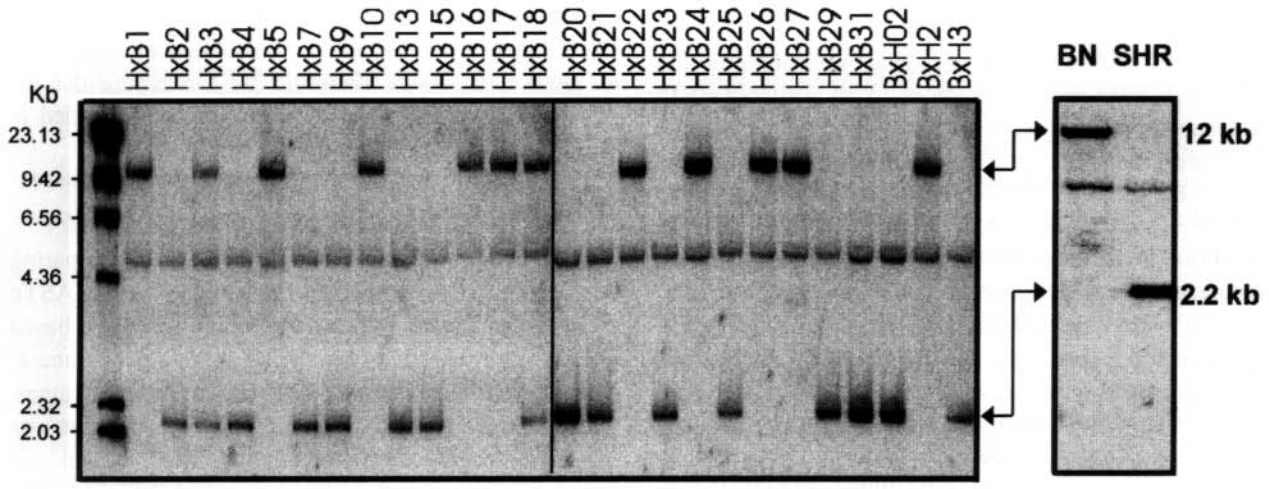

Fig. 2. Strain distribution pattern of RFLP. Strain distribution of *Hcarg* DNA polymorphism from progenitors SHR and BN.*lx* rats (right) as well as in a subset of 26 of their 33 RIS (left). Genomic DNA was extracted, digested with *Bgl*II and resolved on gel. The *Hcarg* radiolabelled probe revealed bands with B, 12 kb or H, 2.2 kb genotype from BN.*lx* and SHR, respectively (right). Strains heterozygous at the *Hcarg* locus were excluded from the SDP analysis.

Results

Chromosomal localization of HCaRG

From several restriction digestions of rat genomic DNA (Fig. 1), a *Bgl*II polymorphism was found near the *Hcarg* gene by the presence of a 12-kb genomic DNA band in BN.*lx* and 2.2 kb in SHR rats (Figs. 1 and 2 right). We used this polymorphism to determine the chromosomal localization of rat *Hcarg* in RIS derived from reciprocal crosses (BXH and HXB) between SHR and BN.*lx* progenitors (Fig. 2, left). The allelic distribution of the *Bgl*II polymorphism in 31 out of the 33 RIS homozygous for *HCaRG* was compared to the strain distribution of 475 markers in the same RIS (see Material and Methods). We found that the *Hcarg* locus co-segregated with D7Cebrp187s3/D7Cebr77s1 on chromosome 7 with recombination fraction of zero in these 31 RIS. Therefore, the *Hcarg* gene was assigned to a 4.4-cM long region between Mit3 and Mit4 of rat chromosome 7 (Fig. 3 left). A possible position of human homologous gene region, based on conserved linkage of rat chromosome 7, is chromosome 8q21-24. In a search of *Hcarg* homologous sequences in Genbank, homologies were found with three clones from chromosome 8 containing ZFP7. It was therefore possible to narrow down the localization of *HCaRG* to chromosome 8q24.3, confirming our initial prediction (Fig. 3 center). This region contains loci involved in several bone diseases, including osteopetrosis and multiple exostosis and several human neoplasms (McKusick et al., 1994; Knuutila et al., 1998). Comparison of human *HCaRG* to the available human genome sequence at NCBI permitted us to identify two other sequences homologous to *HCaRG* on chromosome 4 and on chromosome 6. These sequences are identical to each other, 641 a.a long (while complete human *HCaRG* is 672 a.a), and are 95% identical to human *HCaRG*. It is possible that these sequences code for a homologue of *HCaRG* or a member of the same family. However, it remains to be verified that the same sequence is found on two different chromosomes, since the human genome sequence is still in its draft form. A suggestive position of mouse *Hcarg* was determined based on the conservation of linkage between human and mouse. Mouse *Hcarg* localizes to chromosome 15 between 32 and 44 cM.

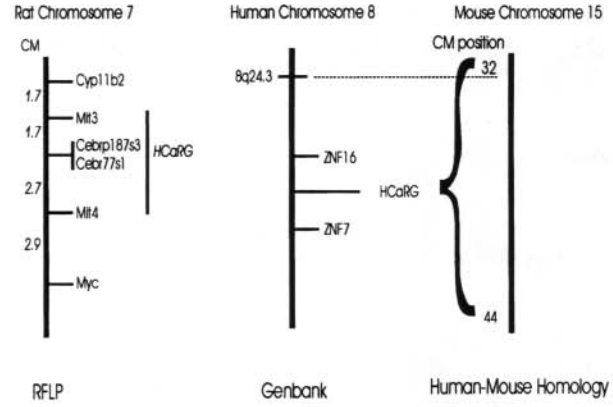

Fig. 3. Chromosomal localization of *HCaRG* in different species. A *Hcarg Bgl*II polymorphism was used as marker of genomic DNA from SHR or BN.*lx*. The allelic distribution of this polymorphism in RIS was compared to the strain distribution of 475 markers. *Hcarg* co-segregated with D7Cebrp187s3/D7Cebr77s1 of rat chromosome 7 in 31 out of 33 strains. cM represents the distance in centimorgans. In the middle, the position of the human homologous gene is depicted based on information obtained from Genbank. On the right side a suggestive position on mouse chromosome 15 is derived from the human–mouse homology database at NCBI.

Sequence comparison

In order to find similar sequences we compared the human HCaRG protein sequence to all available ESTs in Genbank using the tBLASTn programme. This programme dynamically translates all ESTs and compares the sequence to all possible open reading frames. We identified similar sequences in pig and cow. The pig sequence is 152 a.a and the cow sequence is 139 a.a while the complete sequence of HCaRG is 224 amino acids. Nucleotide comparison shows 79% identity between the human and the rat sequences with 93% identity between the rat and mouse HCaRG. Figure 4 shows the amino acid sequence alignment of rat, human, and mouse HCaRG cloned in our laboratory to the pig and cow sequence. At the amino acid level, 80% identity was found between the human and the rat and 95% between the rat and the mouse, 86% identity was calculated between the pig and the cow. The EF-hand domain is conserved in all five species with 8 out of the 10 most conserved a.a. The nuclear-receptor interaction domain is composed of the consensus sequence LxxLL (where x is any a.a) (Heery et al., 1997; Montminy, 1997). This motif is conserved in rat, mouse and human. However, in pig and cow an isoleucine replaces the last leucine (LxxLI). Some reports showed that the receptor interaction domain could be composed of isoleucine motifs: LxxII (Webb et al., 2000). An imperfect LxxLL sequence was recently identified in the chicken p160 coactivator molecule. This imperfect sequence, LxxIL, kept its capacity to interact with the nuclear receptor (Arai et al., 2001). Therefore, we can assume that the putative nuclear-receptor interaction domain identified in HCaRG is conserved in all five species.

Taxonomy report

The taxonomy report was generated by comparing human *HCaRG* to translated ESTs using the tBLASTn programme. Figure 5 shows the report generated based on the information in the NCBI taxonomy database. Most hits (identical sequences) are in eutheria (mammals) with the highest score in *Homo sapiens*. This was expected since we used the human HCaRG sequence to search ESTs, a score of 367 corresponds to an identity of 100%. This figure also shows hits in other organisms, but the score is below 100, and usually corresponds to a very short fragment being similar. No hits were observed in prokaryotes, using this search or with more precise searches (data not shown). *HCaRG* is therefore a gene mostly expressed in mammals.

Discussion

We identified a RFLP between BN.*lx* and SHR rats; this permitted us to localize rat *Hcarg* to chromosome 7. On the basis of conservation of synteny, we suggested the assignment of *HCaRG* on human chromosome

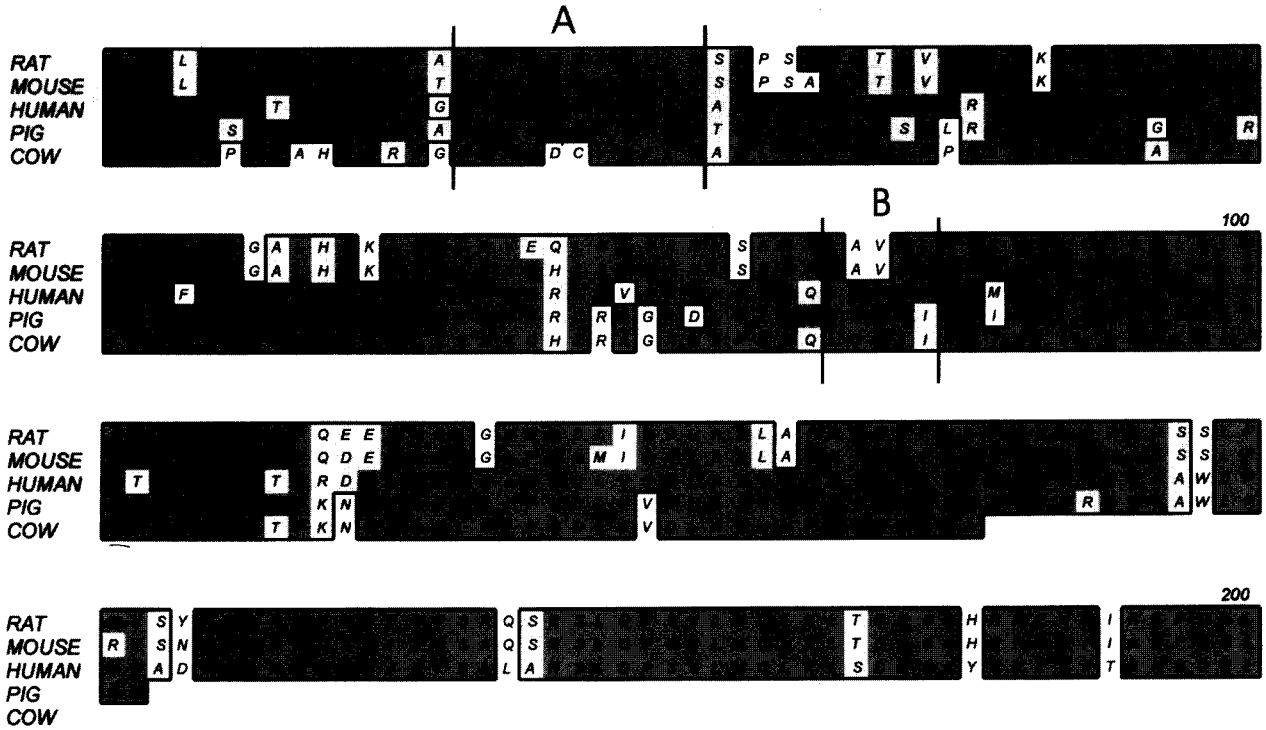

Fig. 4. Multiple species alignment. Similar sequences were found with the tBLASTn programme comparing the human HCaRG protein to translated ESTs. The sequences were then aligned using ClustalX1.81. Identities are shown in dark gray; similarities are in light gray. The putative EF-hand motif (A) and the nuclear receptor-binding domain (B) are conserved in all five species.