed arrows represent nuclear cyclin B1 immunoreactivity, and open arrows show cytoplasmic cyclin B1 immunoreactivity. (c) In morphologically normal mammary glands, immunoreactivity for cyclin B1 was detected in some epithelial cells (arrows). Scale bar = 50 μm.

Association between cyclin B1 immunoreactivity and clinicopathological parameters in breast carcinoma. Associations between cyclin B1 immunoreactivity and clinicopathological parameters in 109 breast carcinomas are summarized in Table 1. Total cyclin B1 immunoreactivity was significantly associated with histological grade ($P = 0.001$), mitotic count ($P = 0.0001$) or Ki-67 LI ($P < 0.0001$), and inversely correlated with ER status ($P = 0.003$) or PR status ($P = 0.04$). There were no significant correlations between total cyclin B1 immunoreactivity and other clinicopathological parameters, such as patient age, menopausal status, clinical stage, tumor size, lymph node metastasis and HER2 status in this study.

However, immunoreactivity for nuclear cyclin B1 was positively associated with tumor size ($P = 0.01$), lymph node metastasis ($P = 0.003$), histological grade ($P = 0.003$), mitotic count ($P < 0.0001$) or Ki-67 LI ($P < 0.0001$), but no other significant association was detected. Cytoplasmic cyclin B1 immunoreactivity was positively associated with histological grade ($P = 0.001$), mitotic count ($P = 0.0001$) or Ki-67 LI ($P < 0.0001$), and an inverse association was detected between cytoplasmic cyclin B1 immunoreactivity and ER ($P = 0.003$) or PR status ($P = 0.01$), which was a similar tendency as that detected in the total cyclin B1 immunoreactivity.

Correlation between cyclin B1 immunoreactivity and its regulatory proteins in breast carcinoma. Previous studies have demonstrated that expression or intracellular localization of cyclin B1 is regulated by various proteins, including p53,[14,15] c-myc,[16]

doi: 10.1111/j.1349-7006.2007.00444.x