



Characterisation of the human Iroquois (IRX) homeobox genes: proposal for a systematic nomenclature

Lars Hansen ^{1,*}, Qingfa Wu ^{1,2}, Niels Tommerup ¹

¹) Wilhelm Johannsen Centre for Functional Genome Research, Department of Medical Genetics, Institute of Medical Biochemistry and Genetics, Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark; ²) Beijing Genomic Institute/Human Genome Centre, Institute of Genetics, Chinese Academy of Sciences, Beijing 100101, China; *) E-mail: larsh@medgen.ku.dk

Introduction

The group of Iroquois (IRX) genes is a small family of homeo box genes of the TALE group encoding DNA binding proteins and regulatory transcription factors important for pattern formation in all higher organisms including vertebrates, insects and nematodes.

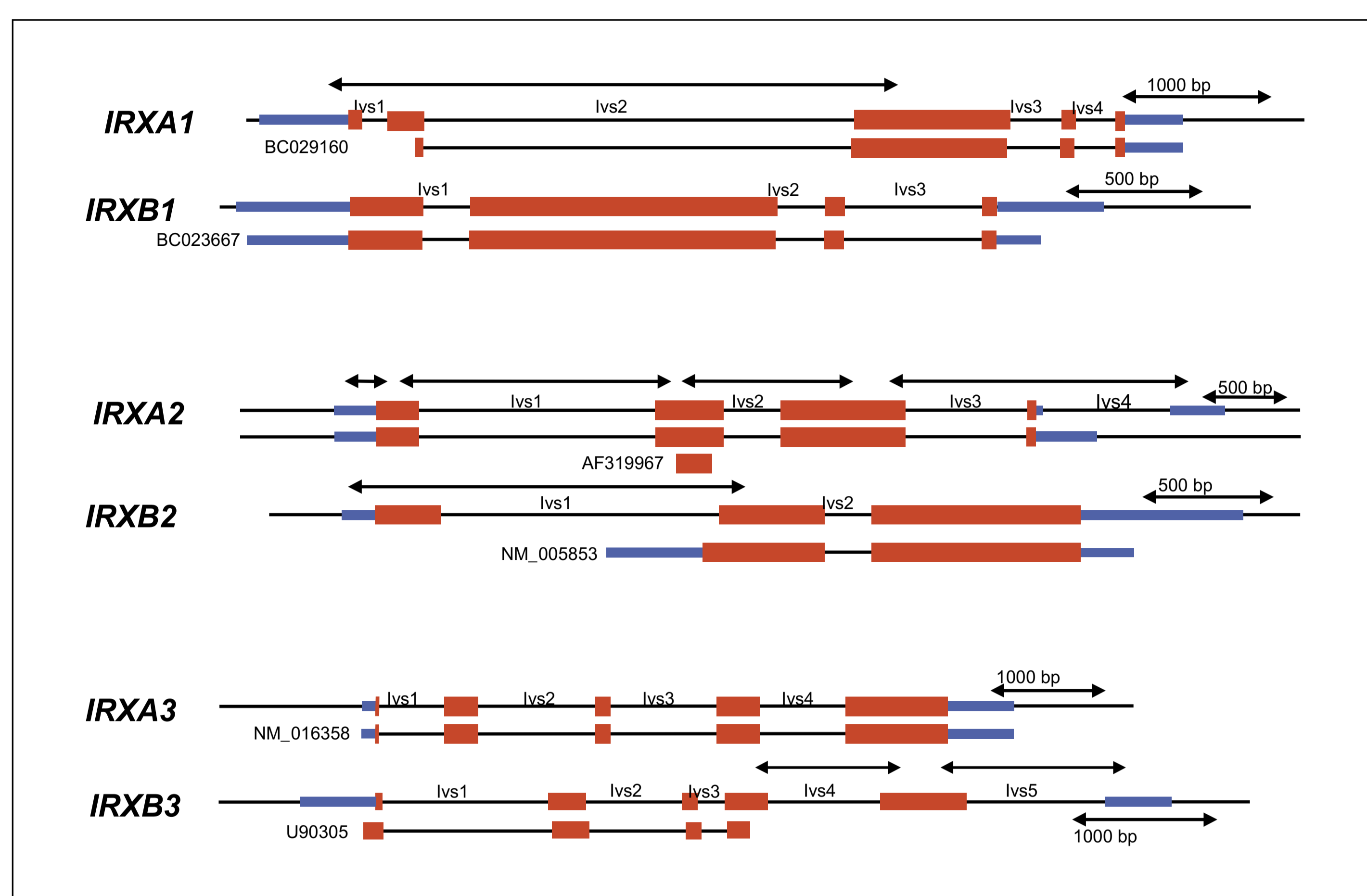


Figure 1. Orthologous IRX genes in human, mouse and rat share identical gene structures, paralogous genes in the same organism show less conservation. The gene structure for the three paralogous gene groups (IRXA1/IRXB1, IRXA2/IRXB2 and IRXA3/IRXB3) is presented. Full-length cDNA sequences are cloned by RT-PCR and gene and cDNA structure confirmed by sequencing. Expression profiles for all six genes have been studied in human adult and fetal tissues and coordinated expression is observed (results not shown). RT-PCR products - black arrows; exons - red boxes; UTR - blue boxes; introns and up- and down-stream regions - black lines; GenBank acc. no. represents cDNA, IMAGE and EST clones.

The IRX proteins contain one homeo domain, one IRO domain and frequently a poly amino acid domain. The evolutionary conservation of the IRX genes is reflected in their genomic organisation of three clustered genes, one cluster is present in *Drosophila* and two clusters in vertebrates with the six genes named *IRX1-IRX6* or *IRO1-IRO6*, depending on the organism.

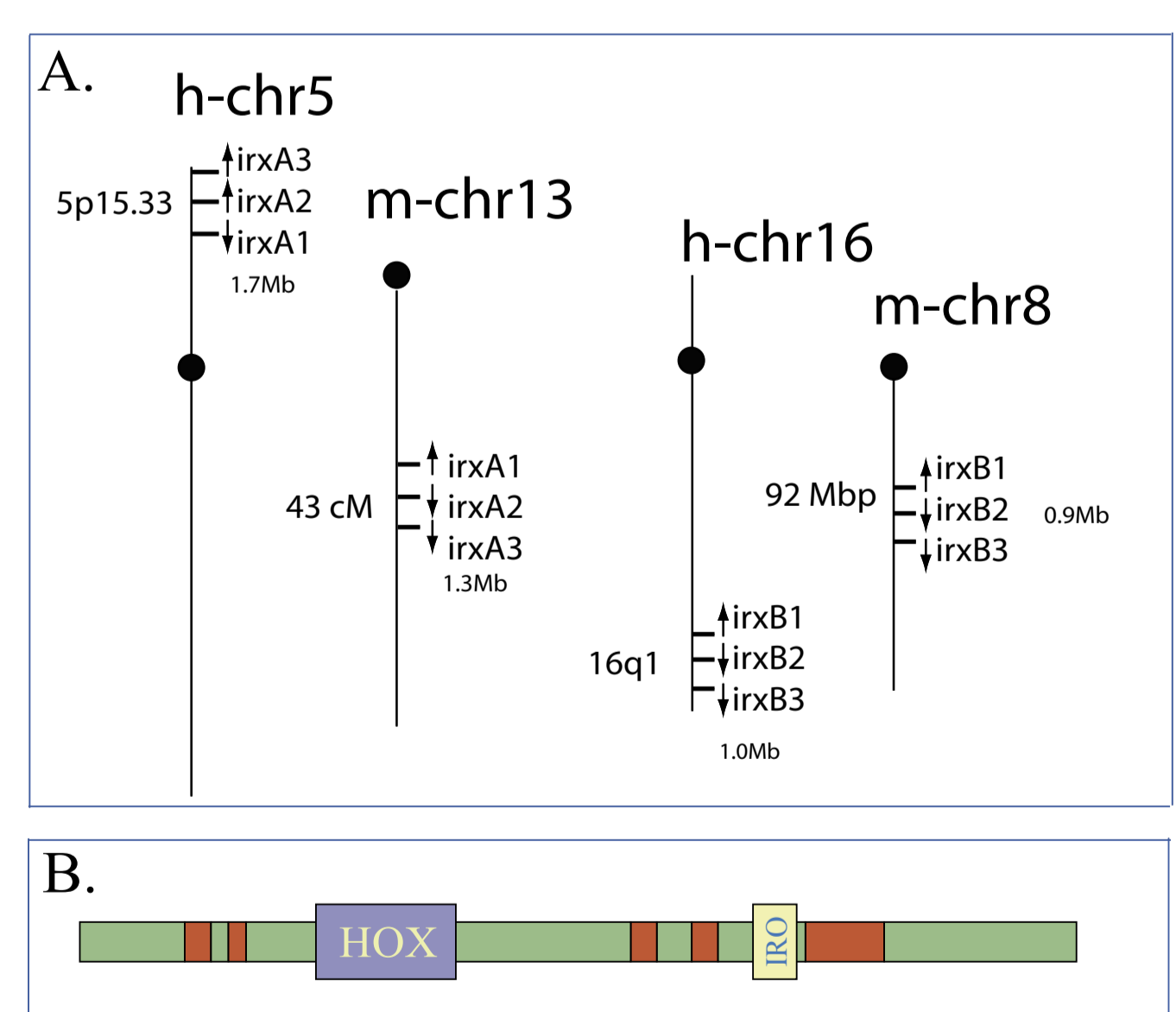


Figure 2. A: The IRX genes located as two clusters on different chromosomes suggest a systematic nomenclature for the genes; h- Homo sapiens, m- Mus musculus. **B:** The IRX-protein structure is highly conserved: homeo domain - blue, IRO domain - yellow, repeated regions - red.



C: Part of the six paralogous human IRX proteins: IRXA1/IRXB1, IRXA2/IRXB2 and IRXA3/IRXB3 are aligned; the homeo (yellow) and IRO (blue) domains and repeated regions (red) are highlighted.

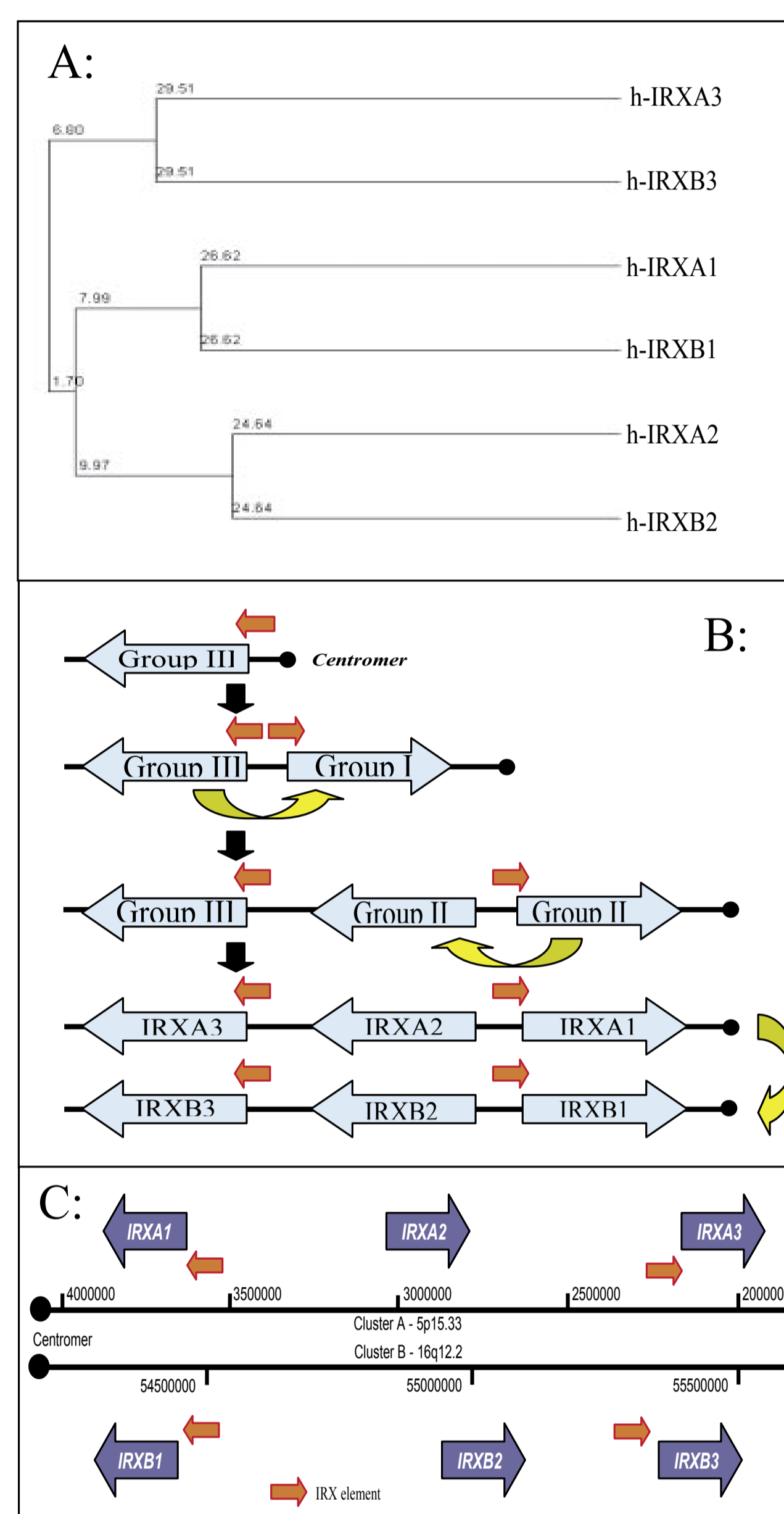


Figure 3. A: Phylogenetic tree for the six human IRX proteins. The tree shows the evolutionary relations between the paralogous proteins from cluster A and B, suggesting group III (IRXA3 or IRXB3) to be the oldest of the tree groups. B: Possible evolutionary events resulting in the origin of the cluster organisation, 1: duplication and inversion of an ancient IRX gene from group III (IRXA3 or IRXB3) resulting in group III and I, 2: duplication of a group I gene resulting in the group II gene, 3: finally a chromosome duplication leads to the two IRX clusters. Red arrows - IRX elements. C: Two inverted identical copies of the IRX element are found in each IRX cluster. The IRX element is a 250 bp intergenic DNA sequence without homology to any known regulatory element or gene, the element is conserved in the human, mouse and rat genome, its biological function is unknown.

Results

Two of the human genes are fully characterised, we have characterised the cluster organisation and the gene structure and expression pattern for all six genes and cloned and sequenced the missing parts of the remaining four human Iroquois homeo box genes. Comparative analysis of the Iroquois genes in ten different organisms (human, mouse, rat, Zebrafish, Fugu, *Ciona intestinalis*, chicken, *Xenopus*, *Drosophila*, *C.elegans*) demonstrate conservation of both genes and clusters, with intra-gene and inter-gene structural conserved features, that may have importance for the function and regulation of the Iroquois genes.

Table 1. Proposed symbols for IRX gene nomenclature are systematic and logic to the chromosome location of the IRX clusters and is analogue to the HOX nomenclature. The proposed nomenclature has been presented to the HUGO Gene Nomenclature Committee; the gene and mRNA sequences for all six human IRX genes are submitted to GenBank. Accession numbers refer to partial cDNA, IMAGE or EST clones; a full-length cDNA clone represents IRXA3/IRX4, IRX-2A/IRX5 represents an amino terminal truncated alternating sequence to our cloned IRXB2.

Iroquois cluster A	Centromere		
	Group I	Group II	Group III
Proposed Human symbol	IRXA1	IRXA2	IRXA3
Approved Human Symbol, aliases	IRX1, IRX-5	IRX2	IRX4
Human, chr.	5p15.33	5p15.33	5p15.33
GenBank ID ¹⁾	XM_098320, U90307	AF319967	NM_016358, AF124733, U90306
Mouse ortholog	<i>Irx1</i>	<i>Irx2</i>	<i>Irx3</i>
Approved Mouse Symbol, aliases	<i>Irx1</i>	<i>Irx2</i>	<i>Irx4</i>
Chromosome	13	13	13
GenBank ID ²⁾	AF165984, Y15002	AF295369, Y15000, AF165986	NM_016358, AF124733
Iroquois cluster B	Centromere		
	Group I	Group II	Group III
Proposed Human symbol	IRXB1	IRXB2	IRXB3
Approved Human Symbol, aliases	IRX-1, IRX3	IRX-2A, IRX5	IRX-3, IRX6, IRX7
Human, chr.	16q12.2	16q12.2	16q12.2
GenBank ID ²⁾	U90308, NM_005853		U90305
Mouse ortholog	<i>Irx1</i>	<i>Irx2</i>	<i>Irx3</i>
Approved Mouse Symbol, aliases	<i>Irx3</i>	<i>Irx5</i>	<i>Irx6</i>
Chromosome	8	8	8
GenBank ID ³⁾	NM_008393, Y15001	NM_018826, AF165985, AF230074, AJ271054	AJ271055

1) The orientations of the three genes in the clusters are depicted and the nomenclature is based on the orientation to the centromere. Cluster A is located on the short arm of chromosome 5 and cluster B is located on the long arm of chromosome 16.
2) The group nomenclature is based on homology structure of the six *Irx* genes, which can be divided into 3 paralog groups.
3) Acc. no. refers to Genbank accession numbers, only human *Irx3* and *Irx2* are represented by full-length clones.

Conclusion

The homeo box IRX genes are organised in clusters in mammals, the cluster topography is conserved and conserved inter-gene elements can be identified. Their function in embryo genesis is under investigation and identification of mutations in the genes will discover the character of genetic disorders with non-functioning IRX genes.

References

Peters T, Dildrop R, Ausmeier K, Ruther U. (2000) Organization of mouse Iroquois homeobox genes in two clusters suggests a conserved regulation and function in vertebrate development. *Genome Res* 10:1453-62
HUGO Gene Nomenclature Committee - <http://www.gene.ucl.ac.uk/nomencla>

Acknowledgement:

Wilhelm Johannsen Centre for Functional Genome Research is established by the Danish National Research Foundation