NOTE TO USERS

A study of the transcriptional regulation of the mouse *Indian Hedgehog* gene in ATDC5 and COS7 cells

by Anthony Ciarallo

Department of Human Genetics
McGill University, Montreal
February 2005

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Your file Votre référence
ISBN: 0-494-12416-4
Our file Notre référence
ISBN: 0-494-12416-4

Table of Contents

Abstract

The signaling peptide Indian Hedgehog (Ihh) is secreted by prehypertrophic chondrocytes of the growth plate, the growth center of long bones. It functions to negatively regulate differentiation and positively regulate proliferation of chondrocytes. Thus, Ihh ultimately controls the rate of bone growth.

The transcriptional regulation of the *Ihh* gene, however, remains uncharacterized. In order to study the gene's regulation, the genomic *Ihh* sequences from several species were aligned to identify conserved regions that may contain regulatory sites. Two putative Stat transcription factor binding sites were identified, one of which is conserved across all species studied while the other is rodent-specific.

In addition, an *in vitro* system was established to test the upstream region of the gene for transcriptional activity. ATDC5 chondrogenic cells were stably transfected with a plasmid containing 5kb of sequence located upstream of *Ihh* as well as a *luciferase* reporter gene. The presence of the *Ihh* sequence induced expression of the *luciferase* reporter 50 fold above expression from a control plasmid. COS7 and ATDC5 cells transiently transfected with similar *Ihh-luciferase* constructs resulted in unique induction patterns. Thus, the *Ihh* upstream genomic region contains sequences that regulate expression in a tissue-specific fashion.

Résumé

Le peptide de signalisation Indian hedgehog (Ihh) est sécrété par les chondrocytes prehypertrophiques de la plaque de croissance, le centre de croissance des os longs. Ihh agit de manière négative sur la differentiation et de manière positive sur la prolifération des chondrocytes. En conséquence, Ihh contrôle la croissance des os.

Toutefois, le contrôle de la transcription du gène *Ihh* demeure méconnu. Pour commencer l'étude du contrôle de l'expression de ce gène, les séquences génomiques *Ihh* de plusieurs èspèces ont été comparées de manière à identifier des séquences non-codantes conservées qui pourraient contenir des sites régulateurs. Deux sites présumés d'attachement des facteurs de transcription Stat ont été identifiés, l'un conservé chez toutes les èspèces étudiées et l'autre étant spécifique aux rongeurs.

De plus, un système *in vitro* a été mis en place pour tester l'activité transcriptionnelle de la région en amont du gène. Des cellules chondrogéniques de la lignée ATDC5 ont été transfectées de manière stable avec un plasmide contenant 5 kb de la séquence localisée en amont de *Ihh* ainsi qu'un gène marqueur *luciferase*. La présence de la sequence *Ihh* a induit l'expression du marqueur *luciférase* à un niveau 50 fois plus élevé que le plasmide témoin. Les cellules des lignées ATDC5 et COS7 transfectées de manière temporaires avec différents plasmides *Ihh-luciférase* équivalents ont démontrées des modes d'induction différents. Ceci indique que la région génomique en amont de *Ihh* contient des séquences contrôlant l'expression de manière différente selon les tissues.

Acknowledgements

I would like to thank Meg Desbarats for subcloning the p1879B construct, Munazzah Ambreen for preparing some of the p1879B deletion constructs, Barbara Miedzybrodzki for engineering the p1879B-gap2 deletion, Ken Dewar and Raman Minhas (Genome Quebec Innovation Centre, McGill University, Montreal) for providing the resources to align the human, chimp, mouse, rat, and chicken *IHH* sequences, Omar Akhouaryi for technical aid and consultation for several experimental manipulations, and Dr. René St-Arnaud for his advice and insight. Lastly, I would like to extend my gratitude to my supervisor, Dr. Benoit St-Jacques, for giving me the opportunity to work in his lab. During the last two years, he has provided me with the guidance that enabled me to develop my scientific skills and the support I needed to pursue my academic endeavours.

List of Abbreviations

ACH	Achondroplasia
ATDC5 d0	ATDC5 cells in an immature, pre-confluent, fibroblastic state (day 0)
ATDC5 d21	ATDC5 cells in a mature, confluent, chondrogenic state (day 21)
BMP	Bone Morphogenic Protein
BMPR	Bone Morphogenic Protein Receptor
CAT	Chloramphenicol acetyl transferase
cDNA	complementary DNA
ChIP	Chromatin Immunoprecipitation
COL	Collagen
DHH	Desert Hedgehog
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
EMSA	Electrophorectic Mobility Shift Assay
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
FGFR	Fibroblast Growth Factor Receptor
$Fgfr3^{ACH}$	FGFR3 with the activating achondroplasia mutation
HCH	Hypochondroplasia
Hh	Hedgehog
Hint	Hedgehog/Intein
IHH	Indian Hedgehog
MTf	Membrane-bound Transferrin-like Protein
NNPP	Berkeley Neural Network Promoter Prediction
PBS	Phosphate Buffered Saline
PTCH	Patched
PTHR-1	Parathyroid Hormone Receptor-1
PTHrP	Parathyroid Hormone related Peptide
RACE	Rapid Amplification of cDNA Ends
RLU	Relative Light Units
RNA	Ribonucleic Acid
SH2	Src Homology 2
SHH	Sonic Hedgehog
SMO	Smoothened
STAT	Signal Transducer and Activator of Transcription
TD	Thanatophoric Dysplasia
TF	Transcription Factor
TFII	Transcription Factor II (complex)
TFBS	Transcription Factor Binding Site
TGF-β	Transforming Growth Factor-β
TIS	Transcription Initiation Site
UCSC	University of California, Santa Cruz

List of Figures and Tables

1. Rationale and Objectives

The Hedgehog family of secretory proteins is critical for normal embryonic development. One of these factors, Indian Hedgehog (Ihh), is assumed to have multiple functions as it is expressed in a number of tissues, but is most recognized for its regulatory role in endochondral ossification. Ihh is a negative regulator of chondrocyte differentiation in the growth plate of developing long bones. The timing and duration of *Ihh* expression are thought to be critical determinants of long bone growth. Certain genes downstream of Ihh signalling have been identified and their functions elucidated, such as *Tgf-β* and *Pthrp*. In contrast, the regulation of *Ihh* gene expression remains undetermined. The identification of regulatory domains within the upstream region of the *IHH* gene would contribute to a better understanding of the regulation of bone development. In essence, the discovery of novel osteogenic regulatory pathways is particularly relevant to medical research exploring new frontiers of treatment for diseases affecting bone development.

This project focuses on analyzing the regions within and surrounding the *Ihh* gene for any transcriptional regulatory activity. The objectives of this project are (1) to map the location of the transcription initiation site of murine *Ihh*, (2) to establish an *in vitro* model for studying the regulation of *Ihh* expression, (3) to locate enhancer/silencer regions in the non-coding sequence upstream of the *Ihh* gene, and (4) to determine whether Fibroblast growth factor receptor-3 (Fgfr3) signalling modulates *Ihh* expression.

2. Introduction

2.1. Indian Hedgehog

Indian Hedgehog is a three exon gene (see figure 1) that codes for a protein that is involved in a variety of biological processes across numerous species. In mammals, it shares homology with two other proteins, Sonic and Desert, all of which belong to the Hedgehog (Hh) family of secreted proteins [1]. The mammalian *Hhs* are related to the segment polarity gene, *hedgehog,* which was originally identified in *Drosophila melanogaster* [2]. The gene's name, *hedgehog*, was based on the mutant phenotype which is manifested by the emergence of denticles on the larval cuticle [3]. Mammalian Hh proteins are key mediators of many fundamental processes in embryonic development, such as growth, patterning, and morphogenesis. Sonic Hedgehog (Shh) is postulated to regulate multiple facets of early embryogenesis [1], although Desert Hedgehog (Dhh) is mainly restricted to male germ line and gonadal development [4-6]. Ihh is best known for its regulatory role during bone development, however it is also expressed in various mammalian embryonic tissues and organs, including the visceral endoderm, cartilage, pancreas, uterus, retinal epithelium, liver, and kidney, the latter three of which maintain expression into adulthood [7-12]. However, apart from the gut, the expression domains of Ihh do not overlap with those of other Hh family members [13].

2.2. Hedgehog Signalling

Hh proteins are synthesized as 45kDa precursors and subjected to a series of post-translational modifications. Hhs undergo autoproteolytic cleavage that is mediated by

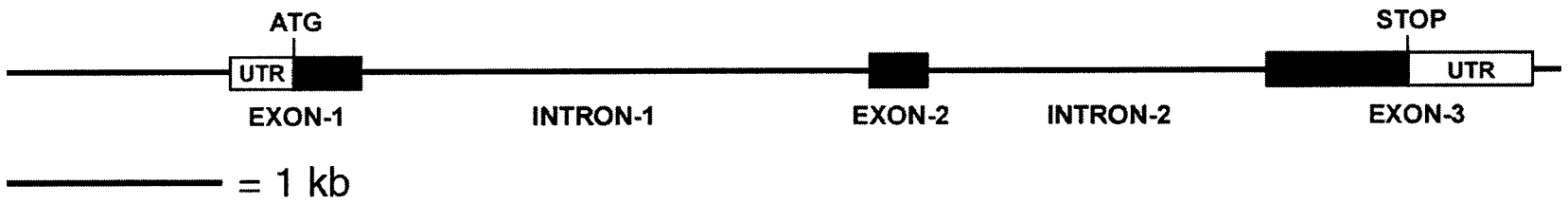


Figure 1: The *Indian Hedgehog* gene. The exon/intron structure is conserved across mammals, rodents, and avians.

their carboxy-terminal Hint (*H*edgehog/*int*ein) module [1, 14]. The Hint module consists of a class of autocatalytic splicing domains that share a homologous crystal structure [15]. The self-excision of the 25kDa Hint module generates a 20kDa Hh amino-terminal fragment with a cholesterol molecule covalently attached to its carboxyl terminus [16]. Prior to secretion, the Hh protein is also palmitoylated at the cysteine residue nearest to its amino terminus [17]. The amino fragment of the Hh precursor confers all known Hh protein activity [15], and the biochemical modifications are presumed to alter its range of signalling [18].

Hhs all bind a common receptor, Patched (Ptch), which also happens to be a transcriptional target of Hh signalling [19, 20]. In vertebrates, the Patched family consists of two transmembrane receptors, Ptch1 and Ptch2, each having 12 transmembrane domains and two large extracellular loops responsible for binding Hh proteins [21]. It has been demonstrated that Hh proteins can diffuse over considerable distances until they reach cells expressing Ptch, where they are then sequestered [22]. Furthermore, it was determined that Ptch transduces the Hh signal after observing that the deletion of the Ptch extracellular domain abolishes cell responsiveness to Hh [23]. However, signal transduction is initiated through another transmembrane protein called Smoothened (Smo). Smo is a constitutively active protein, complexed to and repressed by Ptch, that is essential for Hh signal transduction [24-26]. The binding of Hh to Ptch lifts the inhibition of Ptch over Smo resulting in Hh signal transduction [1]. The intracellular signal transduction cascade initiated by Smo is eventually executed by transcription factors of the Gli family [1]. Overexpression of Gli1, in particular, induces transcription of Hh target genes in the absence of Hh activity, making Gli1 the most likely activator of Hh target gene transcription [27, 28]. Similar to the Ptch family of receptors, some Gli family

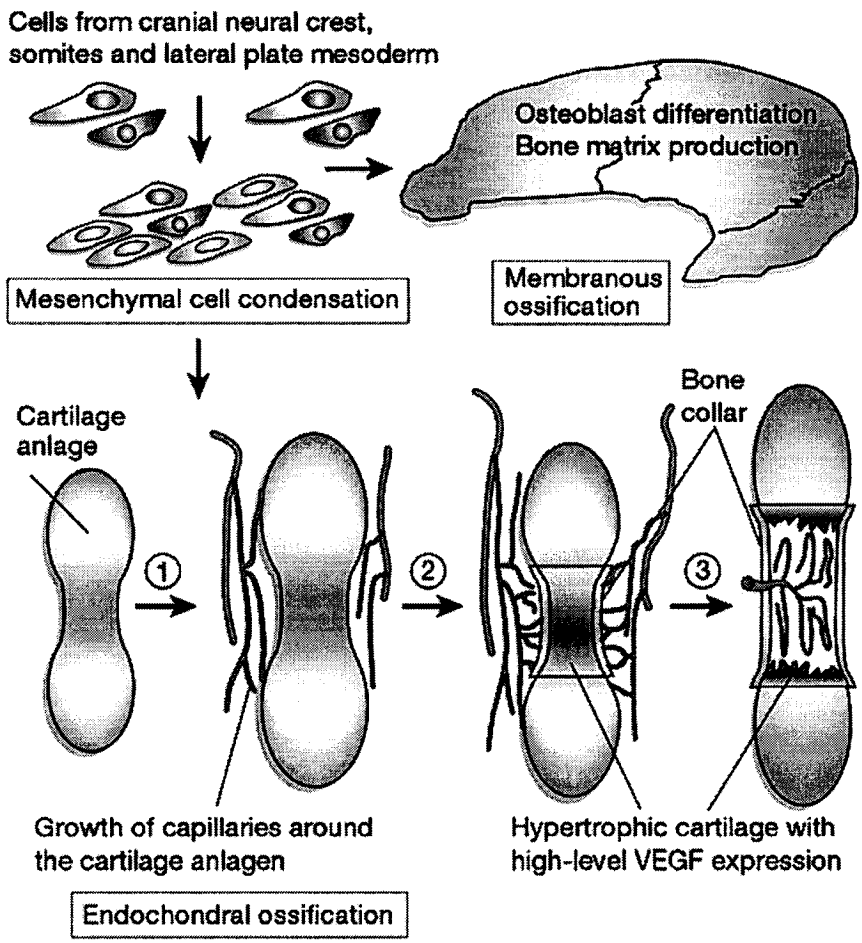
members are also transcriptional targets of Hh signalling. The expression of Gli1 and Gli3 is induced and repressed, respectively, in response to Hh activation [29, 30]. With the exception of some minor differences, the Hh signal transduction pathways are generally conserved between Sonic, Indian, and Desert Hedgehog.

2.3. The Formation and Development of Bone

Bone formation is an intricate and delicately balanced process that is coordinated by a network of signalling pathways, including the IHH pathway.

The process of skeletal formation is mediated by mesenchymal intercellular interactions and the selective expression of tissue-specific genes [31]. Initially, dispersed undifferentiated cells of the embryonic mesenchyme migrate towards predetermined sites of bone formation [31]. At each of these sites, mesenchymal cells aggregate to form a nodule, or "condensation," surrounded by an epithelial cell lining [32]. The condensation enlarges through cell proliferation and may give rise to one or more self-contained bone precursors. Once the nodule has surpassed a critical size threshold, the condensed mesenchymal cells situated at the core downregulate the expression of the genes controlling proliferation and upregulate those controlling differentiation to favour the onset of osteogenesis [31].

The two distinct modes of osteogenesis are *endochondral* and *intramembraneous* ossification, distinguishable by the respective presence or absence of an intermediary cartilaginous template (see figure 2, [33]). During endochondral ossification, mesenchymal cells at the centre of the condensed nodule differentiate into chondroblasts, a specific cell type that secretes cartilage matrix [34]. Eventually, all the cartilage is replaced by bone. Intramembraneous ossification differs from endochondral in that the



(Zelzer et al. 2003)

Figure 2: An overview of intramembranous and endochondral ossification. Both processes stem from a common mesenchymal origin and result in the formation of bone. However, endochondral ossification involves a cartilaginous intermediate whereas membranous ossification proceeds to bone formation directly from the mesenchymal cell condensation.

cartilage intermediary step is circumvented by the direct differentiation of condensed mesenchymal cells into osteoblasts, a specific cell type that secretes bone matrix. Both types of osteogenesis ultimately produce calcified bone, albeit with unique characteristics.

Intramembraneous ossification is responsible for forming the cranial vault, some facial bones, and parts of the mandible and clavicle [34]. Intramembraneous bone is initially formed during early skeletal development by osteoblasts as cancellous (spongy) bone through the sparse deposition of matrix [35]. Subsequent bone remodelling replaces the cancellous bone by compact (dense) bone through the successive deposition of new lamellae and trabeculae [35].

Endochondral ossification, on the other hand, is responsible for the formation of the axial and appendicular skeletons and the bones of the skull base [34]. Contrary to intramembraneous ossification, the formation of the latter skeletal elements is preceded by chondrogenesis. The first chondroblasts appear as a consequence of the Sox9-induced differentiation of mesenchymal cells within the avascular core of the condensed nodule [36]. Sox9 coincidentally promotes chondroblast expression of *Collagen* (*Col*) *2a1, 9a2*, and *11a2* and *Aggrecan*, genes that encode proteins required for the synthesis of the cartilaginous extracellular matrix [36]. At the periphery of the chondroblast nodule, the perichondrium forms from the stacking of flattened cells which serves to separate the cartilage from the surrounding mesenchyme and most likely contributes to appositional growth [35]. Chondroblast proliferation and the ensuing secretion of cartilaginous matrix are responsible for the general growth of the cartilage element. The secretion of copious amounts of matrix inevitably secludes each chondroblast in its own lacuna. This mature, immobilized cell type is called a chondrocyte. The majority of chondrocytes are destined

to undergo a differentiation process that is manifested by a transition in phenotype and gene expression.

Chondrocyte differentiation is a process that is characterized by the formation of a highly organised structure known as the growth plate, the border between cartilage and bone. The columnar alignment of differentiating chondrocytes in the cartilage matrix makes the visualization of the growth plate quite evident on a histological section of a developing bone (see figure 3). There are three identifiable zones that define the growth plate and each is indicative of the sequential state of chondrocyte differentiation, namely proliferation, maturation, and hypertrophy [37]. The proliferative zone is the most proximal region of the growth plate and is easily identified by chondrocytes that have acquired a flattened appearance due to self-compression within the matrix. The hypertrophic zone is the most distal region in the growth plate and consists of chondrocytes that have experienced a drastic increase in cellular volume. The proliferative and hypertrophic zones characteristically express *ColII* and *X*, respectively [37]. As they progress through the differentiation process, chondrocytes gradually downregulate *ColII* expression and upregulate *ColX* expression. The zone of maturation is located between the proliferative and hypertrophic zones, but the determination of its exact location is subjective because there are no obvious boundaries as to where it starts and ends [37]. This zone is commonly referred to as the zone of transition, or the pre-hypertrophic zone, since it consists of several chondrocytes initiating hypertrophic differentiation.

The transition in collagen secretion that occurs during chondrocyte hypertrophy alters the properties of the cartilage matrix to favour perichondrial vascularization and the onset of osteogenesis. This critical change in collagen secretion coincides with the

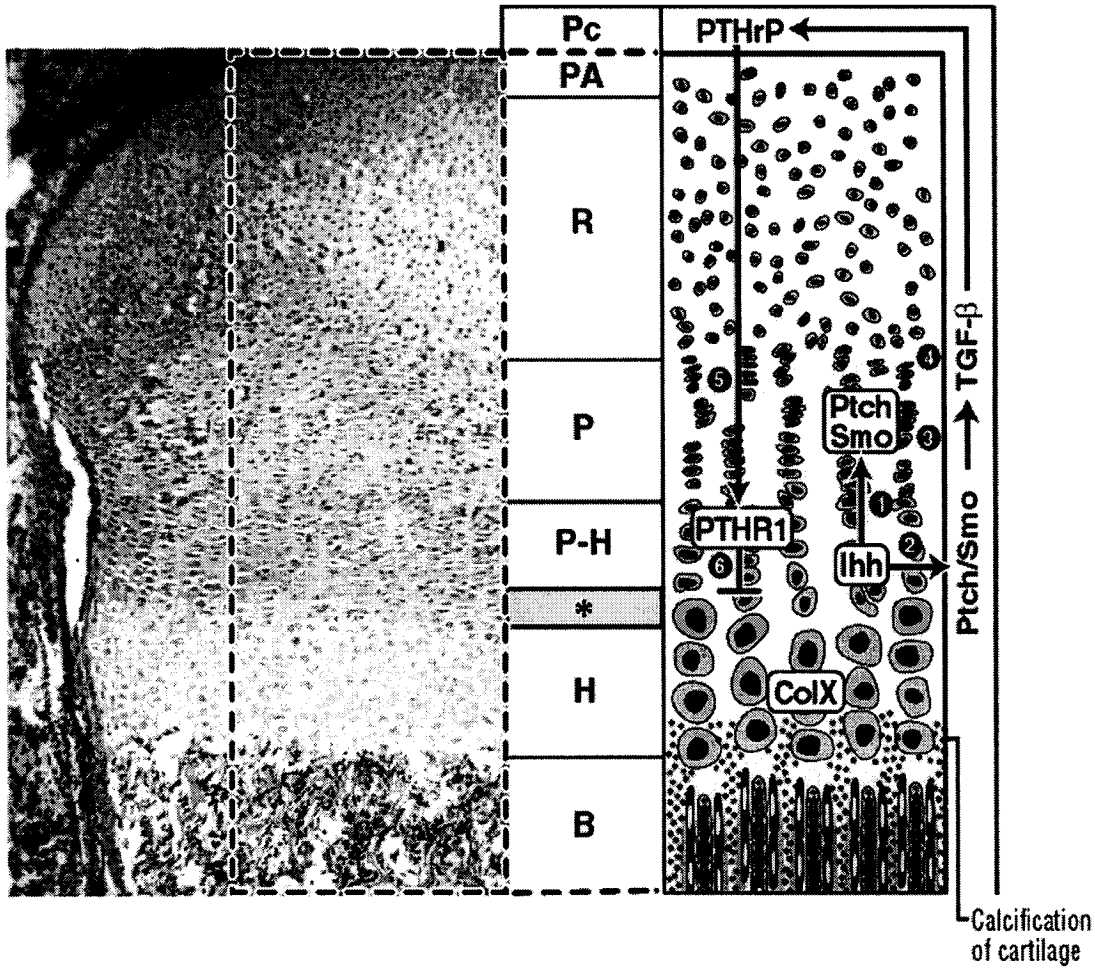


Figure 3: Visualization of the growth plate. *Left panel:* A longitudinal section through a long bone of an E15.5 mouse (St-Jacques, unpublished). Chondrocytes at different stages of differentiation are demarcated. *Right panel:* The corresponding diagrammatic representation of the sectioned long bone from the left panel. *Ihh* is expressed from prehypertrophic chondrocytes and signals to its receptor Ptch expressed by proliferating chondrocytes (1) and the perichondrium (2). The Ptch/Smo signal cascade in the perichondrium activates Tgf-β expression (3) which mediates the upregulation of Pthrp expression in the periarticular perichondrocytes (4). Pthrp diffuses across the growth plate and binds its receptor, Pthr-1, expressed by prehypertrophic chondrocytes (5). The activation of Pthr-1 receptors reduces the rate at which chondrocytes enter a hypertrophic state (6), thus regulating the rate of bone growth. ***Pc:*** Perichondrium, ***PA:*** Periarticular chondrocytes, ***R:*** Resting chondrocytes, ***P:*** Proliferating chondrocytes, ***P-H:*** Prehypertrophic chondrocytes, ***H:*** Hypertrophic chondrocytes, ***B:*** Bone. * (asterisk): represents the zone of transition from prehypertrophic to hypertrophic chondrocytes.

appearance of actively secreting osteoblasts along the vascularized perichondrium [35]. Consequently, the section of the perichondrium that surrounds the hypertrophic chondrocytes is replaced by a bone collar covered by the periosteum. The extracellular cartilage matrix secreted by hypertrophic chondrocytes becomes progressively calcified in preparation to serve as a scaffold for bone formation. Capillaries from the periosteum invade the hypertrophic zone and hypertrophic chondrocytes terminally differentiate and undergo apoptosis, leaving behind the calcified matrix [35]. Primary ossification is established as osteoblasts infiltrate the calcified cartilage. They synthesize bone matrix to replace the cartilage scaffold as it is being resorbed by osteoclasts. Osteoblasts and osteoclasts are functionally specific cells that work in concert to maintain bone integrity and heal bone microfractures by continually remodelling bone matrix [38]. During remodelling, existing bone matrix is resorbed by osteoclasts and subsequently restored by osteoblasts [38]. In fact, the average human skeleton is entirely remodelled throughout every decade of life [39].

Bone growth occurs both longitudinally and appositionally. Longitudinal growth is dependent on the activity at the growth plate, more specifically the expansion of the chondrocyte population through proliferation and the increase in intracellular volume attained during hypertrophy [35]. Despite its deceivingly static appearance, the growth plate is an undoubtedly dynamic structure whose size is determined by the rate at which growth plate chondrocytes initiate proliferation and undergo apoptosis [35]. Longitudinal bone growth ceases when all the chondrocytes in the resting zone, the region proximal to the growth plate that consists of chondrocytes proliferating at a basal rate, fully differentiate and are subsequently replaced by bone matrix. The peri-articular chondrocytes, located on the epiphysial surfaces, are the only chondrocytes never to

undergo differentiation as they are required for a protective, low-friction surface for joint articulation [37]. Contrary to longitudinal growth, appositional bone growth relies on the continual lamellar deposition of cortical bone by the osteoblasts in the periosteum [35]. The combination of all these components determines the size, shape, and mechanical properties of the bone.

2.4. Regulatory Signalling Pathways in the Growth Plate

Endochondral ossification relies on various signalling pathways to ensure the proper execution of chondrogenic and osteogenic programs, including their onset, rate, and duration. A particular gene of interest involved in the regulation of chondrogenesis is *Ihh*. Embryonic mice affected by the *Ihh*, *Parathyroid hormone-related peptide (Pthrp)*, or *Parathyroid hormone receptor-1 (Pthr-1)* null mutations similarly display accelerated chondrocyte differentiation, which contributes to their dwarfed phenotype [40-42]. These mutations have all led to the conclusion that *Ihh* and *Pthrp* are involved in a regulatory loop that is designed to negatively control the rate of chondrocyte hypertrophy during endochondral ossification. Conversely, overexpression of *Pthrp* and *Ihh* led to a delay in chondrocyte maturation and bone formation [43, 44]. *In situ* hybridization of *Ihh* in the growth plate revealed that this gene is expressed by chondrocytes once they reach a pre-hypertrophic state of differentiation [42, 43, 45]. As described in figure 3, Ihh, following its secretion, diffuses through the growth plate to the adjacent perichondrium where it binds and upregulates the expression of its receptor Ptch1 [43]. The activation of Ptch1 induces the expression of *Transforming Growth Factor-β (Tgf-β)* throughout the perichondrium which mediates the signal to periarticular chondrocytes to express *Pthrp* [41-43, 46]. Pthrp diffuses down the growth plate and binds its receptor, Pthr-1, located

on pre-hypertrophic chondrocytes that are slightly proximal to the zone of Ihh expression [41-43]. The activation of the Pthr-1 receptor is thought to slow the rate of chondrocyte hypertrophic differentiation, thereby regulating the size of the growth plate and its distance from the apex of the bone [41-43].

In addition to negatively regulating chondrocyte differentiation, *Ihh* is required for normal chondrocyte proliferation [42]. The rate of proliferation in growth plate chondrocytes is reduced by half in mice affected by the *Ihh* null mutation [42]. Conversely, the overexpression of a constitutively active *Smo* allele in proliferative chondrocytes yields an approximate two-fold increase in cell proliferation [47]. Ihh is believed to induce proliferation by acting directly on chondrocytes, although the precise mechanism has not been described.

2.5. Transcriptional Regulation in Eukaryotes

In growth plate cartilage, *Ihh* expression is limited to pre-hypertrophic chondrocytes. However, the regulation of *Ihh* expression has not been well characterized. Thus far, only a handful of genes are potentially implicated in the transcriptional regulation of *Ihh*, namely *Fibroblast growth factor receptor 3 (Fgfr3)*, certain *Bone morphogenic proteins (Bmps)*, and *Runx2/Cbfa-1* [48-50].

Eukaryotic gene expression is regulated at multiple stages to ensure that the necessary gene products are present at the appropriate time and location. Transcription is one of the processes that collectively regulate gene expression. Comparably, gene transcription is also subject to several regulatory mechanisms.

The transcription of a eukaryotic gene is most dependent on the critical regulatory elements within the promoter region, such as the TATA box which is about 30 base pairs

upstream from the transcription initiation site, the CAAT box, and sometimes additional elements further upstream [51-53]. Essentially, the promoter functions to bind RNA polymerase and coordinates gene transcription, although it requires transcription factors to bind efficiently [54]. Of the three RNA polymerases present in eukaryotic cells, it is the second type, RNA Polymerase II, that transcribes messenger RNA precursors [55]. The transcription factor II (TFII) complex, comprised of at least six nuclear proteins, mediates the interaction between the RNA polymerase and the TATA box and coordinates the proper DNA configuration, energy supply, and polymerase releasing activity needed for the initiation of transcription [56-59]. The associated factors in the complex also serve as co-activators, bridging enhancer-bound proteins to the transcription complex through protein–protein interactions [60].

Not surprisingly, there exists a subclass of genes whose promoter regions do not conform to the classical TATA sequence. In the latter situation, a protein that recognizes a GC-rich sequence binds to the promoter region; for instance, a generic promoter-binding protein such as Sp1. Once bound to the promoter, such a protein enables the binding of the transcription factor II complex [61]. The TFII complex subsequently instigates the cascade of factors that will form the transcription initiation complex and bind RNA polymerase II to the promoter region [62, 63].

Tissue specificity and timing of expression are inherently controlled by *cis*-regulatory elements that are dispersed within genes and their flanking regions. These elements exert their effects through transcription factors which influence interactions of the transcriptional machinery with the gene's promoter. In general, transcription factors are *trans*-regulatory proteins or steroids that bind the canonical DNA sequences of regulatory elements. In addition to having a sequence-specific DNA-binding domain,

most transcription factors also contain a *trans*-activating domain that interacts with neighbouring transcription factors, including those involved in binding RNA polymerase [64]. Furthermore, transcription factors can modify the physical structure of the DNA strand in order to reposition remote regulatory elements within functional range of the promoter region [65]. Thus, transcription factors modulate the expression of specific target genes by initially binding to the gene's enhancer, silencer, or promoter elements and subsequently interacting with the transcription initiation complex at the promoter. These regulatory interactions can be extremely complex, involving dozens of factors and taking place over extended regions of DNA. Extensive experimental analysis has elucidated the logic of this complex regulatory circuitry for only a handful of developmentally important genes [66, 67].

Genes that are frequently transcribed are loosely packed into *euchromatin*, permitting the factors required for transcription to access target sequences [68]. Conversely, the chromosomal segments that are seldom or never transcribed are tightly packed into *heterochromatin* [69]. Nevertheless, chromatin is a dynamic structure and it can be altered through epigenetic modifications. For example, the core histone proteins that constitute a nucleosome are susceptible to acetylation [70, 71]. The acetyl group destabilizes the nucleosome-DNA interaction, rendering genetic regulatory elements more accessible for transcription factors [70]. Thus, histone acetylation produces favourable conditions for gene transcription [71].

Gene transcription is also subject to epigenetic control via DNA methylation. The methyltransferase enzyme covalently modifies DNA through the addition of a methyl group to a cytosine residue within a 5'- CG – 3' dinucleotide, commonly known as a CpG site. Moreover, a genomic region that contains a dense population of CpG sites is

referred to as a CpG island. The significance of CpG sites arises from the correlation between DNA methylation and gene repression [72]. Promoters commonly consist of one or more CpG islands, however most remain unmethylated in eukaryotes [73]. Nevertheless, the CpG sites in the promoter region of a gene are believed to regulate gene activity [73]. In fact, DNA methylation is the basis for genomic imprinting. The methyl group possibly prevents the binding of transcription factors to target sequences, though the exact mechanism is unknown.

In summary, the transcription of eukaryotic genes is coordinated by an array of regulatory processes that act simultaneously to determine the appropriate amount of transcript, if any, that is required for each cell throughout an entire organism. Consequently, a regulatory imbalance may lead to catastrophic repercussions. However, such an imbalance may also be informative. For this reason, research models are engineered to simulate disruptions in the regulatory mechanisms of transcription for the purpose of identifying novel elements and factors that contribute to the control of gene expression.

2.6. Bone Morphogenic Proteins

IHH expression is suspected to depend on Bone morphogenic protein (Bmp) signalling. Bmps are a conserved sub-group of the Transforming growth factor-β (Tgf-β) superfamily of secreted proteins that regulate diverse embryonic patterning events [74], but were originally identified to induce ectopic bone formation [75]. Null mutations of Bmps and Bmp receptors (Bmprs) in mice revealed little information concerning the role of these proteins other than that they result in lethal or very mild phenotypes [76, 77]. Despite this shortcoming, Bmps were subsequently shown to be involved in the

regulation of the mesoderm and apical ectodermal ridge formation, mesenchymal condensation, endochondral ossification, and a number of non-osteogenic developmental processes [78-81]. More specifically, Bmps have been identified as regulators of cell growth, differentiation, apoptosis, cell lineage determination, patterning, and morphogenesis [74]. Investigators have demonstrated using mouse limb explants, either supplemented with Bmps or subjected to *Bmpr* misexpression, that Bmp signalling stimulates cartilage formation [76, 77, 81, 82], possibly through recruitment of undifferentiated mesenchyme [76]. Conversely, Bmp antagonists give rise to abnormally shortened cartilage elements [83, 84]. The exact mechanisms through which Bmps control chondrogenesis and osteogenesis have yet to be characterized.

Bmps generally bind to a serine/threonine kinase class of transmembrane receptors composed of two distinct types, each having three subtypes [85]. Ligands binding type I receptors (i.e.: Bmpr-IA, Bmpr-IB, and the activin receptor ActR-IA) have been shown to do so with less affinity than those binding type II receptors (i.e.: Bmpr-II, and the activin receptors ActR-II and ActR-IIB) [85]. Type I receptors can be readily distinguished from type II receptors by their GS domain, characterized by a SGSGS motif, that is transphosphorylated in a ligand-dependent manner by the type II receptor following heterotetramerization [85]. The phosphorylation of this motif in the type I receptor leads to the release of the receptor-associated Smad proteins that form a heterotrimeric complex with the cytosolic Smad4 protein, a common mediator of the Bmp signal transduction pathway [85]. This complex translocates into the nucleus to induce the transcription of target genes [85].

Bmps and their receptors are expressed in spatially and temporally dynamic patterns throughout vertebrate development. During chondrogenesis, several Bmps are

expressed in the growth plate. For example, most are expressed within the perichondrium [81] and others, such as Bmps-6 and -7, are specifically expressed by prehypertrophic and proliferative chondrocytes, respectively [43, 86]. Similarly, Bmpr-IA and Bmpr-II receptors are expressed by proliferative and maturing chondrocytes [87], whereas Bmpr-IB expression is restricted to the perichondrium [88].

The size of the Bmp family and the diversity in expression patterns of its members illustrate that its function during vertebrate development is quite complex. Indicative of this, efforts to resolve the nature of the interaction between Bmps and Ihh have yielded inconclusive findings. For instance, there has been evidence provided to suggest that Bmps may regulate *Ihh* expression [76, 77, 82, 84] and vice versa [42, 49, 83, 89].

The claim that Ihh regulates *Bmp* expression is supported by i*n situ* hybridizations that revealed an upregulation of *Bmp2* and *Bmp4* expression along the perichondrium following ectopic misexpression of *Ihh* in chick embryonal wings [83]. Similarly, Minina *et al.* (2001) showed upregulation of *Bmp3* and *Bmp7* in the perichondirum and proliferating chondrocytes of cultured mouse embryonic forelimb explants misexpressing *Ihh* driven by a *ColII* promoter. The proposed interaction between Ihh and Bmp3 is further supported by the fact that *Ihh* null mice fail to express any Bmp3 throughout the entire bone rudiment [42]. Despite the compiled research concerning Bmps, there has not been any convincing evidence provided to establish a mechanism whereby Ihh directly upregulates the expression of *Bmps*. Even if Ihh were to induce *Bmp* expression *in vivo*, there is no certainty of a direct causative effect by Ihh until *Bmp* promoter assays have been conducted to substantiate such a claim.

Alternatively, it has been postulated that Bmps directly upregulate the expression of *Ihh* which activates the Ihh/Pthrp regulatory loop, resulting in delayed chondrocyte

maturity in Bmp-treated cartilage elements [77, 82]. Indeed, the phenotype described from the misexpression of *Ihh* in chick embryonal wings [43] resembles that of Bmp treatment or Bmp receptor misexpression [77, 81, 82]; nonetheless, these findings are inadequate to corroborate such a conjecture. Until recently, there have been no attempts to irrefutably prove that the increased zone of *Ihh* expression that is observed in Bmp-treated mouse and chick embryonic limbs is directly caused by induction of *Ihh* transcription. Seki *et al.* (2004) have demonstrated that *Ihh* promoter activity is successfully induced by Bmp signalling using luciferase transfection assays. Eight Smad binding motifs, conserved between mouse and human, were identified in the immediate upstream region of *Ihh* [49]. It seems all eight are required for maximal response, though five Smad motifs were found to be more integral for transcriptional activity upon mutation analysis [49]. Notwithstanding the limitations of such an *in vitro* model, this is the most convincing set of data to suggest that there may be a direct interaction between Bmps and *Ihh* expression *in vivo*.

2.7. Runx2/Cbfa-1

Another potential upstream regulator of *Ihh* expression is **Runx2**/Cbfa-1. *Runx2* belongs to a family of transcription factors with strong homology to the *Drosophila* pair-rule gene *runt* [90]. Initially, Runx2 was thought to be an osteoblast-specific transcription factor since it was shown to be essential for osteoblast differentiation during embryogenesis [91]. Mice null for *Runx2* lacked osteoblasts and were devoid of all intramembraneous and endochondral bone [50, 92, 93]. Nevertheless, recent findings have uncovered an additional role for Runx2 during chondrocyte differentitation [50, 92-94]. Closer examination of *Runx2* homozygous null mice revealed a delay in chondrocyte

differentiation in some areas, while other areas were devoid of hypertrophic chondrocytes [92, 93]. Consistent with this model, *Runx2* misexpression in chick embryos displayed accelerated hypertrophy and bone formation when compared to wild-type embryos [94]. Futhermore, *Runx2* is expressed in prehypertrophic and hypertrophic chondrocytes and in the perichondrium [92-94]. Therefore, it was determined that Runx2, contrary to Ihh, positively regulates chondrocyte differentiation during chondrogenesis. Investigators also observed an altered pattern of *Ihh* expression in *Runx2* homozygous null mice [50, 92, 93], suggesting that either the chondrocyte differentiation process has been disturbed in these mice or that Runx2 regulates *Ihh* expression. Therefore, using a chromatin immunoprecipitation (ChIP) assay, seven putative Runx2 binding sites were identified in the upstream region of the *Ihh* gene, five of which were conserved between mouse and human [50]. Transcriptional activation was confirmed through a dual luciferase assay system by transfecting a 1.2kb promoter region construct containing the Runx2 binding sites into the chondrogenic ATDC5 cell line [50]. The specificity of transcriptional induction by Runx2 was additionally confirmed through mutational analysis [50]. Thus, Runx2 induces the expression of *Ihh in vitro*. Consequently, it is possible that Runx2 is somehow involved with Ihh in a finely balanced regulation that controls the rate at which growth plate chondrocytes hypertrophy. On the other hand, Runx2 may also independently regulate chondrocyte differentiation in parallel to the Ihh/Pthrp regulatory loop. Needless to say, further investigation is necessary to determine the exact regulatory mechanisms involved in controlling hypertrophy and the significance of this interaction *in vivo*.

2.8. Fibroblast Growth Factor Receptor-3

Fibroblast growth factors (Fgfs) are most recognized for their functionality as potent mitogens and differentiation factors for endothelial and neuronal cells [95]. In addition to this, Fgf signalling pathways are essential during the early stages of limb development and throughout skeletal development [96]. The Fgf family comprises 22 genes encoding structurally related proteins which signal through a class of cell-surface tyrosine kinase receptors, also known as Fibroblast growth factor receptors (Fgfrs) [97]. The Fgfr family is comprised of four membrane-spanning receptors [96]. The binding of the Fgf ligand to its receptor results in Fgfr dimerization accompanied by activation of the tyrosine kinase domains and the subsequent transphosphorylation of the receptor monomers [98]. The phosphorylated tyrosine sites bind Src Homology 2 (SH2) domain-containing proteins with high affinity [99]. Upon binding, these signalling proteins trigger downstream cascades which eventually result in biological responses, often altering gene transcription [95].

Fgfrs are expressed in many tissues throughout embryonal development, including cartilage and bone [96]. It is well documented that Fgfr3 is distinctly implicated in the regulation of endochondral ossification during late embryonic and postnatal development [95]. Fgfr3 is expressed in the growth plate on the surface of proliferating and prehypertrophic chondrocytes and has been identified as a negative regulator of bone growth, as demonstrated by skeletal overgrowth in *Fgfr3* knockout mice [48, 100, 101]. More specifically, Fgfr3 is a negative regulator of chondrocyte proliferation [102]. The commonly studied mutations in human *FGFR3* are autosomal dominant, frequently arise sporadically, and result in varying degrees of constitutive receptor activation that are manifested by several dwarfing chondrodysplasias such as achondroplasia (ACH),

hypochondroplasia (HCH), and thanatophoric dysplasia (TD) [103]. The severity of chondrodysplasia directly relates to the intensity of FGFR3 activation [96]. Despite the subtle phenotype in embryonic mice affected with *Fgfr3* dwarfing chondrodysplasias, such as achondroplasia ($Fgfr3^{ACH}$), the shortening of the growth plate in post-natal stages is much more prominent when compared to wild-type mice [48]. The growth plate narrowing in $Fgfr3^{ACH}$ mice is a consequence of the diminished pool of proliferating chondrocytes [48]. Additionally, it was demonstrated that the level of *Ihh* and *Ptch* expression in the growth plate of $Fgfr3^{ACH}$ mice is very significantly reduced in comparison to wild-type [48, 104]. Taking into account the evidence that Ihh positively regulates the proliferation of chondrocytes [42, 47], the latter finding suggests that FGFR3 may suppress chondrocyte proliferation by negatively regulating *Ihh* expression.

Furthermore, an SH2 domain-containing signal transducer, Signal transducer and activator of transcription-1 (Stat1), has been identified as a downstream mediator of the Fgfr3 signalling pathway in TD mice [105]. Therefore, Stat1 may, at least in part, be responsible for inhibiting cell proliferation by repressing *Ihh* expression. However, Stat binding sites have yet to be identified in the upstream promoter region of *Ihh*. Hence, the hypothesis that Fgfr3 regulates *Ihh* expression through Stat1 signalling was tested in our study of the regulation of *Ihh* transcription.

3. Materials and Methods

All manipulations that required the use of commercial reagents and kits were performed according to the manufacturer's protocol, except where indicated.

3.1. Chicken *Ihh* Isolation from a Genomic Library

Genomic clones containing the *Ihh* sequence for chicken (*Gallus gallus)* were isolated from a White Leghorn (10 day old male) chicken genomic library in the Lambda Fix-II vector (Stratagene). The library was plated on XL-1 Blue MRA bacteria using standard protocols [106]. Three rounds of screening were performed in order to isolate single phage clones that contained the *Ihh* gene. In the primary screen, twelve positive clones (designated 2.1, 2.2, 5.1, 14.1, 15.1, 15.2, 16.1, 17.1, 17.2, 17.3, 18.1, and 19.1) were isolated from a total of $1 \cdot 10^6$ plaque forming units on twenty 150mm agar plates. Those clones were subsequently re-screened on 100mm agar plates. The secondary screening produced eight positive clones (designated 2.1.1, 2.1.2, 2.2.1, 5.1.1, 17.3.1, 17.3.2, 18.1.1, and 19.1.1). Two positive clones were picked from each 100mm agar plate from the tertiary screening and their DNA was extracted.

Hybond-N 132mm or 82mm nylon membranes optimized for nucleic acid transfer (Amersham Pharmacia Biotech) were used to lift plaques from the agar plates. The membranes were hybridized in Church buffer (7% SDS, 300mM Sodium Phosphate Buffer pH 6.8, 5mM EDTA) with 10^6 cpm/ml of radioactive probe at 65°C overnight. The *Ihh* probe was synthesized from a mouse *Ihh* cDNA template (B. St-Jacques, unpublished) using 50μCi of [α ^{32}P] dCTP (Amersham Pharmacia Biotech) and the Random Primers DNA Labelling System (Invitrogen). The membranes were washed

under medium-stringency conditions (1% SDS, 140mM Sodium Phosphate buffer pH 6.8) at 65°C. Following the wash, the membranes were exposed on X-OMAT AR scientific imaging film (Kodak) in a Source One cassette (Picker International Health Care Products) at -80°C, and later developed in the M35A X-OMAT Processor (Kodak).

3.2. Plasmids

The DNA of phage clones obtained from the tertiary screening was digested with BamHI, EcoRI, NotI, or ClaI restriction enzymes (New England Biolabs). The digest products were separated by agarose gel electrophoresis and blotted on Hybond-N+ membranes (Amersham Pharmacia Biotech) according to standard protocols [106]. The blotted digests were hybridized overnight at 65°C in Church buffer with $5x10^5$ cpm/ml of the mouse *Ihh* exon1-specific radioactive probe. The blot was stripped and later hybridized with an equal amount of exon3-specific radioactive probe. The mouse *Ihh* 562bp exon1 and 369bp exon3 probe templates were obtained from EcoRI/SacI and ApaLI/BsmBI double digests of the mouse *Ihh* cDNA plasmid (B. St-Jacques, unpublished), respectively. The blots were washed under stringent conditions (1% SDS, 40mM Sodium Phosphate Buffer pH 6.8) at 65°C, exposed on film at -80°C, and developed. A 6kb band from the NotI digest of clone 2.1.1.1 and a 600bp from the NotI digest of clone 17.3.1.2 hybridized with the exon1-specific probe. A 7kb fragment from the NotI digest of clone 17.3.1.2 hybridized with the exon3-specific probe. These fragments were gel extracted with the PCR Purification Kit (Qiagen) and were ligated into the pBluescript II KS(-) cloning vector (Stratagene) using Quick T4 DNA Ligase (New England Biolabs). Large stocks of these recombinant plasmids were obtained by

transforming XL-1 Blue competent bacteria (Stratagene) and purifying the cultures using the Plasmid Midi Kit (Qiagen).

Fragments of mouse genomic DNA, previously isolated from the Ihh region [42], were cloned into the pGL3 luciferase expression vectors (Promega). The mouse *Ihh* constructs are summarized in figure 4. A 5.5kb fragment located 133bp upstream of the mouse *Ihh* ATG start codon, named **1879**, was isolated by an XhoI/NcoI digest (see figure 5). This fragment was introduced in the forward direction into the pGL3-Basic vector (Promega) to generate the **p1879B-gap1** construct. Primers 5'-GCTGCCTGCCC-TGGCGCCC-3' and 5'-CGGCCCATGGCC-GGGTAGCC-3' were used to amplify the 133bp fragment located between the NcoI site and the ATG of *Ihh*. The downstream primer was designed to introduce an NcoI restriction site, 5'-CCATGG-3', at the ATG. The PCR product was digested with NcoI and ligated in the unique NcoI site of p1879-gap1 to generate plasmid **p1879B**. The p1879B construct was later digested with CspI and SacII to remove a 464bp fragment and the digested plasmid ends were blunted using the T4 DNA Polymerase (Invitrogen). The plasmid was then re-ligated to itself to generate the **p1879B-gap2** construct.

The p1879B construct was subjected to a series of deletions that incrementally truncated the region immediately upstream of the *Ihh* ATG. All deletion constructs were generated by digesting out fragments of different lengths from the 5'-end of the genomic insert in p1879B and re-ligating the blunted plasmid ends. The first deletion construct, **p1879B-del1**, was generated by an XhoI/EcoRI double digest, removing an 866bp fragment. The **p1879B-del2** construct was created by deleting a 2.7kb fragment with an XhoI/EcoRV double digest. The **p1879B-del3** construct was generated by a SacI digest that resulted in the removal of a 3.9kb fragment. Finally, the **p1879B-del4** construct was

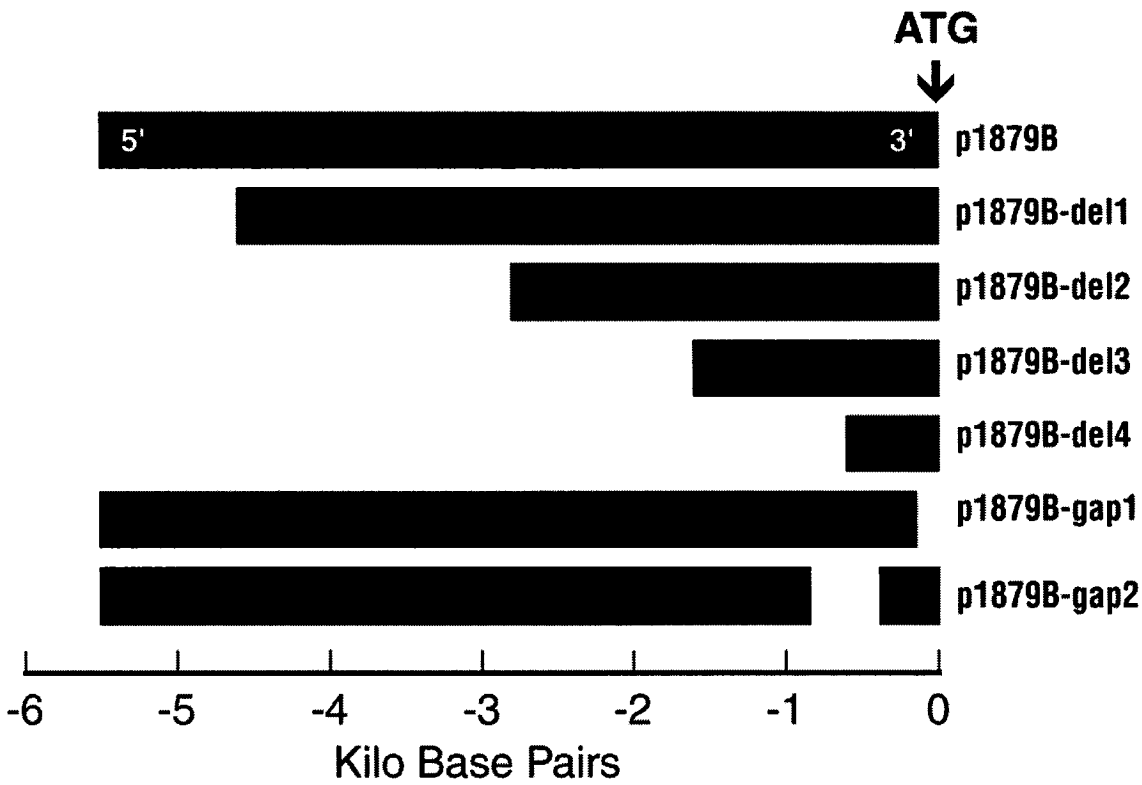


Figure 4: A diagrammatic representation of the p1879B construct and the derived deletion constructs.

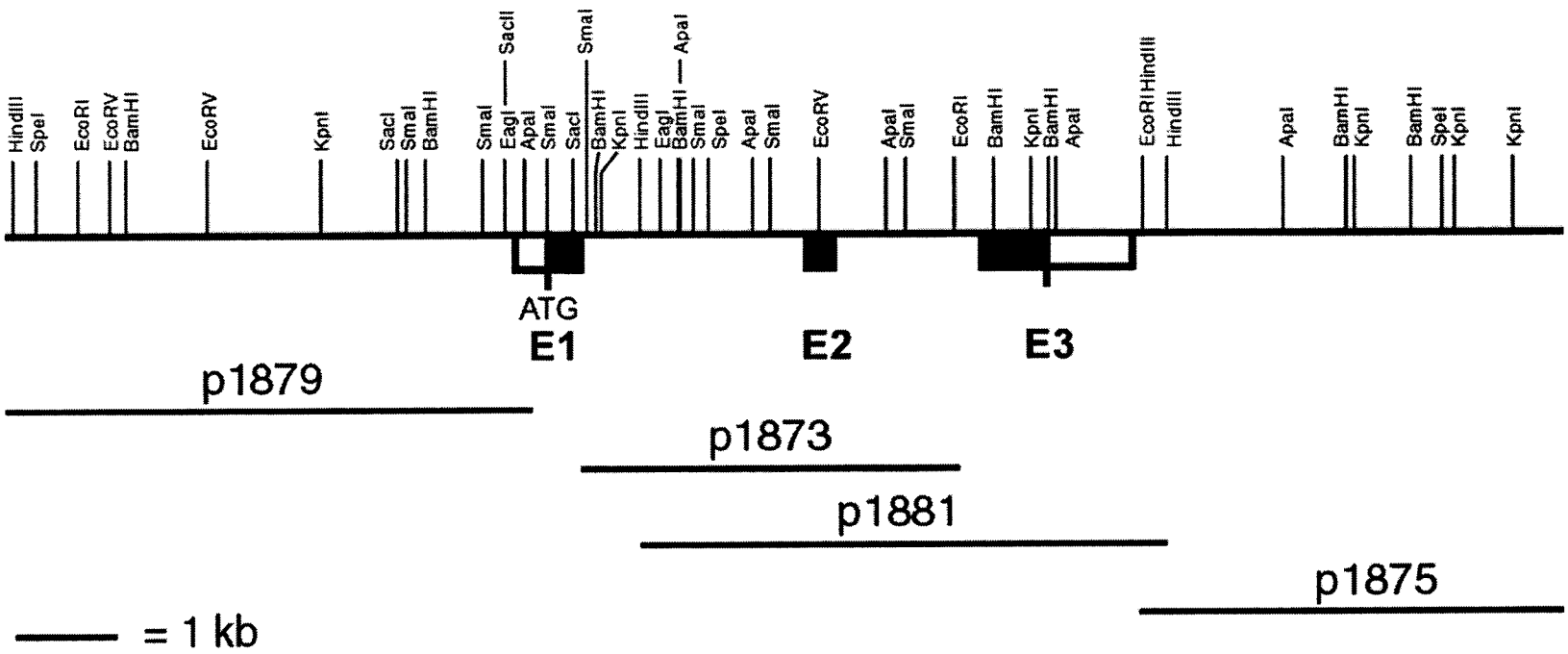


Figure 5: The fragments of mouse genomic DNA, containing various portions of the *Ihh* gene, that were subcloned in pBluescript vectors. The 1879 fragment was subsequently subcloned into the pGL3-Basic vector to generate the p1879B *Ihh-luciferase* construct.

produced from a KpnI/PstI double digest, with T4 DNA Polymerase treatment, which deleted a 4.9kb fragment.

3.3. Interspecies Sequence Analysis

The chick *Ihh* clones were sequenced at the Sheldon Biotechnology Centre at McGill University. Primers used for sequencing chick *Ihh* were synthesized either at the Shriners Hospital for Children or at the Sheldon Biotechnology Centre. The oligos that were used to sequence the region upstream of the chick *Ihh* gene are 5'UTR3 (5'-GTT-TGCGCCCTGGCGGTCAG-3'), 5'UTR4F (5'-GGAAATTGAAGAGATCCG-3'), 5'UTR5F (5'-GGGATCAAACAGAGCAGGGAGG-3'), and 5'UTR7F (5'-CGAGAG-GTGTCTCCATCAGG-3'). The primers used to sequence intron1 are SmallIntron1.1F (5'-GACATCATCTTCAAGGACGAGG-3'), IhhPr3R (5'- CGGAGATGGCCAGGGA-GTTCA-3'), Intron1.1R (5'-GGACCAGTTGTGGTCCCATACGG-3'), Intron1.4R (5'-GGAAGGTGTTGGACAAACAGTGC-3'), and Intron1.5R (5'-CTGCACTGCGTTGT-TCCAGG-3'). The primers used to sequence intron2 are IhhPr4F (5'-GTCCAAGGCGC-ACATCCACTG-3'), Intron2.1F (5'-GGTGTCCTATGGCCATTGCG-3'), Intron2.4F (5'-GCAGCAGAGGGACGAAGCGTGG-3'), Intron2.6F (5'-GTGTATGTATTGCGTC-CCTTTGG-3'), Intron1.5F (5'-GGTGACTCTCTTCATGGGGCTGG-3'), and Intron1.3F (5'-CCTTTCCGCATCCCTGAGTCGTACC-3'). The region immediately downstream of exon3 was sequenced with IhhPr6F (5'-ATGCTGCTGCCCCCCGACAG-3'), 3'UTR1F (5'-GCACCCAAACCAGAGCCTTGTACG-3'), and 3'UTR2F (5'-CCGGAAGCATA-AAGTGTAAAGCC-3').

Analysis of inter-species sequence conservation within and surrounding the *Ihh* gene was performed using the UCSC Genome Browser (http://genome.ucsc.edu) and

multiple-sequence alignment applications [107, 108]. Sequences analyzed were from: human (*Homo Sapiens*; genomic position on chr2:219,744,651-219,750,413, May 2004 UCSC assembly), chimp (*Pan troglodytes*; genomic position on chr13:109,908,158-109,914,081, November 2003 UCSC assembly), mouse (*Mus Musculus*; genomic position on chr1:75,405,759-75,411,741, May 2004 UCSC assembly), rat (*Rattus norvegicus*; genomic position on chr9:74,434,501-74,440,500, June 2003 UCSC assembly), and chicken (*Gallus gallus*; sequence obtained from 2.1.1.1 and 17.3.1.2 chick phage clones and that from the genomic position on chr7:23,074,628-23,083,840, February 2004 UCSC assembly). The Berkeley Drosophila Genome Project Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html) and Center for Biological Sequence Analysis Promoter Prediction Server 2.0 (http://www.cbs.dtu.dk/services/promoter/) online applications were used to identify a putative transcription start site for *Ihh*. The TFSEARCH (http://www.cbrc.jp/research/ db/TFSEARCH.html), TESS (http://www.cbil.upenn.edu/tess/), Transplorer/ Pathsearch (http://www.genomatrix.de), Match (http://www.gene-regulation.com/), Alibaba (http://www.alibaba2.com/), and Celera Discovery System (http://www.celeradiscoverysystem.com/) online applications were used to identify putative transcription factor binding sites within the conserved regions between human, mouse, and chick *Ihh*.

3.4. Promoter Analysis

The Primer Extension System (Promega) and 5' RACE System (Invitrogen) were used to determine the transcription initiation site of the mouse *Ihh* gene. The primers used for Primer Extension are Ihh-PE1 (5'-CAGCAGCAGGAACAGACAGAACCGCAG-TCG-3'), Ihh-PE2 (5'-CGACGCGGCTCAAGGCCGACGGGACTCAGG-3'), Ihh-PE3

(5'-CGAGTGAGAGGGGAAATGGAAGAGATCCGG-3'), Ihh-PE4 (5'-CTGCGGAG-AAGCAAACCCGAAAGCTGGAGG-3'), and MTf PE (5'-TTCCTCTTCCCTGGTC-TCTCTGGCCTTCAC-3'). The primers used for 5' RACE are GSP1 (5'-GCGATCTTG-CCTTCGTA-3') and GSP2 (5'-GGAATTCCGGCTGAACTGCTTGTAGGCAAGA-3'). The primers used for Primer Extension were purified by removing unincorporated labelled nucleotides using the ProbeQuant G-50 Micro Columns (Amersham Biosciences).

3.5. Real-time PCR

The ABI Prism 7700 Sequence Detection System was used for real-time PCR analysis of *Ihh* expression in ATDC5 and COS7 cultured cells and mouse kidneys. Cell and tissue RNA were extracted using Trizol reagent (Invitrogen). The reverse transcriptase products were synthesized using the High Capacity cDNA Archive Reagents (Applied Biosystems), 5μl of which were used for each PCR amplification. The Taqman Gene Expression Assays (Applied Biosystems) Mm00439613_m1 and Mm99999915_g1 primers were used for amplification of *Ihh* and *Gapdh*, respectively. The real-time PCR conditions were programmed by the manufacturer. Similarly, the expression of *Ihh* among samples was calculated in relation to that of subconfluent ATDC5 (d0) cultures according to the manufacturer's Comparative ΔC_t Method.

3.6. Cell Lines

The ATDC5, COS7, and C3H10T1/2 cell lines were used to study *Ihh* expression within an *in vitro* system. The ATDC5 chondrogenic cell line originates from the mouse

AT805 teratocarcinoma cells and were obtained from the RIKEN cell bank, #RCB0565 (Tsukuba, Japan) [109].

ATDC5 pre-confluent (d0) baseline cells were cultured with Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12) (Gibco) supplemented with 5% Fetal Bovine Serum (FBS) and 100μg/mL Penicillin-Streptomycin. Post-confluent ATDC5 cells (d21) were cultured with DMEM/F12 supplement with 5% FBS, 100μg/ml Penicillin-Streptomycin, 10 μg/ml insulin, 10 μg/ml transferrin, and 3 x 10^{-8} M Sodium Selenite [110]. COS7 and C3H10T1/2 cells were a generous gift from Dr. R. St-Arnaud (Shriners Hospital for Children, Montreal) and were cultured in low-glucose Dulbecco's Modified Eagle Medium (DMEM) (Gibco) with 10% FBS.

3.7. Transient Transfections

ATDC5 cultures, d0 and d21, were trypsinized and re-seeded the day prior to transfection in 6-well plates (Fisher) at a density of $3 \cdot 10^5$ cells/well. The cells were transfected with 3μl of either Lipofectamine (Invitrogen) or Fugene6 (Roche), 2μg of the desired mouse *Ihh-luciferase* plasmid, and 20ng of the internal control *renilla luciferase* plasmid (pRL-SV40; Promega). Each *Ihh* construct was transfected in triplicate. Five hours after transfection, the medium was supplemented with FBS to obtain a final concentration of 5%. The cells were lysed 48 hours post-transfection with the Gene Reporter Assay Lysis Buffer (Roche). The transcriptional activity of *Ihh* was quantified using the Dual Luciferase Reporter Assay System (Promega) and the Monolight 2010 Luminometer (Analytical Luminescence Laboratory).

COS7 and C3H10T1/2 cultures were trypsinized and re-seeded one day prior to transfection in 6-well plates (Fisher) at a density of $1.2 \cdot 10^5$ cells/well. In each well, 3μl of

Lipofectamine (Invitrogen) was used to transfect 1μg of the desired mouse *Ihh-luciferase* plasmid with 10ng of the internal control pSv6TK-CAT [111], a generous gift from Dr. R. St-Arnaud (Shriners Hospital for Children, Montreal). Each *Ihh* construct was transfected in triplicate. Five hours after the start of transfection, the medium was supplemented in FBS to obtain a final concentration of 10%. The COS7 lysates were quantified for *luciferase* activity using the Luciferase Assay System (Promega) and the Monolight 2010 Luminometer. Additionally, the lysates were quantified for *chloramphenicol acetyl transferase (CAT)* activity using the CAT-ELISA kit (Roche) and the ELx 808 microtiter plate reader (Bio-TEK Instruments, Inc.).

The luciferase values for each sample were normalized by obtaining the ratio of firefly luciferase to internal control values. The relative induction of transcription for each sample was calculated by comparing the sample ratio to the baseline pGL3-Basic vector ratio.

3.8. FGF9 Treatment of COS7 Cells

COS7 cells were seeded at a density of $1.2 \cdot 10^5$ cells/well and incubated in the conditions stated above for 48 hours in the presence of 20ng/ml FGF9 and 2μg/ml heparin [112]. During transient transfections, FGF9 and heparin were added to the COS7 medium when the cells were supplemented with serum following the initial five hour incubation period.

3.9. Stable Transfections

ATDC5 cells were stably transfected with the p1879 construct. The cells were seeded in a 12-well plate (NUNC) at a density of $1.5 \cdot 10^5$ cells/well. ATDC5 cells were

co-transfected with 3µl of Lipofectamine (Invitrogen) in a 9:1 DNA ratio of the p1879 construct to the pKJ1 neomycin-resistance selection plasmid, a generous gift from Dr. R. St-Arnaud (Shriners Hospital, Montreal). Cultures were selected over a period of two weeks using the Geneticin G-418 Sulfate (Gibco) neomycin antibiotic at a concentration of 400µg/ml, and maintained thereafter at a concentration of 200µg/ml. The surviving clones from the same well were pooled, but each well was cultured independently. Once the stably transfected cell lines were established, *firefly luciferase* activity was quantified from subconfluent (d0) and post-confluent (d21) cultures as described above. However, in order to collect the lysates of the d21 cultures efficiently, the cells were subjected to a treatment, described in subsection 3.10, prior to exposing them to the lysis buffer. The luciferase values were normalized using total protein concentrations which were quantified from the Bradford Assay [113].

3.10. Lysis of Post-Confluent (d21) ATDC5 Cultures

This treatment is modified from an articular chondrocyte extraction protocol for application in ATDC5 cultures [114]. Post-confluent (d21) ATDC5 cultured cells were rinsed with Phosphate Buffered Saline (PBS). They were incubated for 15 minutes at 37°C with a 1.0mg/ml solution of Hyaluronidase (Sigma) in DMEM/F12 medium. The solution was replaced with a 2.0mg/mL solution of Collagenase (Sigma) in 0.1X Trypsin-EDTA (Gibco) and incubated for 15 minutes at 37°C. Then, the cells were incubated with 1X Trypsin-ETDA for 5 minutes at 37°C. The culture was broken up into small aggregates or single cell suspension by repeated pipetting. The cells were centrifuged at 50*g* and resuspended in cell lysis buffer.

4. Results

4.1. Sequence of the Chicken *Ihh* Gene

Initially, the chicken (*Gallus gallus*) *Ihh* genomic sequence could not be examined for inter-species conservation because it was not publicly available. Therefore, a chicken genomic library was screened, using the chicken *Ihh* cDNA as a probe. Inserts in two phage clones (2.1.1.1 and 17.3.1.2) were shown to cover the entire chicken *Ihh* gene. The specific fragments of the inserts from these phages were subcloned into the pBluescript plasmid. Subcloned fragments were sequenced from plasmid DNA in a nested fashion, such that every sequencing reaction provided primer sequences for the next. With the exception of 1.5kb of sequence from intron1, the entire chicken *Ihh* gene was sequenced along with 1.5kb of upstream and nearly 1kb of downstream sequence. In February 2004, the chicken genomic assembly became available on the University of California, Santa Cruz (UCSC) database, though it was incomplete. After comparing the published sequence to the sequence obtained from the chicken genomic library, it was determined that the former was inaccurate. More specifically, the contig arrangement was erroneous, largely due to periodical gaps in the sequence. That notwithstanding, the two sequences are virtually identical in content with exception of a few single nucleotide polymorphisms. Thus, a finalized version of the chicken *Ihh* genomic sequence, which includes 6kb on either side of the gene, was compiled using both sequences.

4.2. Inter-species Sequence Conservation Analysis

The human *IHH* genomic sequence was individually compared to that in chimp, mouse, rat, and chicken using the UCSC Genome Browser (see figure 6). The coding

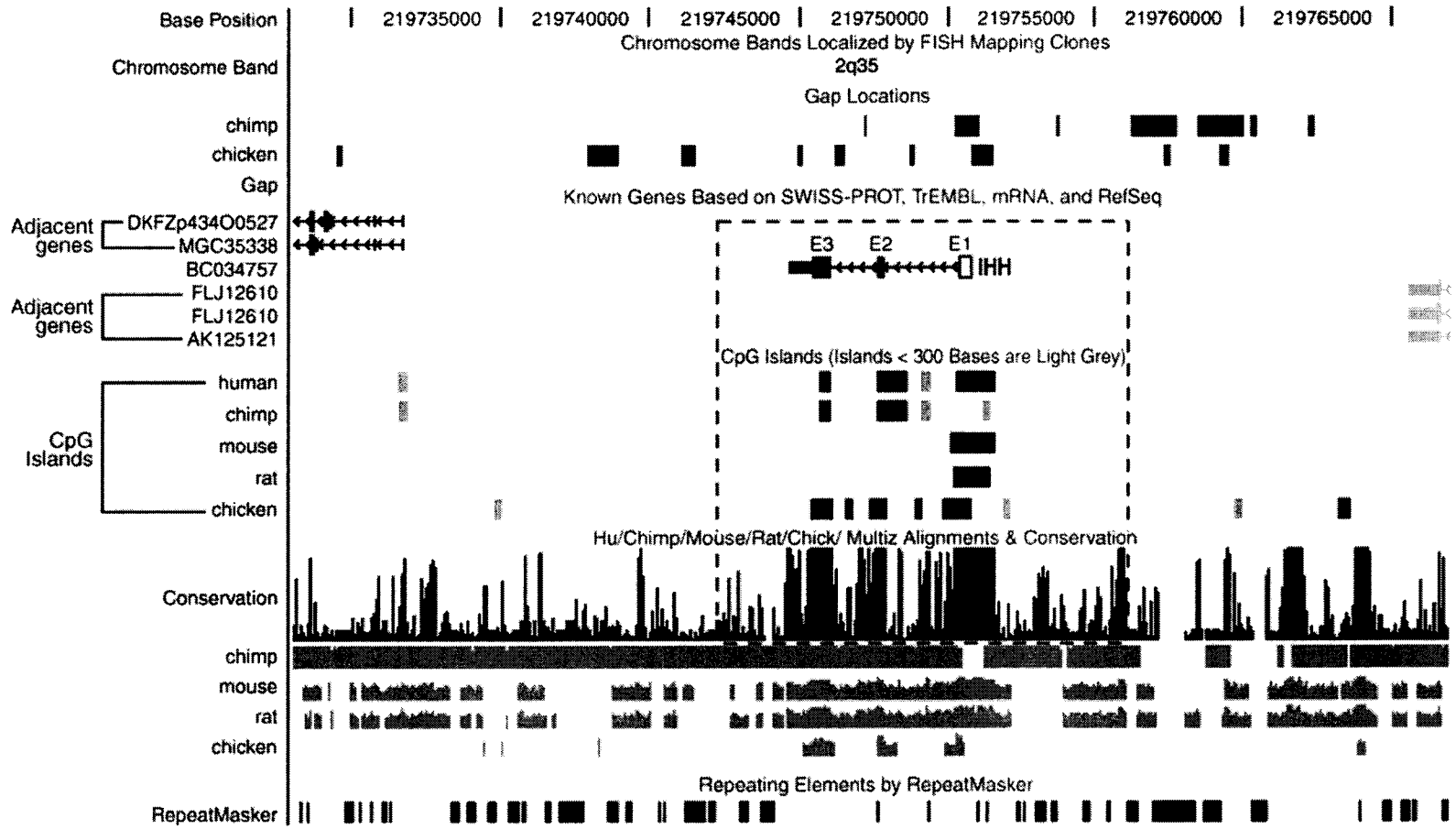


Figure 6: The UCSC Genome Browser cross-species *Indian Hedgehog (Ihh)* gene conservational analysis. Approximately 35kb of genomic sequence from the UCSC human, chimp, mouse, rat, and chicken genome assemblies were compared for sequence, CpG islands, and repetitive sequence conservation. The comparisons are relative to the human *IHH* gene. The nearest adjacent genes are shown approximately 15kb upstream and downstream of *IHH*. The hash-marked area delineating the *IHH* gene represents the sequence used for the *Ihh* multiple species alignment (see appendix A) and demarcates the region with the strongest conservation associated with the gene.

exons of *Ihh* were highly conserved across all species. Note that the apparent lack of conservation in certain regions in a given species is either caused by a gap in the genomic assembly or a fragment that is species-specific. As expected, the human and chimp, as well as the mouse and rat, *Ihh* genes are highly analogous to one another since those respective species are evolutionarily more closely related. Note that the original chicken *Ihh* UCSC genomic assembly was used in this comparison because the UCSC Genome Browser does not provide an option for substituting one genomic sequence for another. Despite the preliminary status of the chicken genomic assembly, it was not anticipated that the non-coding regions of chicken *Ihh* gene would be so poorly conserved in relation to human.

There is a strong conservation of CpG islands in all five species in the promoter region and neighbouring first exon. Unlike the rodents, CpG islands are additionally distributed throughout the gene in human, chimp, and chicken. The highly elevated CpG content in the promoter regions of each species suggests that DNA methylation may not play a regulatory role in *Ihh* gene transcription during early development as it is highly expressed.

The *Ihh* genomic sequences for all species were more closely analyzed with a multiple alignment program for the purpose of identifying potential transcriptional regulatory elements. The hashed box surrounding the *IHH* gene in figure 6 indicates the portion of the genome that was selected for the alignment, which includes 6kb of upstream and 4kb of downstream flanking sequences. The rationale for choosing this area is based on the diminution of inter-species conservation beyond the hashed box and the relative distance to adjacent genes. The data from the multiple sequences alignment reiterate the strong conservation that was previously observed among mammals with the

UCSC Genome Browser. In concordance with the UCSC Genome Browser findings, the coding sequences are well conserved between human and chicken. Furthermore, it was determined that the only significant non-coding conservation between human and chicken *Ihh* is located immediately upstream of the gene and at intron splice sites (see appendix A). The 5' (i.e.: GT) and 3' (i.e.: AG) intron splice site, as well as the exon 3 stop codon (i.e.: TGA), are conserved in all species. Consequently, exons 2 and 3 are both structurally identical between samples. Conversely, the structure of the first exon is somewhat more divergent on account of the distinct ATG start codon sites for each species (see appendix A). In fact, the mouse actually has two in-frame ATG codons distanced 114bp apart, both of which have been suggested to initiate translation (Entrez Protein accessions P97812 and AAH46984). The one farther downstream coincides with the human and rat start codons, which are all slightly upstream to that of the chicken. Effectively, the length of exon1 of *Ihh* is unique to some species.

The Berkeley Neural Network Promoter Prediction (NNPP) program was employed to analyze more than 1kb of human non-coding sequence upstream of *IHH*. The NNPP program identified the most likely *IHH* transcription initiation site to be -230bp from the human ATG, a region that is conserved in all five species (see appendix A). Unfortunately, this prediction can not be corroborated by available sequences because the published human cDNA clone is incomplete and all of the cDNA clones from the other species start from a unique location. In addition, two mouse cDNA sequences and the chicken cDNA sequences extend further upstream than this putative start site, suggesting that the prediction may be inaccurate. On the other hand, the first nucleotide of one published mouse cDNA (GenBank accession AK090147), extracted from the intestinal mucosa of a C57BL/6J mouse, corresponds precisely to the predicted start

location of the *Ihh* transcript. Sequences pertaining to TATA and CCAAT-like elements were subsequently annotated on the alignment at -32bp and -52bp, respectively, upstream of the putative transcription initiation site (see appendix A), which is within the expected range for such regulatory elements.

The transcription factor prediction programs that were used to identify putative regulatory elements within the *Ihh* gene were, for the most part, uninformative. However, the Celera Discovery System search engine revealed two STAT binding sites in the promoter region of *Ihh*. The first site (5'-**TTC**CAT**TTC**-3', STAT1 [115]; the core bases are highlighted in bold), located -241bp from the putative transcription start site, is conserved in all examined species. The second (5'-**GTC**CCC**GAA**-3', STAT3 [116]), located +149bp from the putative transcription start site, is specific to mouse and rat.

4.3. Experimental Promoter Analysis

The transcriptional regulation of the mouse *Ihh* gene is yet to be characterized. In order to begin addressing this issue, a "Primer-Extension" assay was used to identify the mouse *Ihh* transcription initiation site (TIS). Mouse adult kidney RNA was used because *Ihh* is highly expressed in adult kidneys [11]. Four separate primers, spanning approximately 1kb of sequence upstream to the start codon of *Ihh* were used. As a control for assay conditions, a primer and RNA template provided by the manufacturer were also used. The expected 87bp product was obtained using the manufacturer's RNA control template (see figure 7, lane 2). However, all *Ihh* specific primers failed to generate a specific product indicative of the start location of the *Ihh* transcript (see figure 7, lanes 4 and 5; only two of the four *Ihh* primers are shown). The authenticity of the faint signals observed with the *Ihh* primers is questionable because the resulting bands should have

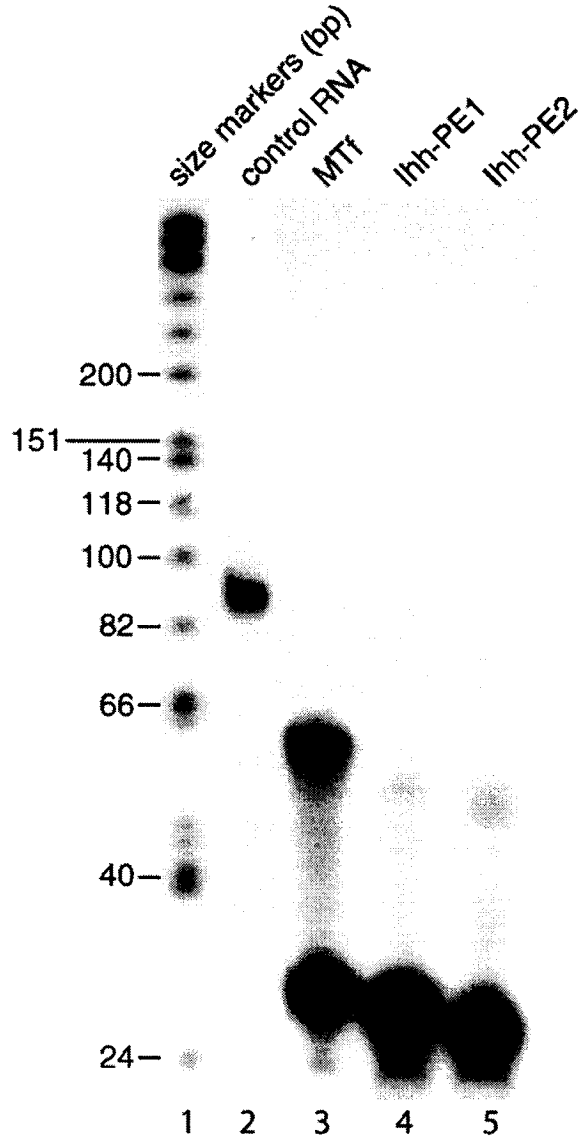


Figure 7: Autoradiogram of the Primer Extension products from the manufacturer's control RNA and primer (lane 2), ATDC5 mature (day 21) RNA with an *MTf* primer (lane 3), and adult mice kidney RNA using two distinct *Ihh* primers (lanes 4 and 5).

been separated by 190bp. As an additional control, a Primer Extension assay was attempted using a previously characterized chondrogenic gene, *membrane-bound transferrin-like protein (MTf)*, to verify whether published results for this gene could be reproduced. The signal obtained in lane 3 did not correspond to the published 115bp product for *MTf* [117], indicating that the Primer Extension assay was not performing as expected. The experimental discrepancies in the *Ihh* and *MTf* TIS determination suggest an overlying problem with the assay.

As an alternative to Primer Extension, a 5'-Rapid Amplification of cDNA Ends (RACE) assay was performed. 5'-RACE also failed to produce a signal corresponding to the *Ihh* ITS. Conversely, the manufacturer's control RNA template produced the expected 711bp band (see figure 8). It is speculated that the difficulties experienced with the 5'-RACE were caused by the high GC-content of *Ihh* promoter region because one of the primers used in the assay is a polydeoxycytosine oligonucleotide that could bind non-specifically a GC-rich region.

4.4. Transient Transfections with *Ihh-Luciferase* Constructs

ATDC5 cells were chosen to study the *in vitro* transcriptional regulation of *Ihh* because they differentiate into chondrocytes when cultured under the appropriate conditions [110], in a process that mimics growth plate chondrocyte maturation. It was determined from real-time PCR assays, that *Ihh* expression peaks in ATDC5 cells after 21 days (d21) in culture, resulting in a 24 fold higher expression than in the baseline d0 cultures (see figure 9). The p1879B *Ihh-luciferase* construct, containing 5.5kb of non-coding sequence upstream of the ATG, was transiently transfected into ATDC5 baseline (d0) and mature (d21) cultures to test for the presence of transcriptional cis-regulatory

Figure 8: Agarose gel electrophoresis of the 5' RACE product using the manufacturer's (Invitrogen) control RNA (left lane). The expected band size is 711bp, as shown.

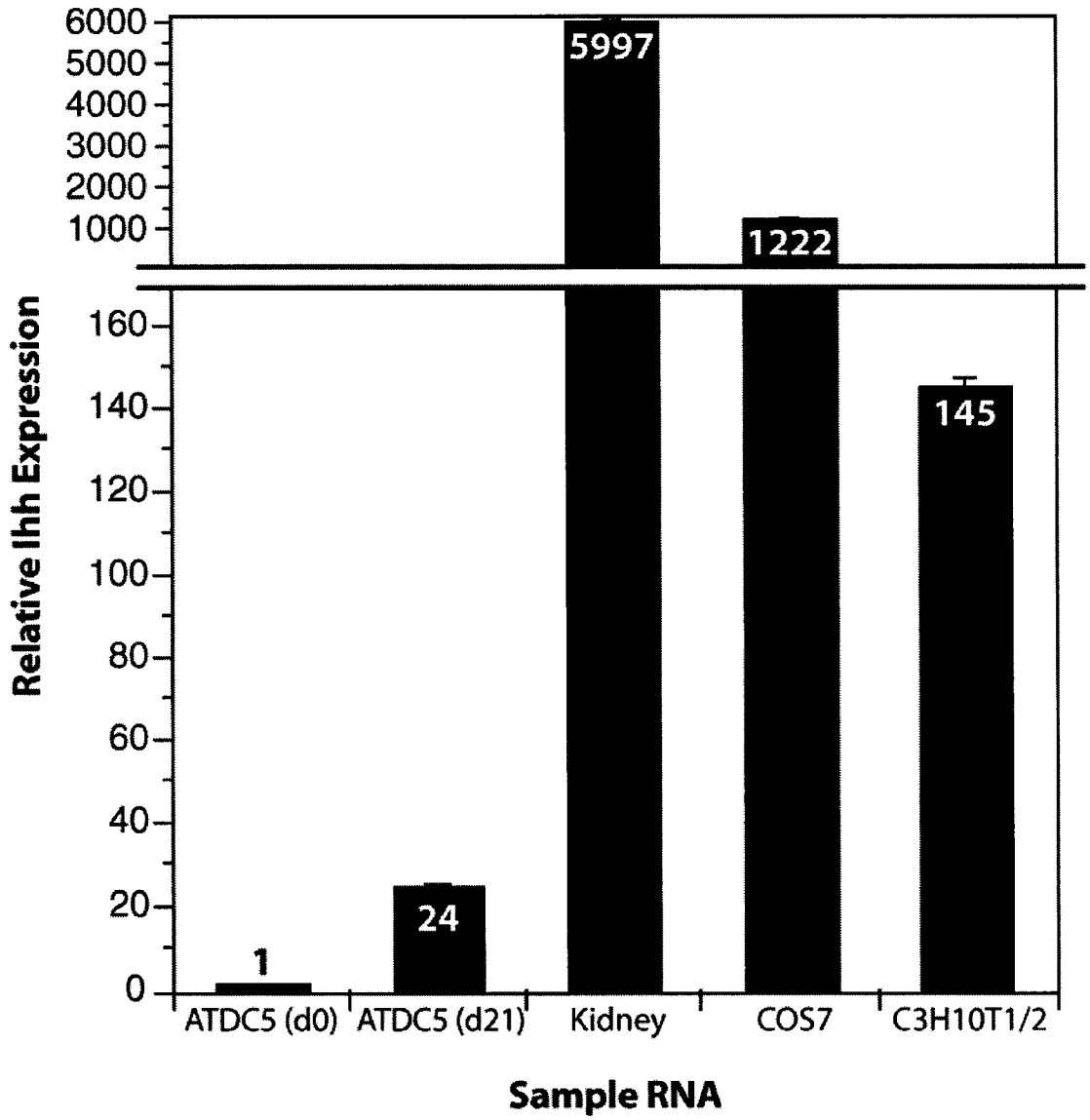


Figure 9: The expression of the Indian Hedgehog gene calculated from the real-time PCR values of various RNA sources in relation to ATDC5 immature (day 0) cells (i.e.: baseline = 1).

elements. An induction in *luciferase* expression was anticipated in ATDC5 d21 cultures due to the higher levels of *Ihh* expression, however problems were encountered during the transfection process.

Originally, the Lipofectamine reagent was used to transfect the *Ihh-luciferase* construct into ATDC5 cultures. A significant decrease in transfection efficiency was observed in mature d21 cultures when compared to that in baseline cultures, making d21 results less reliable (see appendix B). Furthermore, the negative control plasmid transfection efficiencies were dramatically different from those of the experimental plasmid, as judged by the *renilla luciferase* results. Nevertheless, the normalized luciferase values collected from ATDC5 mature cultures for plasmid p1879B were reproducibly close to two fold greater ($p=0.001$) than those from the baseline cultures (see figure 10), indicating enhanced transcriptional activity. In attempts to correct for poor transfection efficiencies in mature ATDC5 cells, the FuGENE6 transfection reagent was also tested. A preliminary transfection trial was performed on ATDC5 d21 cultures with a LacZ-SV40 plasmid to compare the relative transfection efficiencies of Lipofectamine to FuGENE6. The FuGENE6 reagent transfected the LacZ plasmid approximately three fold more efficiently than did Lipofectamine (see figure 11). Similarly, the *renilla luciferase* internal control plasmid transfected more readily with FuGENE6 than with Lipofectamine during transient transfections. Nonetheless, mature ATDC5 cultures responded less strongly to the p1879B *Ihh* construct than did the d0 cultures with FuGENE6 (see figure 10 and appendix B). After closer analysis of the data, it seems that the normalized luciferase values for the d21 cultures are compromised by a high experimental baseline value (see Discussion section for more details).

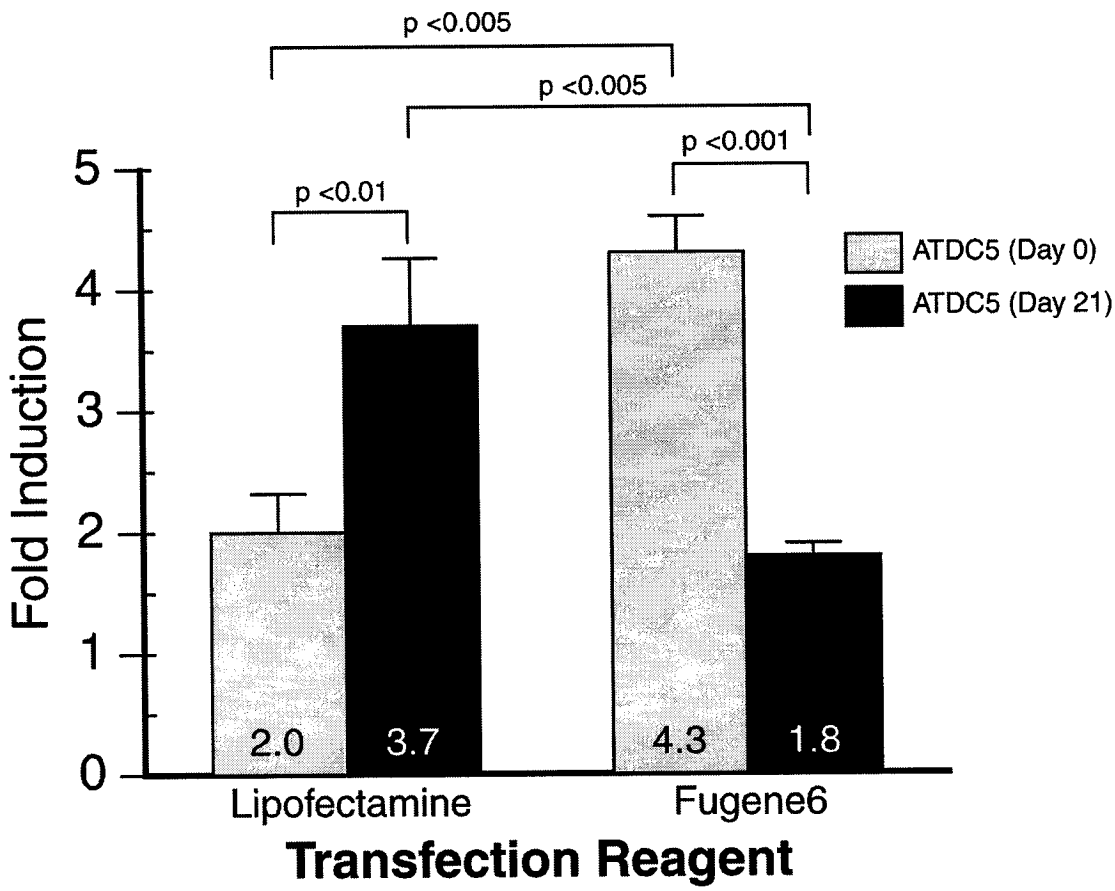


Figure 10: The discrepancy in the *luciferase* expression between transfection reagents in the ATDC5 cell line. Lipofectamine and Fugene6 were used to transiently transfect the p1879B construct into baseline (day 0) and mature (day 21) ATDC5 cells.

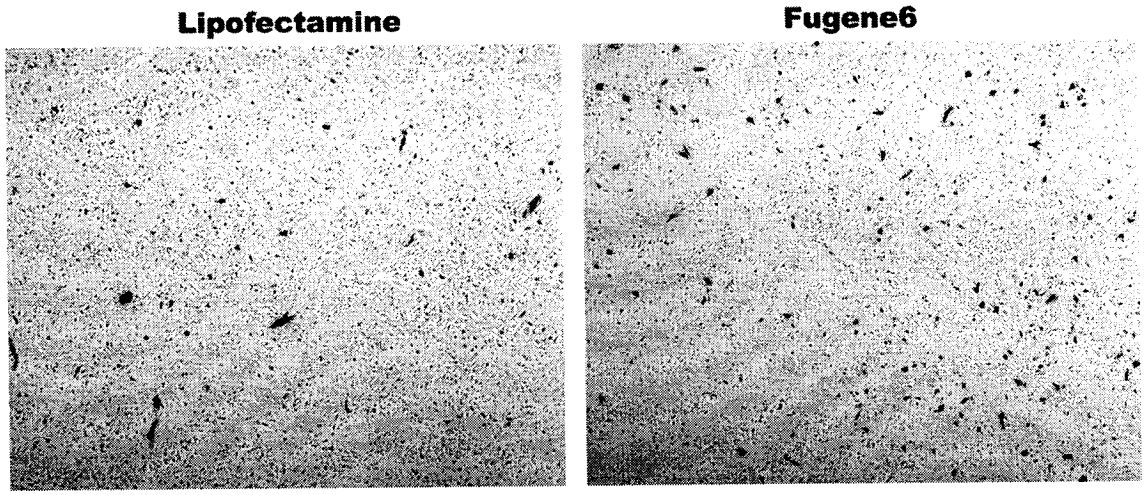


Figure 11: ATDC5 mature (day 21) cells transfected with a LacZ reporter plasmid, using Lipofectamine (left panel) or Fugene6 (right panel), and stained with X-GAL.

Considering the difficulties encountered with mature ATDC5 cultures, only the ATDC5 d0 cells were transiently transfected with *Ihh-luciferase* constructs. The results are summarized in figure 13. In comparison to the negative control plasmid, pGL3-Basic, transfection of either the p1879B construct, which contains 5.5kb of upstream *Ihh* sequence, or the p1879B-del2 deletion construct induced *luciferase* expression by approximately 2-fold ($p<0.01$). Furthermore, the p1879-gap1 construct, from which a Stat3 binding site is deleted, and the p1879B-del3 deletion construct resulted in an average 6-fold *luciferase* induction ($p<0.01$), a response three times stronger than that observed with the p1879B plasmid ($p<0.005$). On the other hand, the p1879B-del1 deletion construct resulted in luciferase levels comparable to baseline values, thus reversing any positive response observed with the complete p1879B construct.

The mature ATDC5 chondrogenic cell line was determined to be an unrepresentative model of *Ihh* transcriptional regulation. Therefore, other cell lines were tested for *Ihh* gene expression. Two cell lines, COS7 and C3H10T1/2, were found to express *Ihh* at levels much higher than ATDC5 cells in any stage of differentiation (see figure 9). These cell lines were transfected with the p1879B *Ihh* construct using the Lipofectamine reagent. The COS7 cells demonstrated a seven fold induction in *luciferase* expression ($p<0.005$) while the C3H10T1/2 cells were unresponsive (see figure 12). Hence, the COS7 cells were selected as an additional *in vitro* model to study *Ihh* transcriptional activity of the 5' upstream region.

The COS7 transfection data are summarized in figure 13. Transfection of the p1879B construct induced *luciferase* expression by more than 2-fold ($p<0.001$) over the pGL3-Basic plasmid. The p1879B-gap1 Stat mutation construct resulted in a *luciferase* induction that almost doubled that of p1879B ($p<0.001$), supporting the putative role of

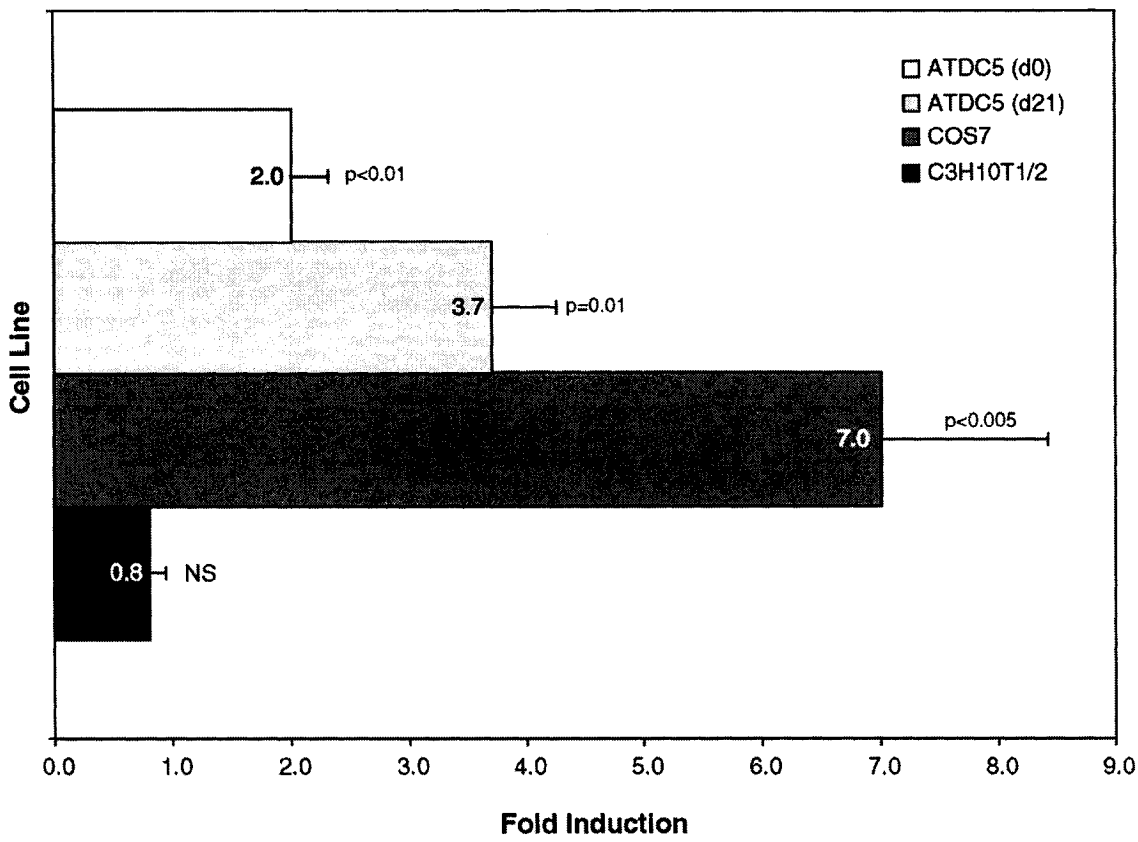


Figure 12: A trial of transient transfections with the p1879B construct to test the induction of *luciferase* expression in the indicated cell lines.

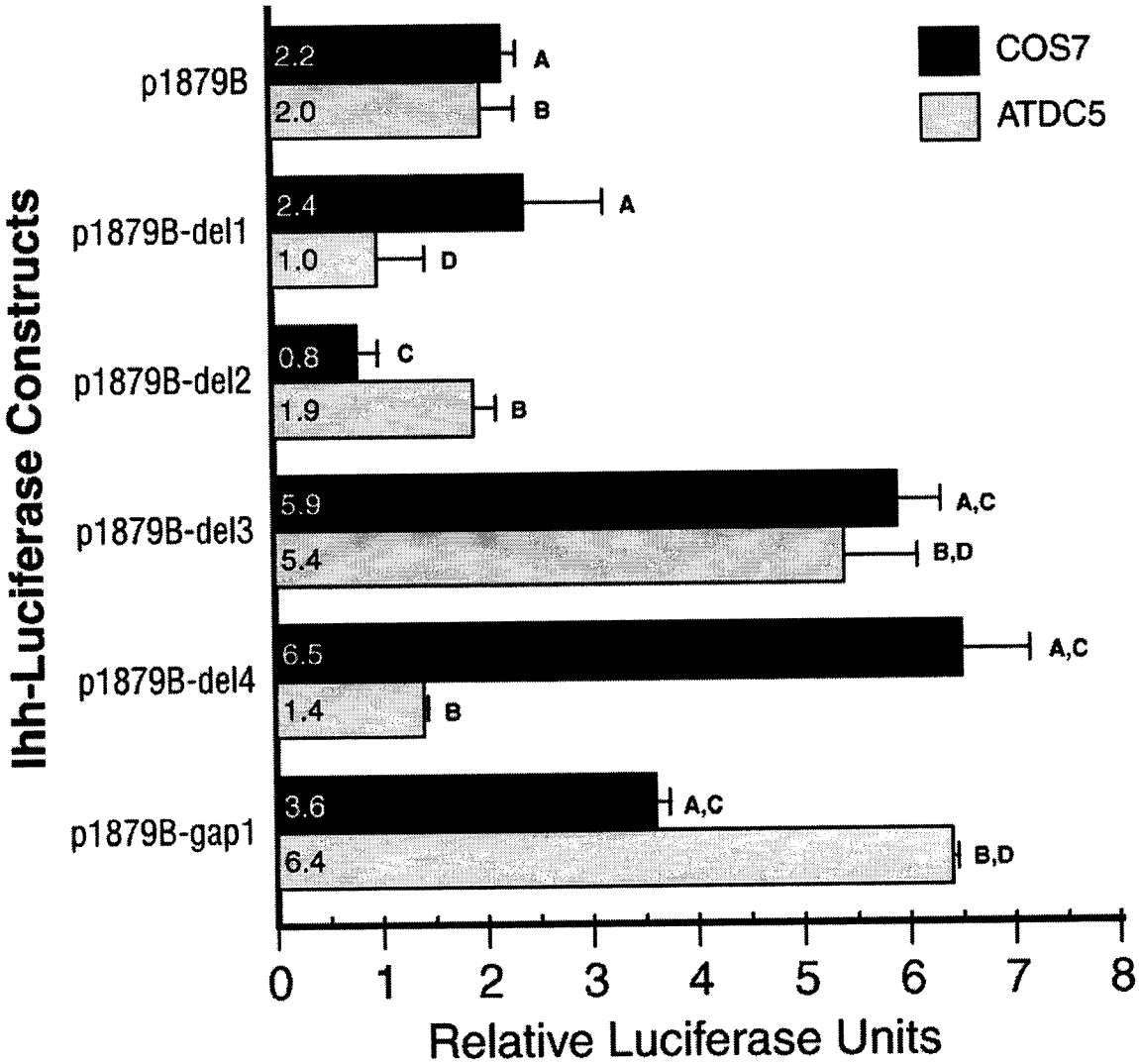


Figure 13: Transient transfections of the complete p1879B and deleted constructs into the COS7 and ATDC5 (day 0) cell lines with the Lipofectamine reagent. Statistical analysis was performed using the Student T Test; **A:** $p < 0.05$ compared to the transfected baseline plasmid (pGL3-Basic) into COS7 cells, **B:** $p < 0.05$ compared to the transfected baseline plasmid (pGL3-Basic) into ATDC5 (day 0) cells, **C:** $p < 0.05$ compared to plasmid p1879B as transfected in COS7 cells, **D:** $p < 0.05$ compared to plasmid p1879B as transfected in ATDC5 (day 0) cells.

Stats to actively repress *Ihh* expression. On the other hand, the second of the Stat mutation constructs, p1879B-gap2, was indistinguishable from p1879B (see figure 14). As for the other deletion constructs, p1879B-del(3 and 4) both resulted in a 6-fold *luciferase* induction, three times greater than the complete p1879B construct ($p<0.001$). Interestingly, the p1879B-del2 deletion construct resulted in a 3-fold decrease of *luciferase* expression in comparison to p1879B ($p<0.001$).

In the aim of further investigating the interaction between *Fgfr3* and *Ihh*, our group has previously demonstrated a significant reduction in *Ihh* expression in ATDC5 chondrogenic cultures treated with recombinant FGF9 protein (St-Jacques, unpublished). Therefore, COS7 cultures were treated with FGF9 and subsequently transiently transfected with the same constructs as stated above. The transfection results are summarized in figure 14. When compared to their untreated counterparts, p1879-del4 and p1879B-gap1 resulted in an additional 2- ($p=0.005$) and 3-fold ($p<0.01$) induction, respectively. The remaining constructs were not affected by treatment.

4.5. Stable Transfections in ATDC5 Cells

As an alternative approach to transient transfections, ATDC5 cells were stably transfected with the p1879B construct. One of ten established neomycin-resistant clones induced *luciferase* expression after 21 days in culture. In comparison to d0 cultures, a 50-fold induction was observed in mature stably-transfected ATDC5 cells (see figure 15). Therefore, stably transfected ATDC5 cells, as opposed to their transiently transfected counterparts, undeniably revealed that the upstream *Ihh* sequence of the p1879B construct contains an active promoter that drives gene expression.

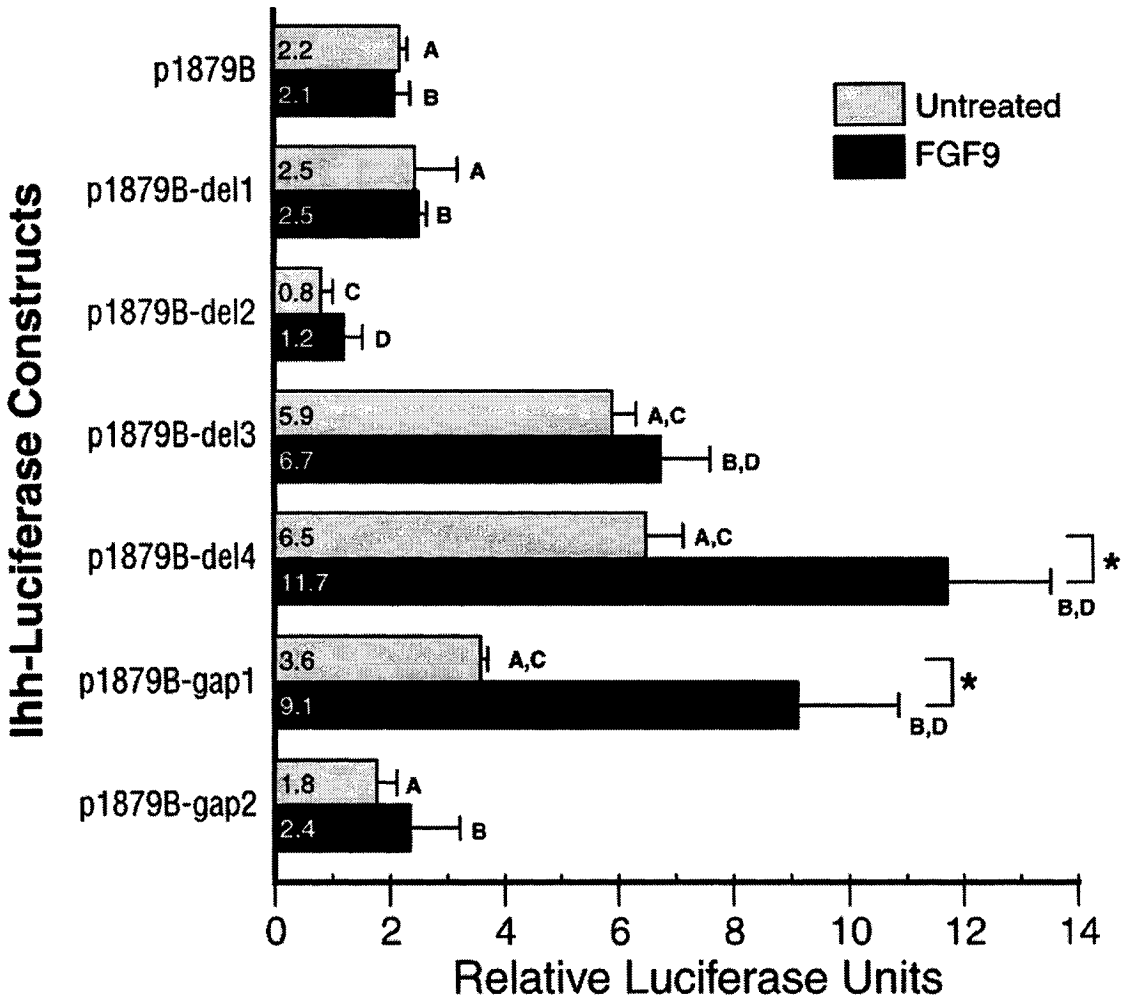


Figure 14: Transient transfections of the complete p1879B and deleted constructs into the untreated and FGF9 treated COS7 cells. Statistical analysis was performed using the Student T Test; **A:** $p < 0.05$ compared to the transfected baseline plasmid (pGL3-Basic) in untreated COS7 cells, **B:** $p < 0.05$ compared to the transfected baseline plasmid (pGL3-Basic) in FGF9 treated COS7 cells, **C:** $p < 0.05$ compared to plasmid p1879B as transfected in untreated COS7 cells, **D:** $p < 0.05$ compared to plasmid p1879B as transfected in FGF9 treated COS7 cells. ★ (star): $p < 0.05$ between the untreated and FGF9 treated transfected COS7 cells for that given plasmid.

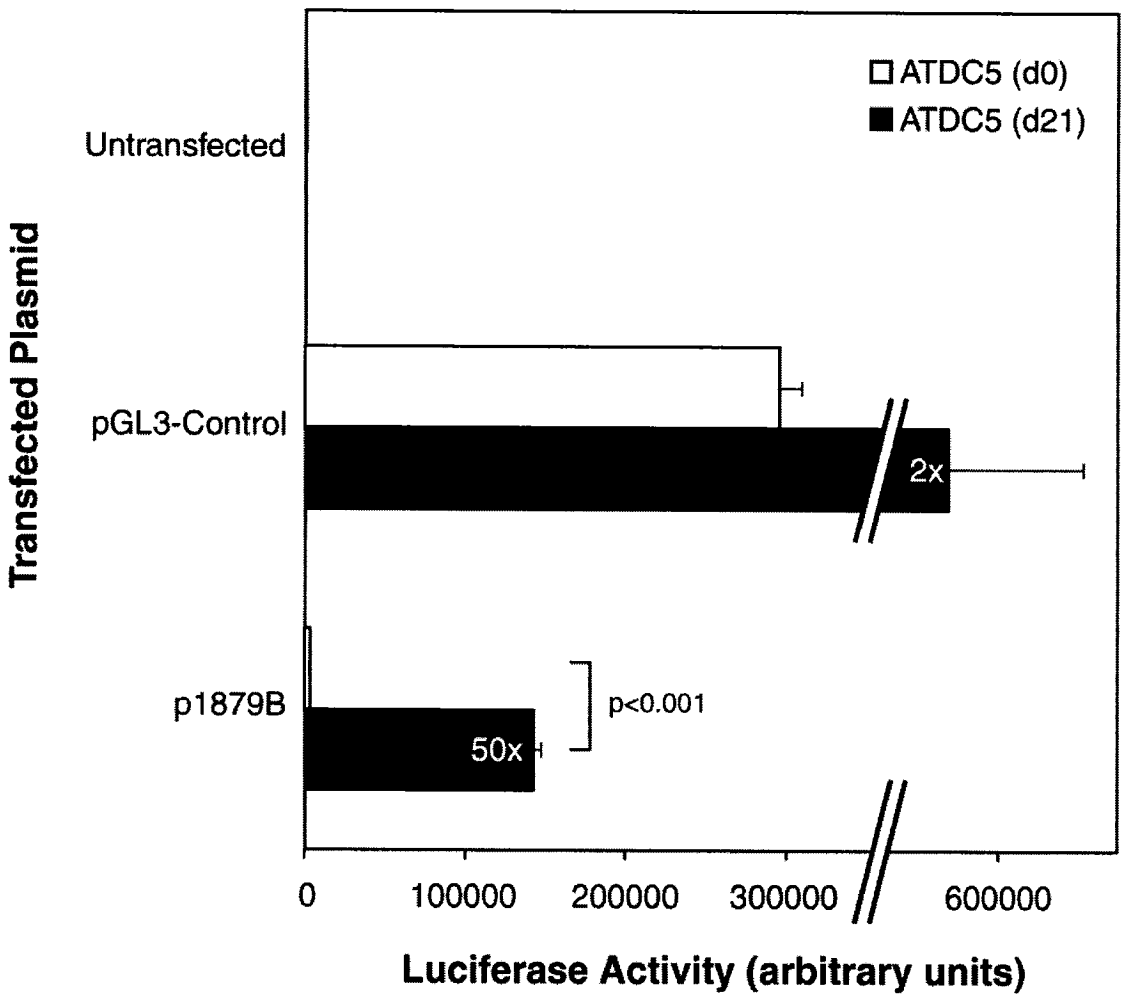


Figure 15: Stable transfections of ATDC5 cells. The numbers within the bars of the mature (day 21) ATDC5 stable transfectants indicate the fold-induction over the immature (day 0) ATDC5 stable transfectants for a given plasmid.

5. Discussion

The best characterized function of Ihh is its role in bone development. Consequently, this project was primarily focussed on the transcriptional regulation of *Ihh* in cartilage cells. Nevertheless, *Ihh* is widely expressed in various types of tissues, suggesting that its transcriptional regulation is likely to be complex. In order to gain better insight as to the location of the *cis*-regulatory elements that control *Ihh* expression, the genomic sequences of multiple species were analyzed for conservation.

5.1. Sequence Conservation

The human and mouse genomic *Ihh* sequences were the only sequences originally analyzed, but non-coding sequence conservation was too high to allow identification of functionally important elements on the basis of sequence comparison alone. Consequently, additional *Ihh* sequences were selected from the genomic assemblies of different species, such as chimp and rat. Sequence similarity between chimp and human, and rat and mouse, respectively, is so strong that conservational analysis between these species did not reveal any new information. Therefore, the sequence in and around the chicken *Ihh* gene was subsequently obtained from a genomic clone with the expectation that it would refine the inter-species sequence analysis. The reason for choosing the chicken is that the function and expression pattern of chicken *Ihh* in cartilage are identical to those in mammals [43]. Furthermore, the evolutionary distance between chicken and mammals is likely to eliminate the redundancies observed among mammalian conservation, facilitating the identification of potential regions important for transcriptional gene regulation.

The five species were initially analyzed using the UCSC Genome Browser online tool to acquire a large-scale overview of the conservation within the *Ihh* genes and their flanking upstream and downstream regions. Specific genomic coordinates required for the multiple species alignment were chosen based on inter-species conservation and the relative distance from the *Ihh* gene. The multiple alignment produced a base-wise analysis of sequence conservation and indicated the nucleotides that were conserved between species hence facilitating the isolation of distinctly conserved sites and the identification of regulatory elements. As predicted, the chicken sequence helped to discriminate which elements are critical to Ihh function, including the coding regions and the promoter region. However, it was disappointing to discover that very little non-coding sequence is actually conserved in chicken. The lack of conservation suggests that this gene is uniquely regulated in chicken even though it codes for a protein that is characteristically identical to that in primates and rodents.

The prediction programs used for identifying potential transcription factor binding sites (TFBS) only recognise canonical elements present in the database search matrix. Thus, many of the functional sites will not conform to the search parameters and a proportion of active regulatory elements will be overlooked. For instance, none of the TFBS prediction applications could identify the recently published Smad and Runx2 binding sites located in the promoter region of the mouse *Ihh* gene[49, 50]. Instead, critical regulatory elements located within conserved regions of the *IHH* gene must be determined experimentally.

The Stat binding sites may have been the only type of regulatory element to be recognized by the prediction applications in the conserved upstream region of *Ihh*, but it is a significant finding with regards to the regulation of bone development. Indeed, Fgfr3

constitutive activation has been reported to downregulate the expression of *Ihh* in chondrocytes [48] likely via activation of Stat factors [104]. Thus, the identification of novel Stat binding elements in the promoter region of *Ihh* substantiates the existence of a direct interaction between the two genes. It should be noted, however, that the putative Stat binding sites (5'-TTCCATTTC-3' and 5'-GTCCCCGCA-3') deviate from the canonical palindromic Stat1 binding sequence, 5'- **TTC**NNN**GAA**-3'. Nonetheless, identical variant sequences have been shown to bind Stat1 and Stat3, respectively, in specific target genes [115, 116].

The transcription initiation site (TIS) of mouse *Ihh* could not be identified experimentally. Examination of available cDNA sequences was also inconclusive since all three of the mouse cDNA sequences and the single chicken cDNA sequence in GenBank have different 5'-termini. The simplest explanation is that no clone, or perhaps only one, accurately represents the true *Ihh* transcript initiation site. Alternatively, the mouse and chicken transcription initiation sites may be separately positioned, due to evolutionary divergence, and two of the three mouse cDNA clones are the products of artefact. In some instances, it has been reported that cDNA clones were isolated from a minor mRNA species and, consequently, did not represent the mRNA population for the genes in question [118]. Lastly, one can not exclude the possibility that the different cDNA clones from mouse and chicken are genuine and may each distinctively represent a transcript that is specifically expressed in its respective tissue, or even strain. It should be noted, however, that only one size transcript (2.6 kb) has been detected on Northern blot [119]. The Berkeley promoter prediction application identified a putative TIS at a position corresponding to -230bp from the human *IHH* ATG. Interestingly, there is one mouse cDNA sequence starts at this exact nucleotide (GenBank accession AK090147).

Another compelling aspect about this candidate TIS is that it has TATA and CCAAT-like sequences positioned within a respectable range. Nonetheless, promoters lacking the TATA sequence exist and they are characterized by a GC-rich sequence. Transcription factors that recognize GC-rich sequences, like Sp-1, may initiate the binding of the transcriptional machinery. The region upstream of *Ihh* exon1 notably contains multiple sites of GC-rich sequence, supporting the possibility that the gene initiates transcription from a TATA-less promoter.

The annotation of the *Ihh* gene revealed that exon length is generally conserved across all species, except for the exon1 coding sequence whose length is dependent on the location of the ATG initiator codon (see figure 16). The function of the first two *Ihh* exons is to encode the 20kDa amino fragment of the protein that confers all Hedgehog signalling activity. Exon3 is responsible for encoding the 25kDa carboxyl fragment, comprising the Hint module, which is responsible for the protein's autoproteolytic activity. As a consequence of having a strongly conserved coding sequence, the structure and function of the Ihh protein is analogous across several species. The only discrepancy lies within the first coding exon of the mouse *Ihh*. Among the mammalian species examined, only the mouse has an additional in-frame ATG located upstream of the conserved translation initiation site. This ATG results from the evolutionary mutation of a cytosine, found in human and rat, to an adenine. All the mouse cDNA clones include both in-frame ATG codons, indicating that either one may potentially initiate translation. In 90% of studied mRNAs, ribosomes initiated translation at the farthest upstream ATG [120]. Nevertheless, there exists certain genes in which inactive ATG codons are situated upstream and in-frame of the translation initiation site [120, 121]. It was determined that the efficiency of functional ATG codons depends on how well the flanking nucleotides

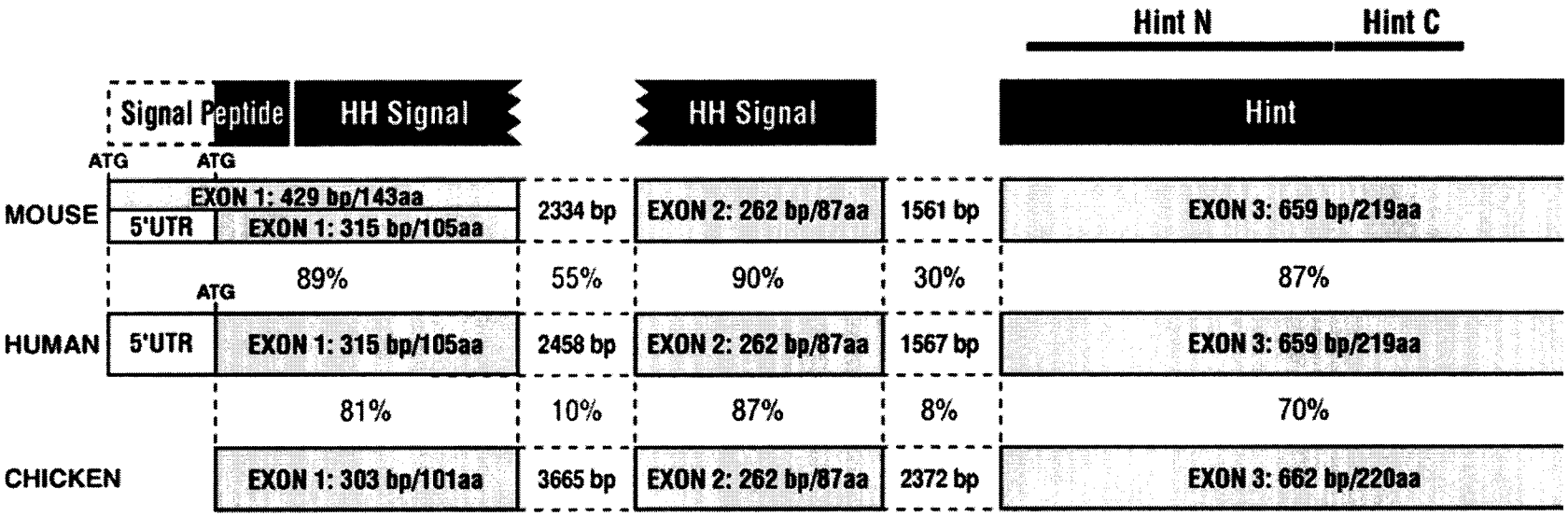


Figure 16: The *Indian Hedgehog* genes of the mouse, human, and chicken. The three exons are shown in grey and their corresponding protein functions in black. The respective exon sizes are indicated in basepairs (bp) and amino acid residues (aa). Note that the mouse *Ihh* exon1 is horizontally bisected to account for the alternate translation start site. The two introns are indicated only by their size and inter-species conservation. The sequence conservation between human and mouse or chicken exons and introns are indicated as a percentage calculated from NCBI Blast sequence alignment (http://www.ncbi.nlm.nih.gov/BLAST/).

match the consensus initiation sequence, 5'-ACC**ATG**G-3', especially in the -3 and +4 positions (N.B.: the A in **ATG** is +1) [120]. According to these criteria, the most efficient starting codon for the mouse *Ihh* would be the downstream ATG, which corresponds to the conserved mammalian translation initiation site [118]. Despite the supporting evidence, nothing can exclude the possibility that the upstream ATG triplet is functional, to some extent, and contributes to a minority of the protein precursor population [120, 121]. In fact, it was hypothesized that the presence of a weak upstream initiation site might function to limit the synthesis of a protein that would be harmful if overproduced [120, 121]. Therefore, the upstream ATG codon in mouse may function as an additional regulatory mechanism to modulate the level of secreted Ihh. Regardless of what occurs *in vivo*, the post-translational modification of both precursor isoforms yields identical mature Ihh proteins since the variably sized signal peptide is cleaved at a conserved site.

5.2. Promoter Analysis

The Primer Extension and 5'RACE assays were used to attempt to identify the *Ihh* transcription initiation site. Neither assay produced conclusive results, and all the necessary precautions were taken to avoid confounding factors. The possibility of experimental error is precluded by successful positive controls. In addition, the quality of the RNA samples was monitored and several different preparations were assayed to eliminate the likelihood that RNA integrity was the source of the problem. All the primers were selected according to the manufacturer's guidelines. Furthermore, the primers designed for Primer Extension together spanned 1kb upstream of the *Ihh* ATG to compensate for the possibility of the gene having an unusually long 5'UTR. The primer radioactive labelling procedures worked well, producing probes with high specific

activities. Considering the GC-rich content of the *Ihh* 5'UTR, one can speculate that the presence of secondary structure may be the cause of the analytical shortcomings.

In order for translation to occur successfully, the size of a gene's 5'UTR must be within an optimal range. A 5'UTR that is excessively short introduces complications during the initiation of translation [118]. For instance, ribosomes are unable to efficiently bind to a short sequence flanking the ATG, resulting in poor translation. Conversely, naturally long 5'UTRs are usually GC-rich, which increases the likelihood of secondary structure formation in the transcript [118]. However, the physical length of an excessively long 5'UTR is proposed to mechanically destabilize non-covalent molecular interactions within the transcript, thereby counteracting the secondary structure introduced by a GC-rich sequence [118]. Presumably, the problems experienced during the promoter analysis assays stem from secondary structure. In fact, the difficulties encountered during the sequencing of the chicken *Ihh* 5'UTR were likely attributable to inhibitory secondary structures caused by density of CpG di-nucleotides in the promoter region.

5.3. Real-time PCR

Analysis of *Ihh* expression in ATDC5 cells revealed that there is a 24-fold increase after 21 days in culture (see figure 9). This conforms with the observation that these cells acquire a mature chondrogenic phenotype when left to differentiate in a supplemented medium [110]. As immature fibroblastic-like ATDC5 differentiate into mature chondrocyte-like cells, they upregulate *Ihh* expression as they begin to hypertrophy similar to that during chondrocyte differentiation in the growth plate [43]. In addition, adult kidneys strongly express *Ihh* [11], as do COS7 and C3H10T1/2 cells, which was confirmed through real-time PCR (see figure 9). In fact, the elevated level of

Ihh expression measured in COS7 cells may be due to the fact that this cell line was derived from simian kidney cells [122]. On the other hand, it is not clear why the mesenchymal-like C3H10T1/2 cells express *Ihh* at high levels.

5.4. ATDC5 Transient Transfections

Our laboratory is interested in the expression and function of *Ihh* within the context of bone development. Therefore, the ATDC5 chondrogenic cell line was used to study the transcriptional control of *Ihh in vitro*. In theory, the inducible chondrogenic phenotype of ATDC5 cells provided the most suitable model for the characterization of the *Ihh* regulatory elements in chondrocytes. ATDC5 mature (d21) cultures were transfected with the mouse *Ihh-luciferase* constructs to determine which plasmids would respond to conditions known to induce *Ihh* transcription, while ATDC5 baseline (d0) cultures were transfected to indicate the basal level of activity of the system. In comparison to the d0 cultures, ATDC5 d21 cultures were expected to respond strongly to the transfected constructs. Instead, the experimental outcomes were influenced by technical difficulties. Mature cultures that were transfected using the Lipofectamine reagent poorly incorporated the mouse *Ihh-luciferase* constructs, thereby compromising experimental data (see appendix B). Alternatively, another transfection reagent, FuGENE6, was tested on the mature cultures. FuGENE6 transfected mature ATDC5 cells with high efficiencies. However, using this reagent, transfection efficiencies for the negative control plasmid (pGL3-basic) were low in mature d21 cultures in comparison to immature d0 cultures, producing an artificially high d21 baseline (see appendix B). Consequently, following normalisation, the relative *luciferase* expression for plasmid p1879B was diminished in d21 ATDC5 cells (see figure 10). Had the values for the

negative control plasmid remained constant, there would have been a 5-fold induction of *luciferase* expression over d0 ATDC5 cells.

Nevertheless, ATDC5 d0 cells sufficiently express *Ihh* in order to study transcriptional regulation of the gene in response to genetic manipulations. Using the data from ATDC5 d0 transient transfections from figure 13, distinct regions in the 1879 *Ihh* upstream fragment were assigned specific regulatory activities (see figure 17). Firstly, since luciferase activity of the p1879B-del1 construct fell to baseline levels, it can be assumed that there is enhancer activity in region A. Furthermore, the strong induction of *luciferase* expression with plasmid p1879B-del3 suggests that there is enhancer activity in region D, which actually may extend into region E since p1879B-del4 is also associated with induced *luciferase* activity but to a lesser extent. Lastly, the suppressed expression of *luciferase* observed with plasmids p1879B-del(2 and 3) in relation to p1879B-del3 can be reconciled by the presence of an antagonist of region D in regions B and C.

Ultimately, the ATDC5 chondrogenic *in vitro* system proved to be unrepresentative of *Ihh* transcriptional regulation *in vivo*. For example, while the transfection efficiencies, assessed by renilla luciferase activity, remained consistent for each replicate of a given plasmid, they varied to a large extent from one plasmid to the next. This discrepancy made it impossible to accurately compare firefly luciferase activity levels between plasmids. Most of the difficulties experienced with transfection efficiencies in mature cultures are thought to be due to the excessive extracellular matrix surrounding the cells. Even after trypsinization, ATDC5 cells from mature cultures were associated with residual cartilaginous matrix which, in turn, might have impeded plasmid incorporation into the cell during transfection. In addition, the ATDC5 model was inconvenient because of the 21-day culture incubation required for every experiment.

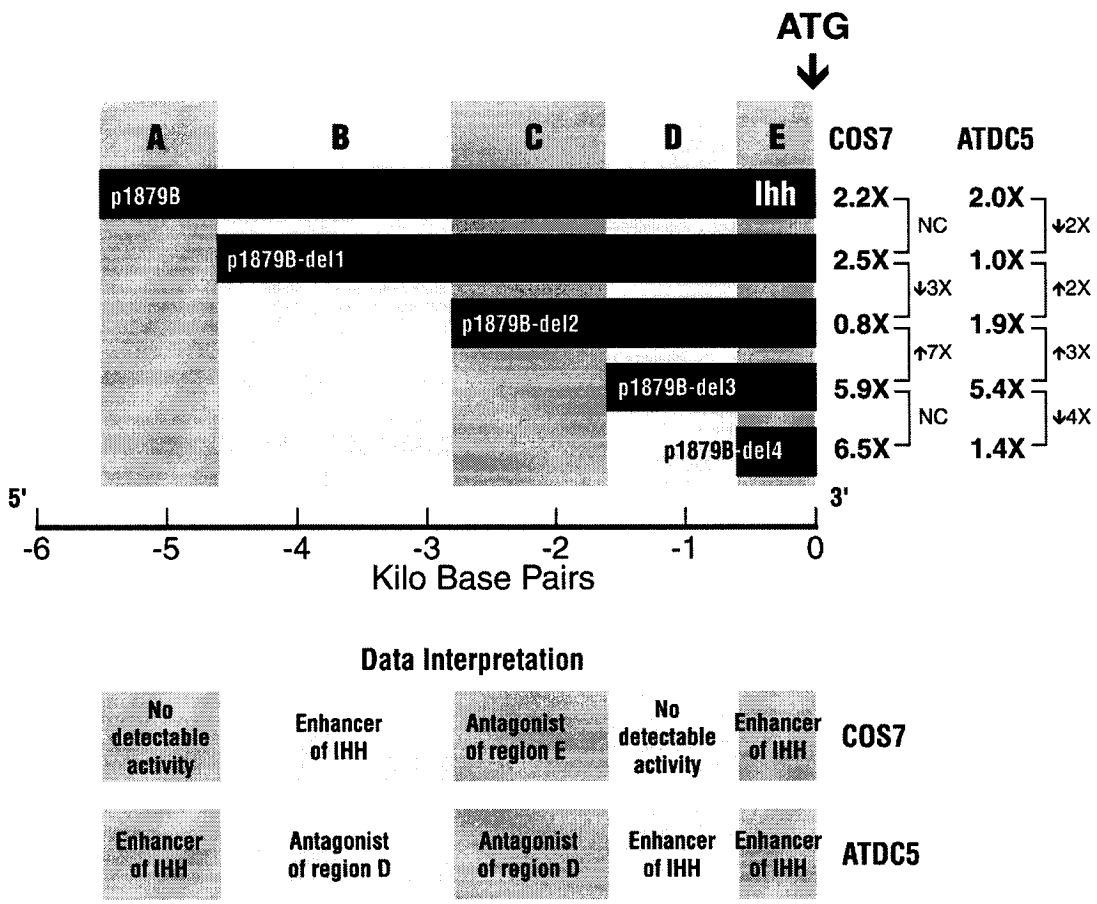


Figure 17: An overview and interpretation of the results of from the transfected p1879B deletion constructs. The regions between the deleted segments have been labelled regions A-E for simplicity. *(Upper panel)* A summary of the data obtained from the COS7 and ATDC5 transient transfections with the p1879B construct and derived deletion constructs. The induction of *luciferase* expression as compared to the negative control plasmid (pGL3-Basic) is indicated for both cell lines as well as the relative expression following each deletion. *(Lower panel)* Data interpretation inferred from the analysis of the *luciferase* expression in each cell line.

5.5. ATDC5 Stable Transfections

The difficulties encountered during the transient transfection of mature d21 ATDC5 cells were unanticipated. Nonetheless, the concept of measuring the induction of *luciferase* expression in mature cultures that have upregulated *Ihh* expression is valid. Stable transfections provided an alternative approach for accomplishing this goal while simultaneously avoiding the problematic delivery of the *Ihh-luciferase* constructs into mature ATDC5 cells. This type of assay proved to be most effective for ATDC5 manipulations with respect to the 50-fold luciferase induction that was observed from a p1879B stably transfected clone.

However, only one of ten stably transfected clones demonstrated any luciferase expression. The remaining nine clones were statistically insignificant upon comparison to their baseline counterparts. There are three possible reasons to explain this occurrence. Firstly, the integrated DNA could have been mutated or excised. Secondly, the neomycin resistance plasmid may have integrated without the luciferase plasmid. Lastly, the plasmid of interest may have integrated into a silenced part of the ATDC5 genome.

Ultimately, the ATDC5 *in vitro* chondrogenic system works best through stable transfections. This approach is more demanding but should be the strategy of choice for future experimentations.

5.6. COS7 Transient Transfections

The COS7 cell line was used to study *Ihh* transcriptional regulation based on its ability to also induce *luciferase* expression in transient transfection assays with *Ihh-luciferase* constructs. The cell line overcame the difficulties encountered using ATDC5

cells and it proved to be a robust *in vitro* system with respect to its time-effectiveness, reliability, and reproducibility.

The deletion clones served to identify segments within the 5' untranslated region of the *Ihh* gene that may contain enhancer or silencer activity. Despite the fact that these deletion clones are too large to specifically identify individual cis-regulatory elements, general scenarios can be deduced from the results obtained during the transfections (refer to figure 17). Upon inspection, it is logical to assume that there is no significant transcriptional regulatory activity in region A since the activity of p1879B-del1 is indiscernible from that of p1879B. On the other hand, a strong silencer region may reside in region C due to the 7-fold increase in transcriptional activity that is observed upon deleting this region from p1879B-del2. In a related matter, the three fold drop in *luciferase* expression that is observed when region B is deleted suggests that this segment may counteract the repressor activity of region C. Region E seemingly contains enhancer activity since the p1879B-del(3 and 4) constructs are equally associated with induced *luciferase* expression. Alternatively, the enhanced expression observed with the latter constructs may be solely due to the absence of region C, supporting the claim that it harbours silencer activity. Evidently, in order to confirm these observations, supplementary deletion constructs will have to be tested to refine the assignment of specific activities to distinct locations.

In order to investigate the postulated interaction between Fgfr3 and *Ihh*, transfected COS7 cells were treated with recombinant FGF9 protein. The mutation of the putative Stat binding domains provides an opportunity to disrupt the direct interaction between Fgfr3 and *Ihh*, should it exist. Treatment of COS7 cells with FGF9 would accentuate any regulatory effect of Fgfr3 over *Ihh* expression since FGF9 preferentially

binds Fgfr3 [112]. The induction of *luciferase* expression with the p1879B construct was expected to be lower in COS7 cells subsequent to FGF9-treatment. In contrast, the transcriptional activity of the p1879B-gap(1 and 2) constructs was expected to be less affected, if at all, by the treatment due to the removal of a Stat binding site in either construct. However, the transcriptional activity of the p1879B construct was not affected by the presence of FGF9, suggesting that treatment of COS7 cells does not affect *Ihh* expression in a similar manner as predicted from the chondrogenic cultures (St-Jacques, unpublished). Considering that the COS7 cell line originates from simian kidney cells, it is not surprising that they behave differently than does a chondrogenic cell line. However, the counterintuitive increase in response to the p1879B-gap1 and p1879B-del4 constructs may be indicative of a novel mode of interaction between Fgf signalling and *Ihh* expression, which is possibly specific to COS7 cells. Nevertheless, any significant findings obtained using COS7 cells can always be compared with results from ATDC5 stably transfected clones.

In concordance with a variant effect of Fgfs on gene expression, Iwata *et al.* (2000) persuasively argued that Fgfr3 may serve as an activator of bone growth in the early stages of embryonic development and as a negative osteogenic regulator only in later embryonic and post-natal developmental stages [123]. In fact, the *in vitro* experimental design is not ideal for reproducing the temporal dynamics of regulated gene expression and protein interactions that occur *in vivo*. Therefore, certain cultured cells may be incapable of providing the necessary environment to properly study Fgfr3 signalling, and may explain why the results observed in FGF9-treated COS7 cultures deviate from expectations.

Additionally, it was taken for granted that the effects of Fgf9 would primarily be potentiated through Fgfr3. In fact, Fgf9 can also bind Fgfr2 [112]. Hence, Fgfr2 signalling may be responsible for the observed effects in COS7 cells via an unknown response. Alternatively, it is possible that Fgfr3 may not mediate its effects through *Ihh*, but rather directly on cells or even through another intermediate.

5.7. Future Directions

One topic that remains controversial is the location of the *Ihh* transcription initiation site. There are different types of approaches, such as RNase Protection, which were not attempted that may provide a solution to this problem. Even though the NNPP program predicted the location of a putative transcription initiation site, it must be experimentally confirmed.

As much as the deletion constructs were useful to map out some regulatory regions within the 5'UTR of *Ihh*, additional deletions of the p1879B construct will need to be engineered and tranfected into COS7 cells to confirm the current findings. A series of constructs bearing progressively smaller deletions are required to precisely identify enhancers, silencers, and sequences that are imperative for promoter activity, along with their respective locations.

The Stat mutation constructs, though informative, also include flanking sequences that may have confounded the transfection data. In order to more accurately investigate the activity of the putative Stat binding sites in the promoter region of *Ihh*, an Electrophoretic Mobility Shift Assay (EMSA) would need to be performed with the Stat transcription factors. If Stat binding is confirmed, then additional *Ihh* constructs can be engineered using *In Situ* Mutatgenesis to introduce specific single-nucleotide mutations

into each of the Stat binding sites. Alterations in the binding activity of these mutated sites could then be verified through EMSA. The transcriptional activity of these constructs can subsequently be quantified through transient transfections with COS7 cells, both untreated and FGF9-treated.

Finally, the 1873, 1875, and 1881 fragments that span the *Ihh* gene and downstream non-coding region (see figure 5) must be subcloned into pGL3-Promoter vectors and transfected to test for enhancer activity. There are certain segments of sequence within these fragments that are conserved between human, mouse, and chicken (see appendix A) and could potentially regulate gene transcription. Regulatory activity can be confirmed through COS7 transient transfections and manipulations similar to those previously described could be used to isolate precise regulatory sites.

6. Conclusion

Currently, there is very little published data regarding the transcriptional regulation of *Ihh* [48-50], none of which have demonstrated a definite relation to *IHH in vivo*.

Despite the ambiguity concerning the promoter region, the sequence analysis of *Ihh* was successful at identifying a couple of Stat putative transcription factor binding sites. This finding has considerable impact with regards to the regulation of chondrogenesis and general skeletal growth, especially in conjunction with the recently identified Bmp and Runx2 binding domains in the upstream region of *Ihh*. The putative Stat sites, though their functional relevance remains untested, provide a new focus for *Ihh* transcriptional experimentation.

The establishment of an effective *in vitro* transfection system to study the expression of the *Ihh* gene was challenging and tedious. Manipulations with the ATDC5 chondrogenic cell line, a seemingly ideal model to study the regulation of *Ihh* expression, were hindered by unforeseen technical difficulties. Nevertheless, this cell line produced stable transfectants that successfully achieved their purpose, though not efficiently. Eventually, an effective working system was designed that employed a reliable cell line, COS7, which inherently expressed *Ihh*. The COS7 cell line was used in the vast majority of the transfection assays, which have been used to define the active regulatory regions of *Ihh*. Distinct regions were delineated upstream of the ATG which contain characteristic properties of enhancers and silencers. Nevertheless, upstream segments exhibiting regulatory activity distinct from that in COS7 cells were found in d0 ATDC5 cells suggesting that *Ihh* may be uniquely regulated in various tissues. Based on the current

findings, additional *Ihh-luciferase* constructs can be engineered to better map out the exact location of regulatory elements throughout the entire gene, including the elusive promoter region.

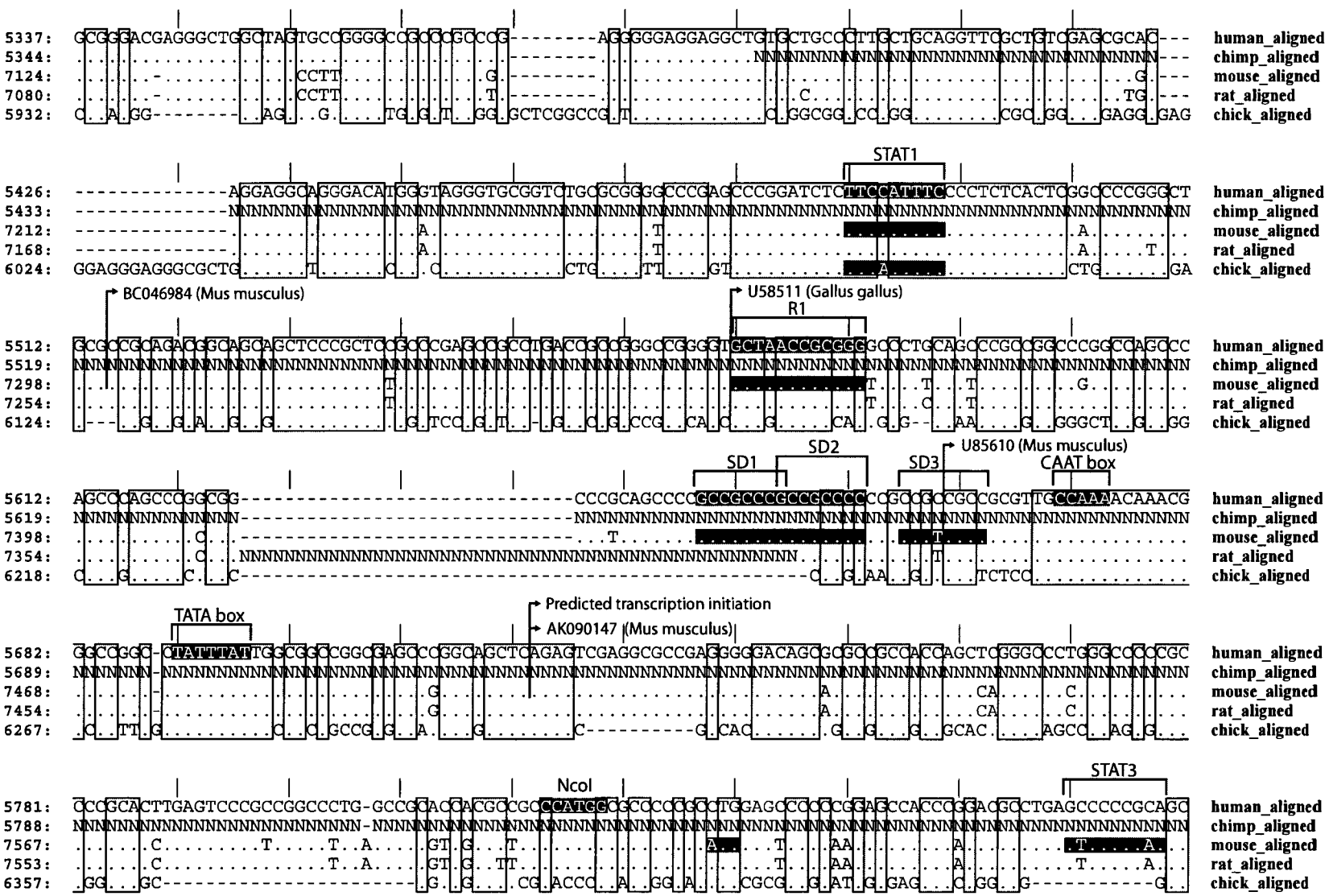
In summary, the experimental progress on the *Ihh* gene is part of a continual effort to identify the signalling pathways involved in the regulation of bone development. A key benefit offered by this project is the introduction of a reliable and reproducible experimental design that is effective for *in vitro* study of the *Ihh* gene. In addition, the data presented contributes towards future experimentation and accomplishing our group's ultimate objectives.

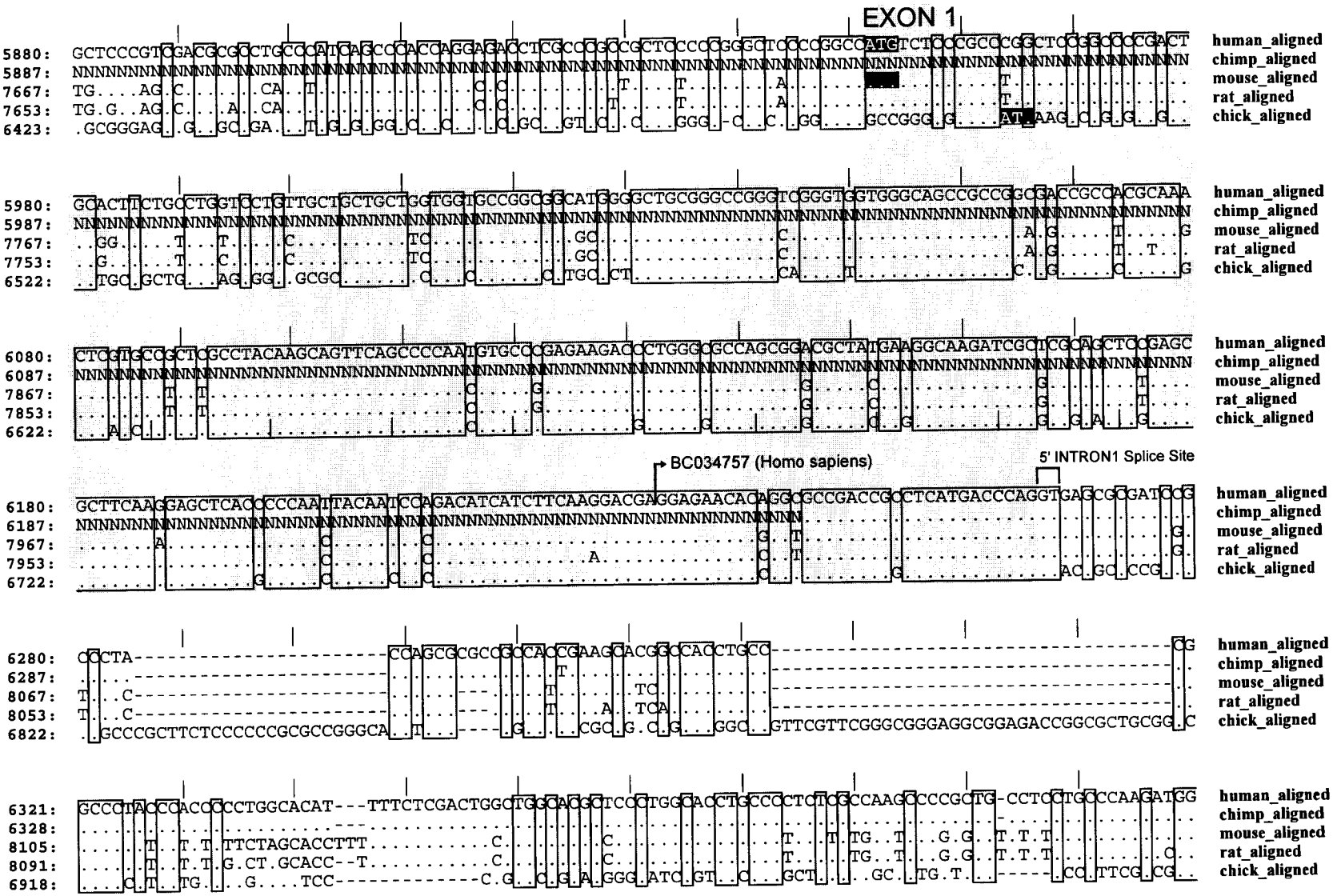
Appendix A

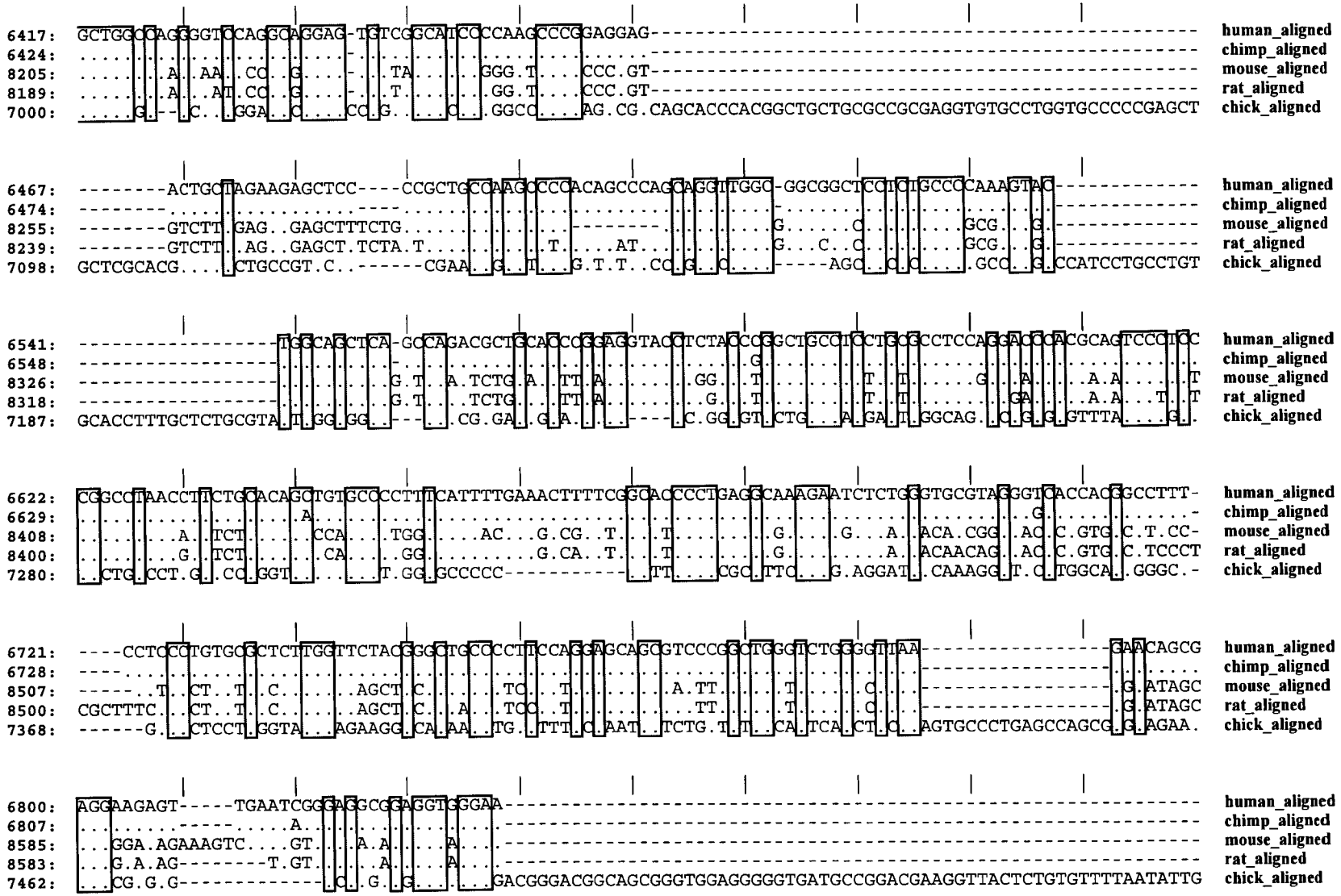
Multiple Species Alignment

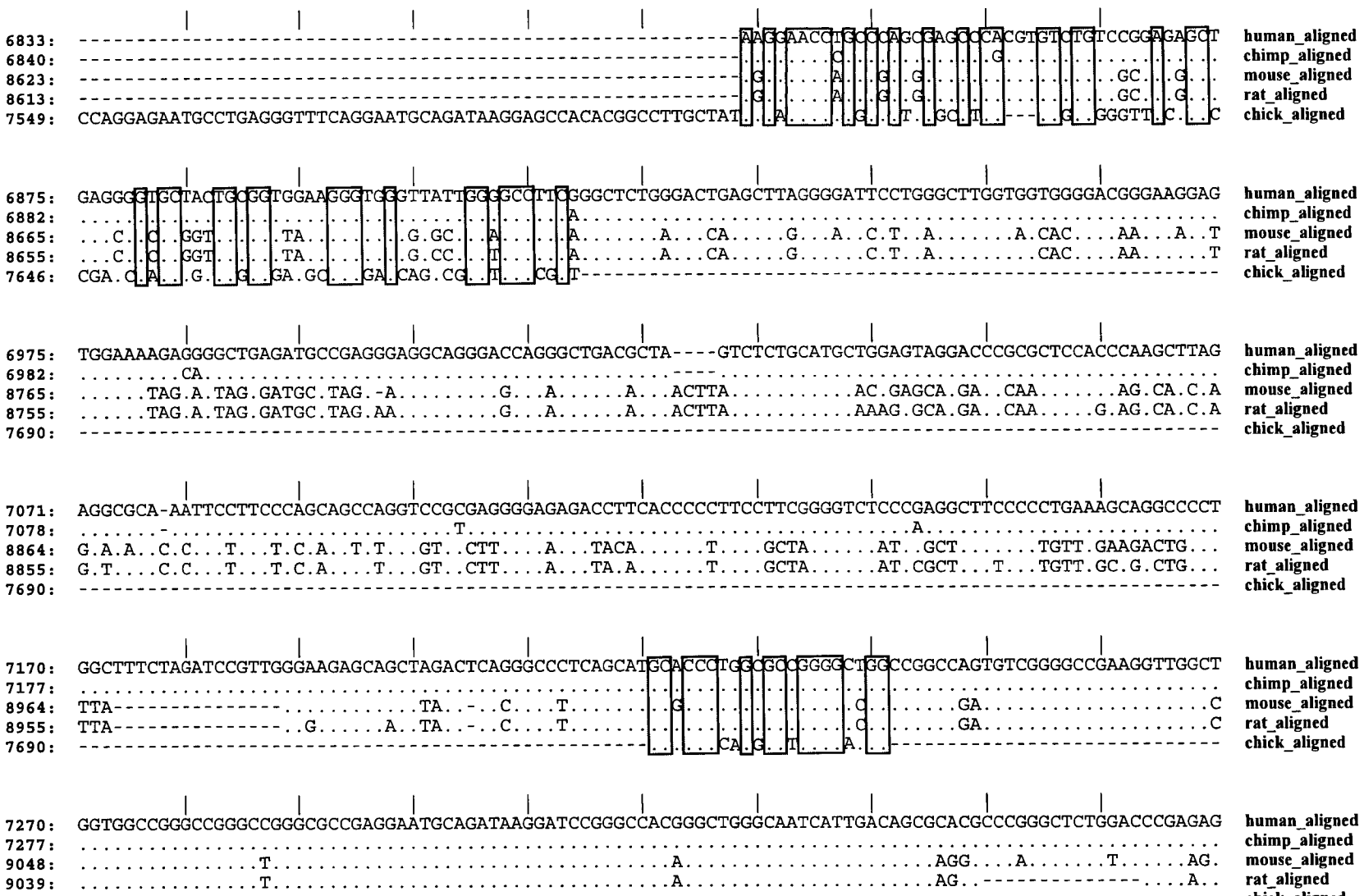
Multiple Species Alignment of the *IHH* gene

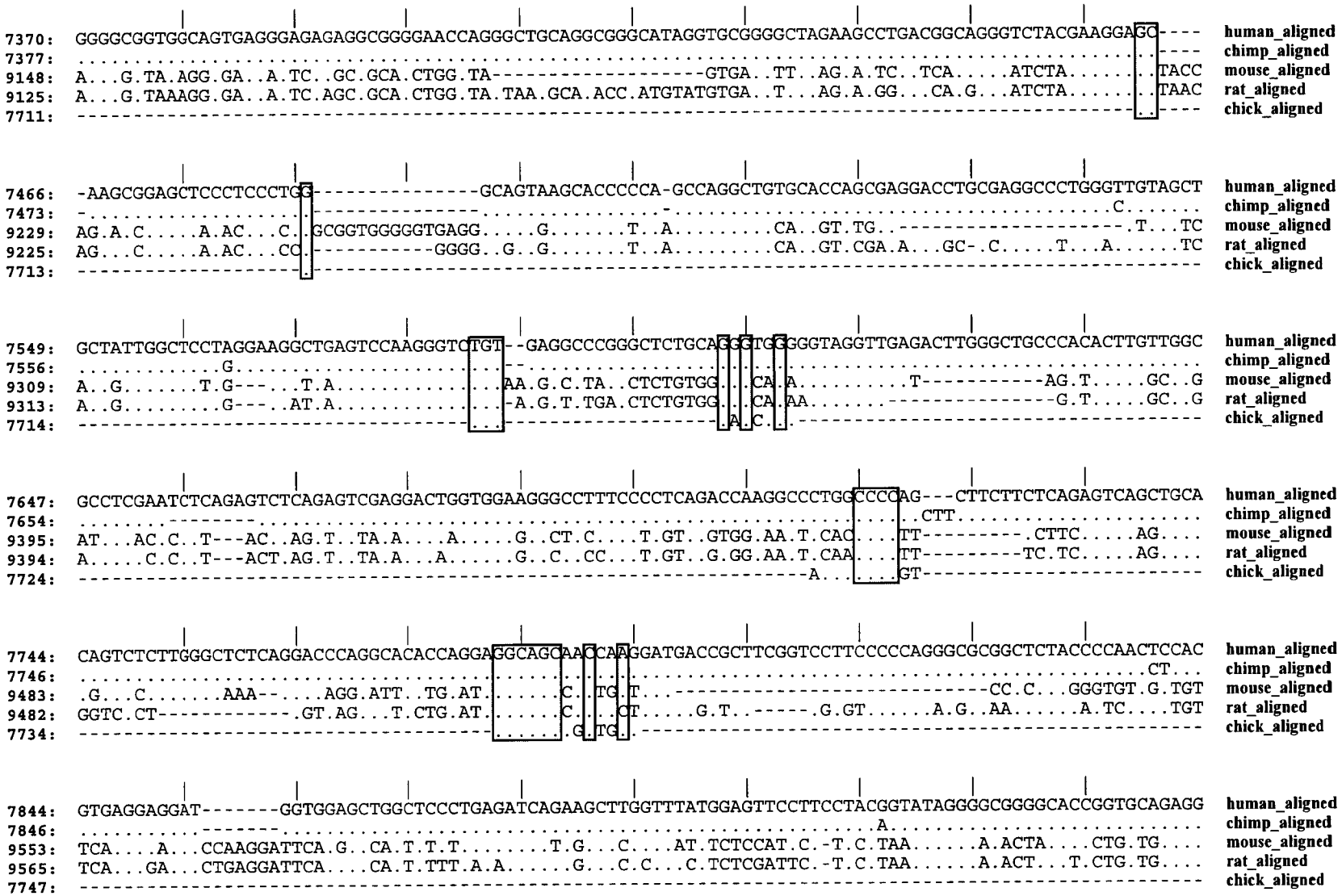
The genomic *Ihh* sequences of human, chimp, mouse, rat, and chicken were aligned to analyze sequence conversation. Despite the fact that 6kb and 4kb of upstream and downstream genomic sequences, respectively, were used in the alignment, the extragenic regions that are not highly conserved between human, mouse, and chicken are excluded from this appendix. The human sequence is used as a reference to compare the sequences from the other species from approximately 600bp upstream of the ATG to 600bp downstream of the TGA. The nucleotides matching those from the human sequences are replaced by a period. A dash is placed at a position where a nucleotide is present in one sequence but absent in another. A box is placed around a segment where the human, mouse, and chicken sequences are identical. The accession numbers for RNA transcripts and the species in which they were cloned are noted at the position of their first nucleotide by an arrow. Furthermore, the NNPP predicted transcription initiation site for human *IHH* is indicated by an arrow at a position that coincides with the first nucleotide of the AK090147 mouse cDNA clone. The *Ihh* exon coding sequences are shaded in grey and are highly conserved across all species. The intron splice sites are also indicated and equally well conserved. All noteworthy sites are labelled and their sequences highlighted in black, including the two mouse ATG sites separated by 114bp, one of the Runx2 binding sites (**R1**) [50], some Smad binding sites (**S1, S2, S3**) [49], and the putative Stat binding sites (**STAT1, STAT3**) containing sequences that were shown to have specific Stat binding activity [115, 116].

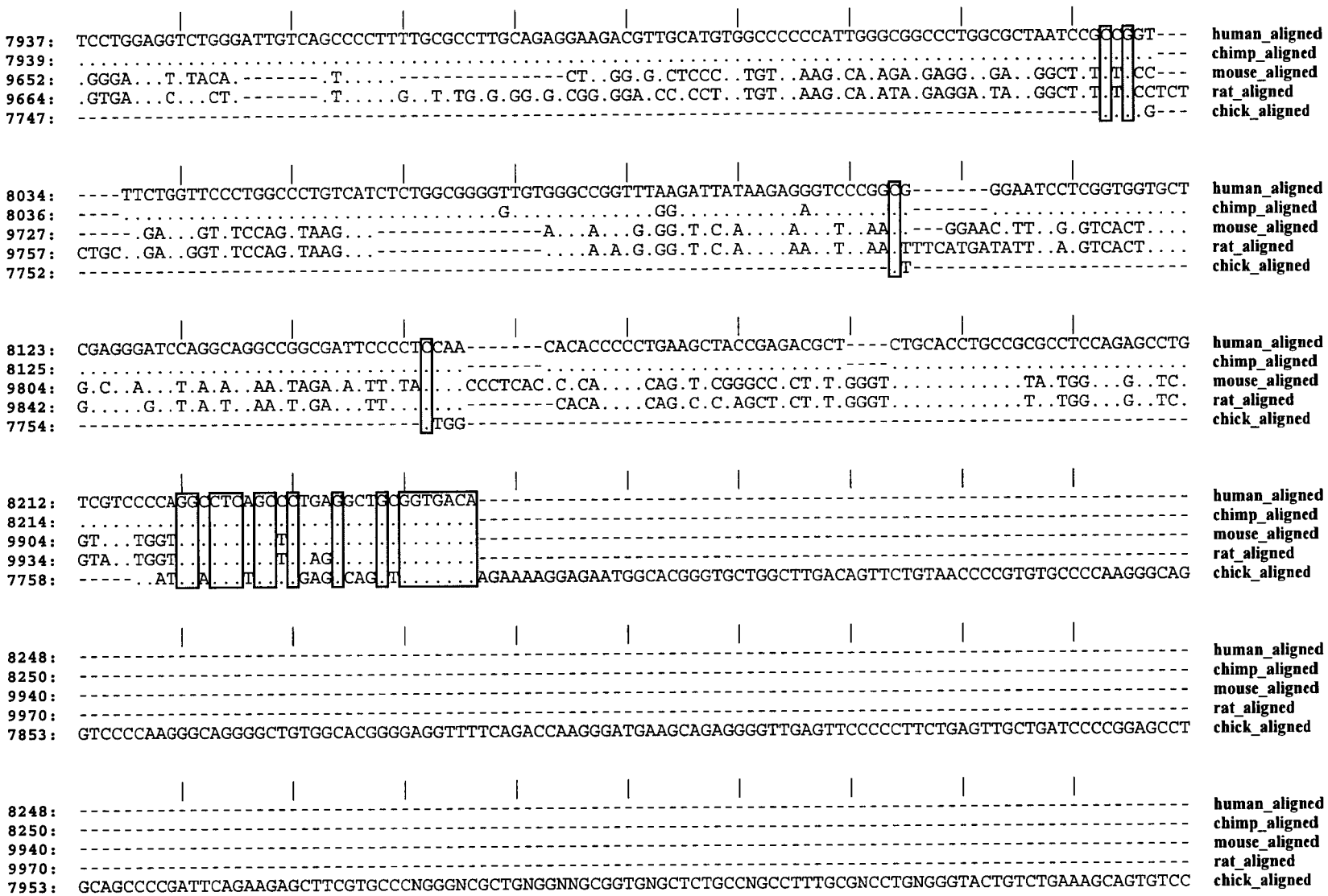

STAT1
BC046984 (Mus musculus)
U58511 (Gallus gallus)
R1
SD1
SD2
SD3
U85610 (Mus musculus)
CAAT box
TATA box
Predicted transcription initiation
AK090147 (Mus musculus)
NcoI
STAT3
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned

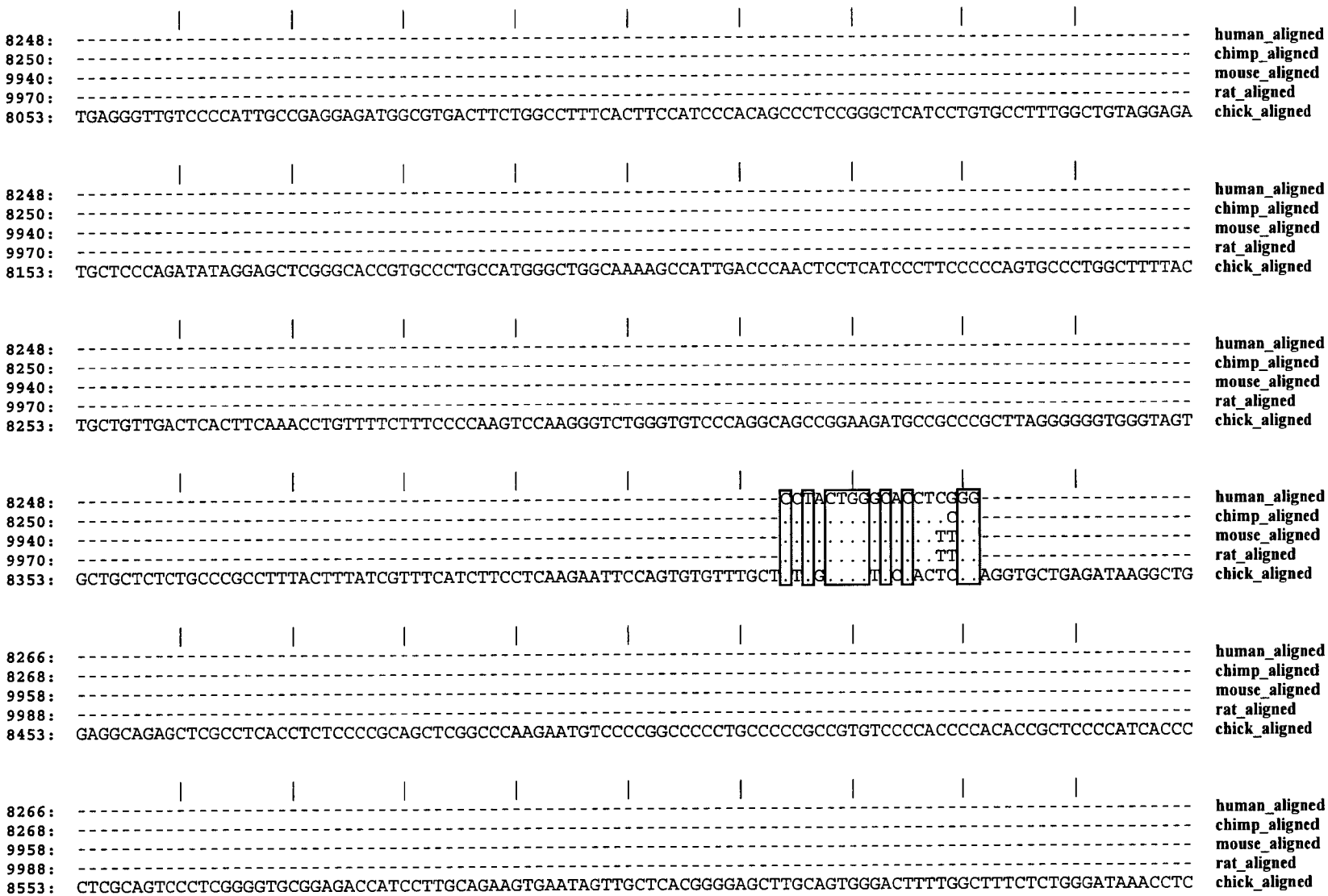

EXON 1
BC034757 (Homo sapiens)
5' INTRON1 Splice Site
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned

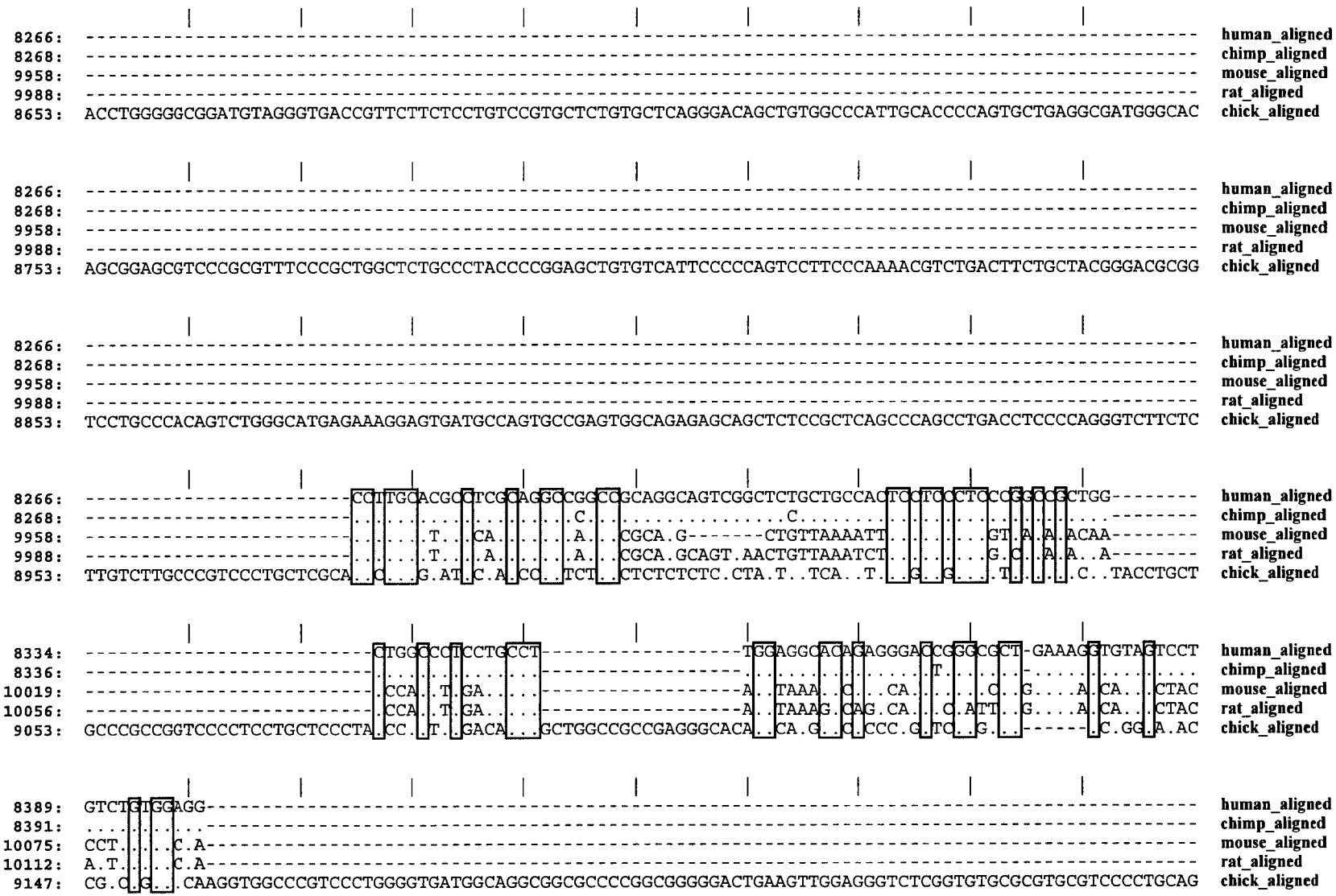

6417: GCTGGCCAGGGGTCCAGGCAGGAG-TGTCGGCATCCCCAAGCCCGGAGGAG human_aligned
6424: chimp_aligned
8205: mouse_aligned
8189: rat_aligned
7000: chick_aligned
6467: --------ACTGCTAGAAGAGCTCC----CCGCTGCCAAGCCCCACAGCCCAGCAGGTTGGC-GGCGGCTCCTCTGCCCCAAAGTAC human_aligned
6474: chimp_aligned
8255: --------GTCTT.GAG..GAGCTTTCTG mouse_aligned
8239: --------GTCTT..AG..GAGCT.TCTA.T rat_aligned
7098: GCTCGCACG....CTGCCGT.C..-------CGAA chick_aligned
6541: human_aligned
6548: chimp_aligned
8326: mouse_aligned
8318: rat_aligned
7187: GCACCTTTGCTCTGCGTA chick_aligned
6622: CGGCCTAACCTTCTGCACAGCTGTGCCCCTTTCATTTTGAAACTTTTCGGCACCCCTGAGGCAAAGAATCTCTGGGTGCGTAGGGTCACCACGGCCTTT- human_aligned
6629: chimp_aligned
8408: mouse_aligned
8400: rat_aligned
7280: chick_aligned
6721: ----CCTCCCTGTGCGCTCTTGGTTCTACGGGCTGCCCCTTCCAGGAGCAGCGTCCCGGCTGGGTCTGGGGTTAA human_aligned
6728: chimp_aligned
8507: mouse_aligned
8500: CGCTTTC rat_aligned
7368: chick_aligned
6800: AGGAAGAGT-----TGAATCGGGAGGCGGAGGTGGGAA human_aligned
6807: chimp_aligned
8585: mouse_aligned
8583: rat_aligned
7462: chick_aligned

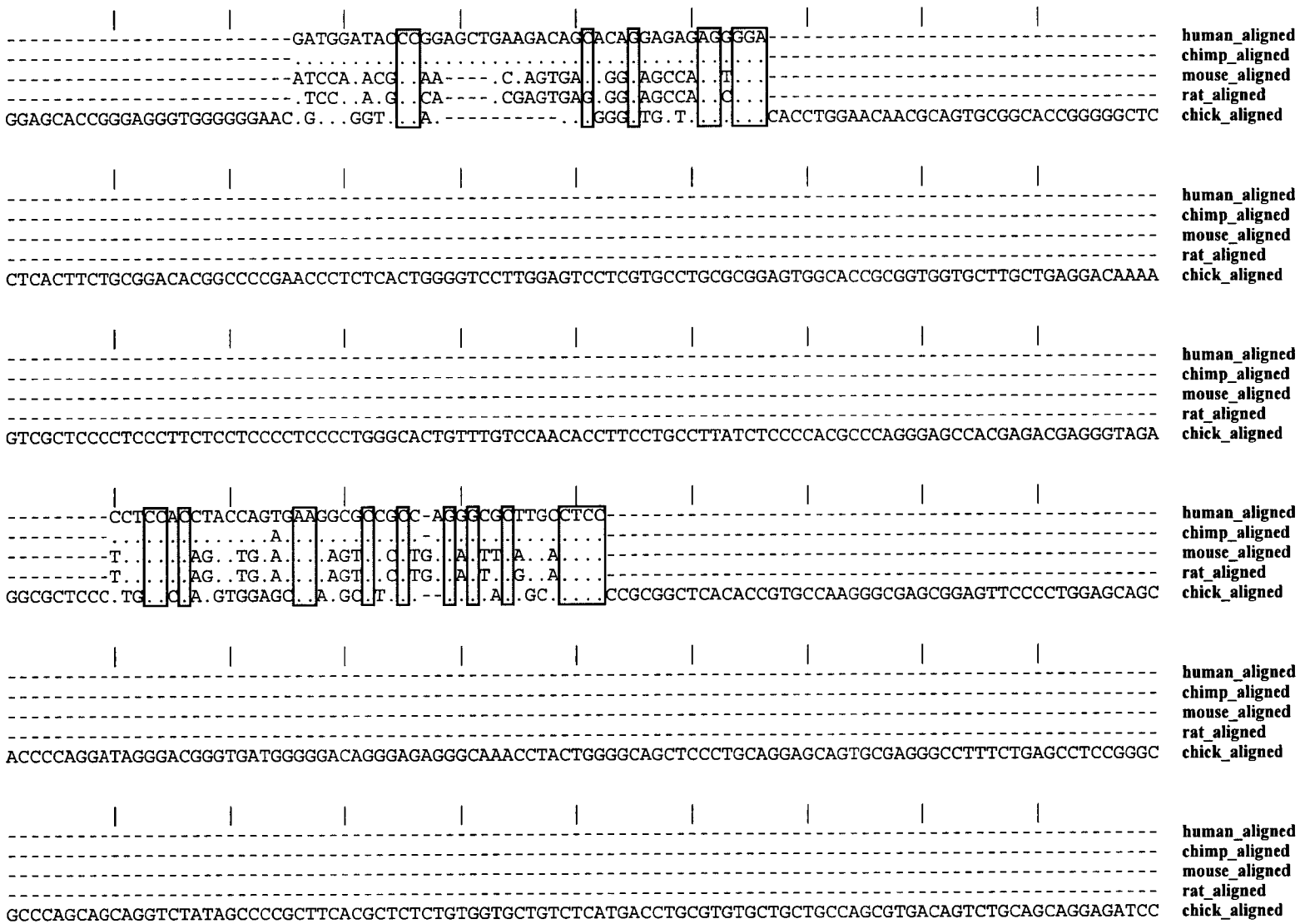

6833: ---------- AAGGAACCTGCCCAGCGAGCCCACGTGTCTGTCCGGAGAGCT human_aligned
6840: ---------- .....C.....G..... chimp_aligned
8623: ---------- .G....A...G..G.....GC...G.. mouse_aligned
8613: ---------- .G....A...G..G.....GC...G.. rat_aligned
7549: CCAGGAGAATGCCTGAGGGTTTCAGGAATGCAGATAAGGAGCCACACGGCCTTGCTAT..A....G..T..GC.T..---..G..GGGTT.C...C chick_aligned
6875: GAGGGGTGCTACTGCGGTGGAAGGGTGGGTTATTGGGGCCTTCGGGCTCTGGGACTGAGCTTAGGGGATTCCTGGGCTTGGTGGTGGGGACGGGAAGGAG human_aligned
6882: .....A..... chimp_aligned
8665: ...C..C..GGT.....TA.....G.GC...A.....A.....A...CA.....G...A..C.T..A.....A.CAC....AA...A..T mouse_aligned
8655: ...C..C..GGT.....TA.....G.CC...T.....A.....A...CA.....G.....C.T..A.....CAC....AA.....T rat_aligned
7646: CGA.C.A...G...G..GA.GC...GA.CAG.CG...T...CG.T---------- chick_aligned
6975: TGGAAAAGAGGGGCTGAGATGCCGAGGGAGGCAGGGACCAGGGCTGACGCTA----GTCTCTGCATGCTGGAGTAGGACCCGCGCTCCACCCAAGCTTAG human_aligned
6982: .....CA.....---..... chimp_aligned
8765: ......TAG.A.TAG.GATGC.TAG.-A.....G...A.....A...ACTTA.....AC.GAGCA.GA..CAA.....AG.CA.C.A mouse_aligned
8755: ......TAG.A.TAG.GATGC.TAG.AA.....G...A.....A...ACTTA.....AAAG.GCA.GA..CAA.....G.AG.CA.C.A rat_aligned
7690: ---------- chick_aligned
7071: AGGCGCA-AATTCCTTCCCAGCAGCCAGGTCCGCGAGGGGAGAGACCTTCACCCCCTTCCTTCGGGGTCTCCCGAGGCTTCCCCCTGAAAGCAGGCCCCT human_aligned
7078: .....-.....T.....A..... chimp_aligned
8864: G.A.A..C.C...T...T.C.A..T.T...GT..CTT....A...TACA.....T....GCTA.....AT..GCT.....TGTT.GAAGACTG... mouse_aligned
8855: G.T....C.C...T...T.C.A....T...GT..CTT....A...TA.A.....T....GCTA.....AT.CGCT...T...TGTT.GC.G.CTG... rat_aligned
7690: ---------- chick_aligned
7170: GGCTTTCTAGATCCGTTGGGAAGAGCAGCTAGACTCAGGGCCCTCAGCATGCACCCTGGCGCCGGGGCTGGCCGGCCAGTGTCGGGGCCGAAGGTTGGCT human_aligned
7177: ..... chimp_aligned
8964: TTA----------.....TA..-..C....T.....G.....C.....GA.....C mouse_aligned
8955: TTA----------..G.....A..TA..-..C....T.....C.....GA.....C rat_aligned
7690: ----------..CA.G..T....A.. chick_aligned
7270: GGTGGCCGGGCCGGGCCGGGCGCCGAGGAATGCAGATAAGGATCCGGGCCACGGGCTGGGCAATCATTGACAGCGCACGCCCGGGCTCTGGACCCGAGAG human_aligned
7277: ..... chimp_aligned
9048: .....T.....A.....AGG....A.....T.....AG. mouse_aligned
9039: .....T.....A.....AG..----------....A.. rat_aligned
7711: ---------- chick_aligned

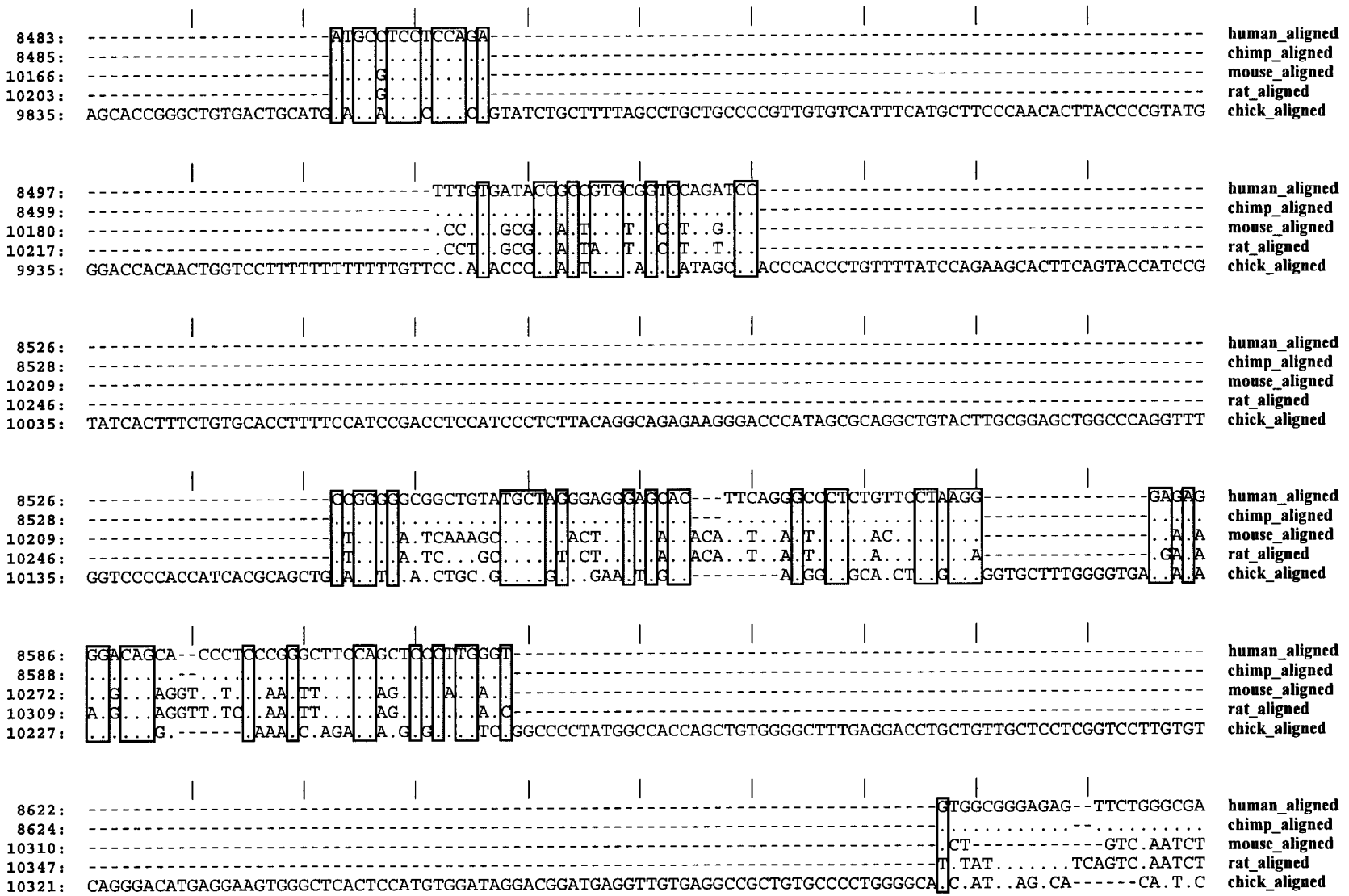

7370: GGGGCGGTGGCAGTGAGGGAGAGAGGCGGGGAACCAGGGCTGCAGGCGGGCATAGGTGCGGGGCTAGAAGCCTGACGGCAGGGTCTACGAAGGAGC---- human_aligned
7377: ..... chimp_aligned
9148: A...G.TA.AGG.GA..A.TC..GC.GCA.CTGG.TA----------GTGA..TT..AG.A.TC..TCA.....ATCTA.....TACC mouse_aligned
9125: A...G.TAAAGG.GA..A.TC.AGC.GCA.CTGG.TA.TAA.GCA.ACC.ATGTATGTGA..T...AG.A.GG...CA.G...ATCTA.....TAAC rat_aligned
7711: ---------- chick_aligned
7466: -AAGCGGAGCTCCCTCCCTGG----------GCAGTAAGCACCCCCA-GCCAGGCTGTGCACCAGCGAGGACCTGCGAGGCCCTGGGTTGTAGCT human_aligned
7473: -..... chimp_aligned
9229: AG.A.C.....A.AC...C..GCGGTGGGGGTGAGG.....G.....T..A.....CA..GT.TG..----------.T...TC mouse_aligned
9225: AG...C.....A.AC...CC.----------GGGG..G..G.....T..A.....CA..GT.CGA.A...GC-.C.....T...A.....TC rat_aligned
7713: ---------- chick_aligned
7549: GCTATTGGCTCCTAGGAAGGCTGAGTCCAAGGGTCTGT--GAGGCCCGGGCTCTGCAGGGTGGGGGTAGGTTGAGACTTGGGCTGCCCACACTTGTTGGC human_aligned
7556: .....G.....-- chimp_aligned
9309: A..G.....T.G----..T.A.....AA.G.C.TA..CTCTGTGG...CA.A.....T----------AG.T.....GC..G mouse_aligned
9313: A..G.....G----..AT.A.....-A.G.T.TGA.CTCTGTGG...CA.AA.....----------G.T.....GC..G rat_aligned
7714: ----------A.C.---------- chick_aligned
7647: GCCTCGAATCTCAGAGTCTCAGAGTCGAGGACTGGTGGAAGGGCCTTTCCCCTCAGACCAAGGCCCTGGCCCCAG---CTTCTTCTCAGAGTCAGCTGCA human_aligned
7654: .....CTT..... chimp_aligned
9395: AT...AC.C..T---AC..AG.T..TA.A....A.....G..CT.C....T.GT..GTGG.AA.T.CAC....TT---------.CTTC.....AG.... mouse_aligned
9394: A.....C.C..T---ACT.AG.T..TA.A...A.....G..C..CC...T.GT..G.GG.AA.T.CAA....TT---------TC.TC.....AG.... rat_aligned
7724: ----------A...GT---------- chick_aligned
7744: CAGTCTCTTGGGCTCTCAGGACCCAGGCACACCAGGAGGCAGCAACCAAGGATGACCGCTTCGGTCCTTCCCCCAGGGCGCGGCTCTACCCCAACTCCAC human_aligned
7746: .....CT... chimp_aligned
9483: .G...C.....AAA-----AGG.ATT..TG.AT.....C..TG.T...----------CC.C...GGGTGT.G.TGT mouse_aligned
9482: GGTC.CT----------.GT.AG...T.CTG.AT.....C.....CT.....G.T.------.G.GT.....A.G..AA.....A.TC....TGT rat_aligned
7734: ----------.....G.TG.---------- chick_aligned
7844: GTGAGGAGGAT-------GGTGGAGCTGGCTCCCTGAGATCAGAAGCTTGGTTTATGGAGTTCCTTCCTACGGTATAGGGGCGGGGCACCGGTGCAGAGG human_aligned
7846: .....A..... chimp_aligned
9553: TCA....A...CCAAGGATTCA.G..CA.T.T.T.....T.G...C....AT.TCTCCAT.C.-T.C.TAA.....A.ACTA....CTG.TG.... mouse_aligned
9565: TCA...GA...CTGAGGATTCA....CA.T.TTT.A.A.....G...C.C...C.TCTCGATTC.-T.C.TAA.....A.ACT...T.CTG.TG.... rat_aligned
7747: ---------- chick_aligned

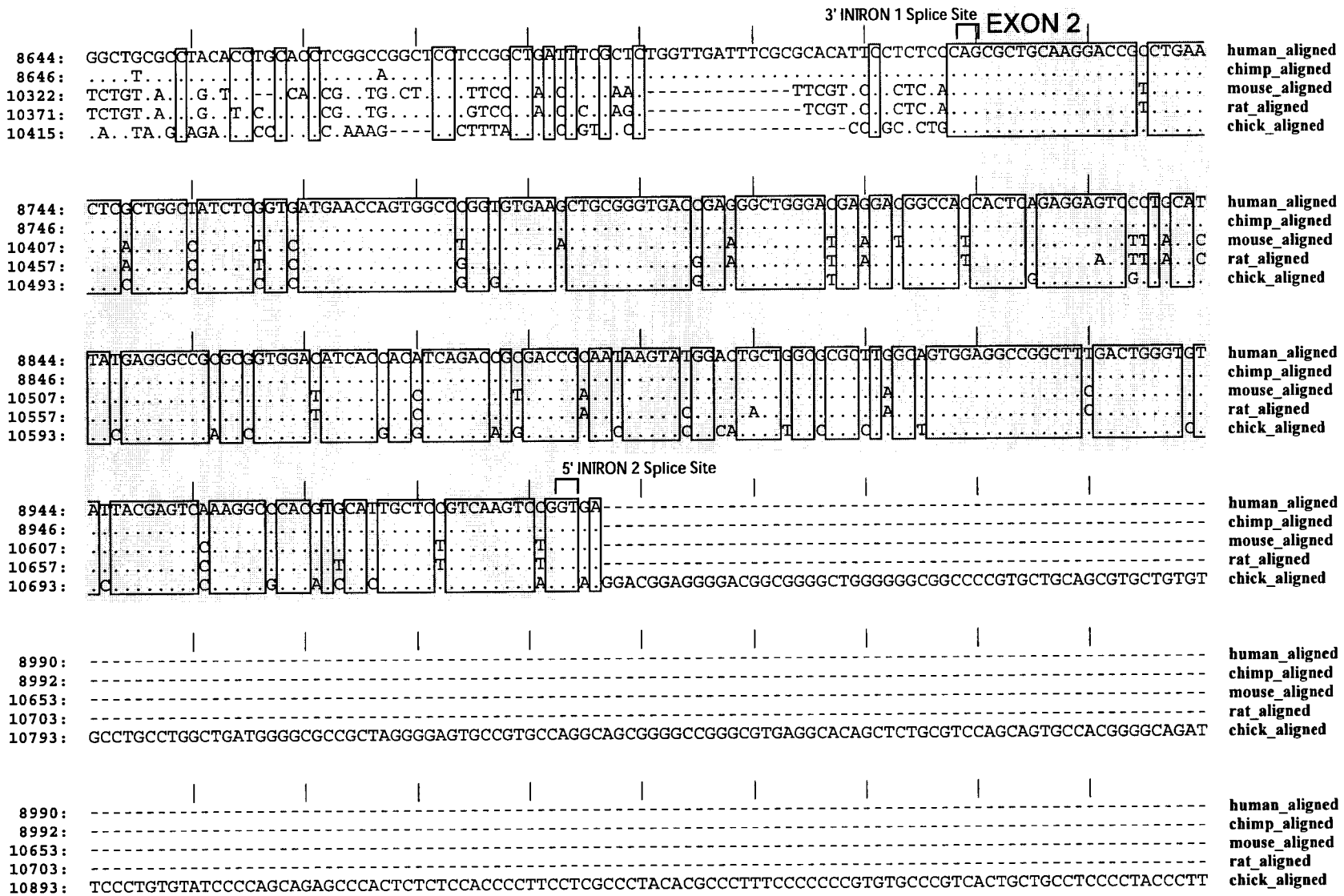

7937: TCCTGGAGGTCTGGGATTGTCAGCCCCTTTTGCGCCTTGCAGAGGAAGACGTTGCATGTGGCCCCCCATTGGGCGGCCCTGGCGCTAATCCGCCCGGT--- human_aligned
7939: ..... chimp_aligned
9652: .GGGA...T.TACA.-------.T.....----------CT..GG.G.CTCCC..TGT..AAG.CA.AGA.GAGG..GA..GGCT.T.T.CC--- mouse_aligned
9664: .GTGA...C...CT.-------.T.....G..T.TG.G.GG.G.CGG.GGA.CC.CCT..TGT..AAG.CA.ATA.GAGGA.TA..GGCT.T.T.CCTCT rat_aligned
7747: ----------.G--- chick_aligned
8034: ----TTCTGGTTCCCTGGCCCTGTCATCTCTGGCGGGGTTGTGGGCCGGTTTAAGATTATAAGAGGGTCCCGGCG-------GGAATCCTCGGTGGTGCT human_aligned
8036: ----.....G.....GG.....A.....---------- chimp_aligned
9727: -----.GA...GT.TCCAG.TAAG...----------A...A...G.GG.T.C.A....A...T..AA..---GGAAC.TT..G.GTCACT.... mouse_aligned
9757: CTGC..GA..GGT.TCCAG.TAAG...----------....A.A.G.GG.T.C.A....AA..T..AA.TTTCATGATATT..A.GTCACT.... rat_aligned
7752: ----------.T---------- chick_aligned
8123: CGAGGGATCCAGGCAGGCCGGCGATTCCCCTCCAA-------CACACCCCCTGAAGCTACCGAGACGCT----CTGCACCTGCCGCGCCTCCAGAGCCTG human_aligned
8125: .....------.....---- chimp_aligned
9804: G.C..A...T.A.A..AA.TAGA.A.TT.TA....CCCTCAC.C.CA....CAG.T.CGGGCC.CT.T.GGGT.....TA.TGG...G..TC. mouse_aligned
9842: G.....G..T.A.T..AA.T.GA...TT.....--------CACA....CAG.C.C.AGCT.CT.T.GGGT.....T..TGG...G..TC. rat_aligned
7754: ----------.TGG---------- chick_aligned
8212: TCGTCCCCAGGCCTCAGCCCTGAGGCTGCGGTGACA---------- human_aligned
8214: .....---------- chimp_aligned
9904: GT...TGGT.....T.....---------- mouse_aligned
9934: GTA..TGGT.....T..AG.....---------- rat_aligned
7758: ------..AT..A...T...GAG.CAG.T.....AGAAAAGGAGAATGGCACGGGTGCTGGCTTGACAGTTCTGTAACCCCGTGTGCCCCAAGGGCAG chick_aligned
8248: ---------- human_aligned
8250: ---------- chimp_aligned
9940: ---------- mouse_aligned
9970: ---------- rat_aligned
7853: GTCCCCAAGGGCAGGGGCTGTGGCACGGGGAGGTTTTCAGACCAAGGGATGAAGCAGAGGGGTTGAGTTCCCCCTTCTGAGTTGCTGATCCCCGGAGCCT chick_aligned
8248: ---------- human_aligned
8250: ---------- chimp_aligned
9940: ---------- mouse_aligned
9970: ---------- rat_aligned
7953: GCAGCCCCGATTCAGAAGAGCTTCGTGCCCNGGGNCGCTGNGGNNGCGGTGNGCTCTGCCNGCCTTTGCGNCCTGNGGGTACTGTCTGAAAGCAGTGTCC chick_aligned

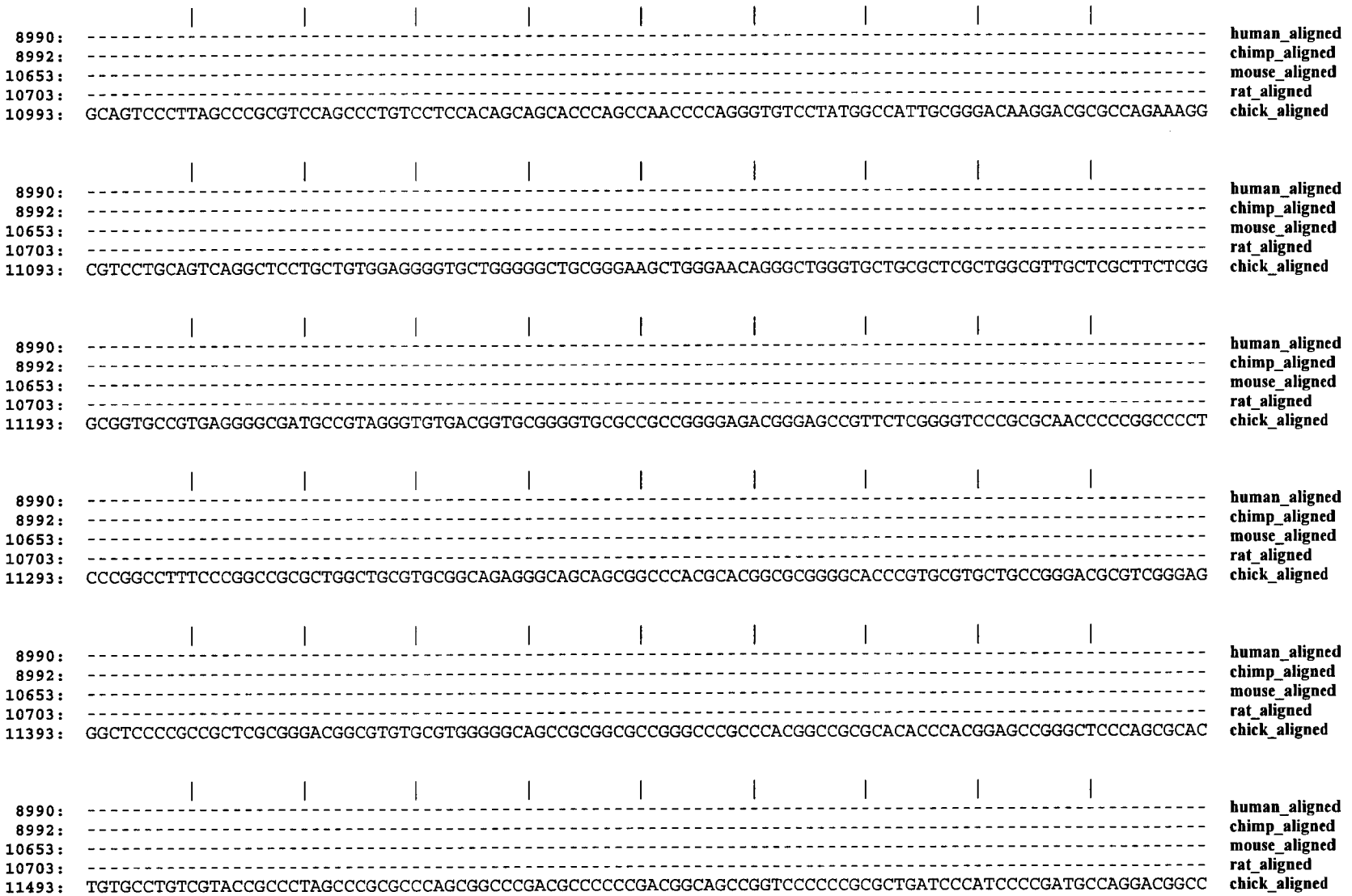

8248: ---------- human_aligned
8250: ---------- chimp_aligned
9940: ---------- mouse_aligned
9970: ---------- rat_aligned
8053: TGAGGGTTGTCCCCATTGCCGAGGAGATGGCGTGACTTCTGGCCTTTCACTTCCATCCCACAGCCCTCCGGGCTCATCCTGTGCCTTTGGCTGTAGGAGA chick_aligned
8248: ---------- human_aligned
8250: ---------- chimp_aligned
9940: ---------- mouse_aligned
9970: ---------- rat_aligned
8153: TGCTCCCAGATATAGGAGCTCGGGCACCGTGCCCTGCCATGGGCTGGCAAAAGCCATTGACCCAACTCCTCATCCCTTCCCCCAGTGCCCTGGCTTTTAC chick_aligned
8248: ---------- human_aligned
8250: ---------- chimp_aligned
9940: ---------- mouse_aligned
9970: ---------- rat_aligned
8253: TGCTGTTGACTCACTTCAAACCTGTTTTCTTTCCCCAAGTCCAAGGGTCTGGGTGTCCCAGGCAGCCGGAAGATGCCGCCCGCTTAGGGGGGTGGGTAGT chick_aligned
8248: ----------CCTACTGGGCACCTCGGG---------- human_aligned
8250: ----------.....C..---------- chimp_aligned
9940: ----------.....TT..---------- mouse_aligned
9970: ----------.....TT..---------- rat_aligned
8353: GCTGCTCTCTGCCCGCCTTTACTTTATCGTTTCATCTTCCTCAAGAATTCCAGTGTGTTTGCT.T.G...T.C.ACTC..AGGTGCTGAGATAAGGCTG chick_aligned
8266: ---------- human_aligned
8268: ---------- chimp_aligned
9958: ---------- mouse_aligned
9988: ---------- rat_aligned
8453: GAGGCAGAGCTCGCCTCACCTCTCCCCGCAGCTCGGCCCAAGAATGTCCCCGGCCCCCTGCCCCCGCCGTGTCCCCACCCCACACCGCTCCCCATCACCC chick_aligned
8266: ---------- human_aligned
8268: ---------- chimp_aligned
9958: ---------- mouse_aligned
9988: ---------- rat_aligned
8553: CTCGCAGTCCCTCGGGGTGCGGAGACCATCCTTGCAGAAGTGAATAGTTGCTCACGGGGAGCTTGCAGTGGGACTTTTGGCTTTCTCTGGGATAAACCTC chick_aligned

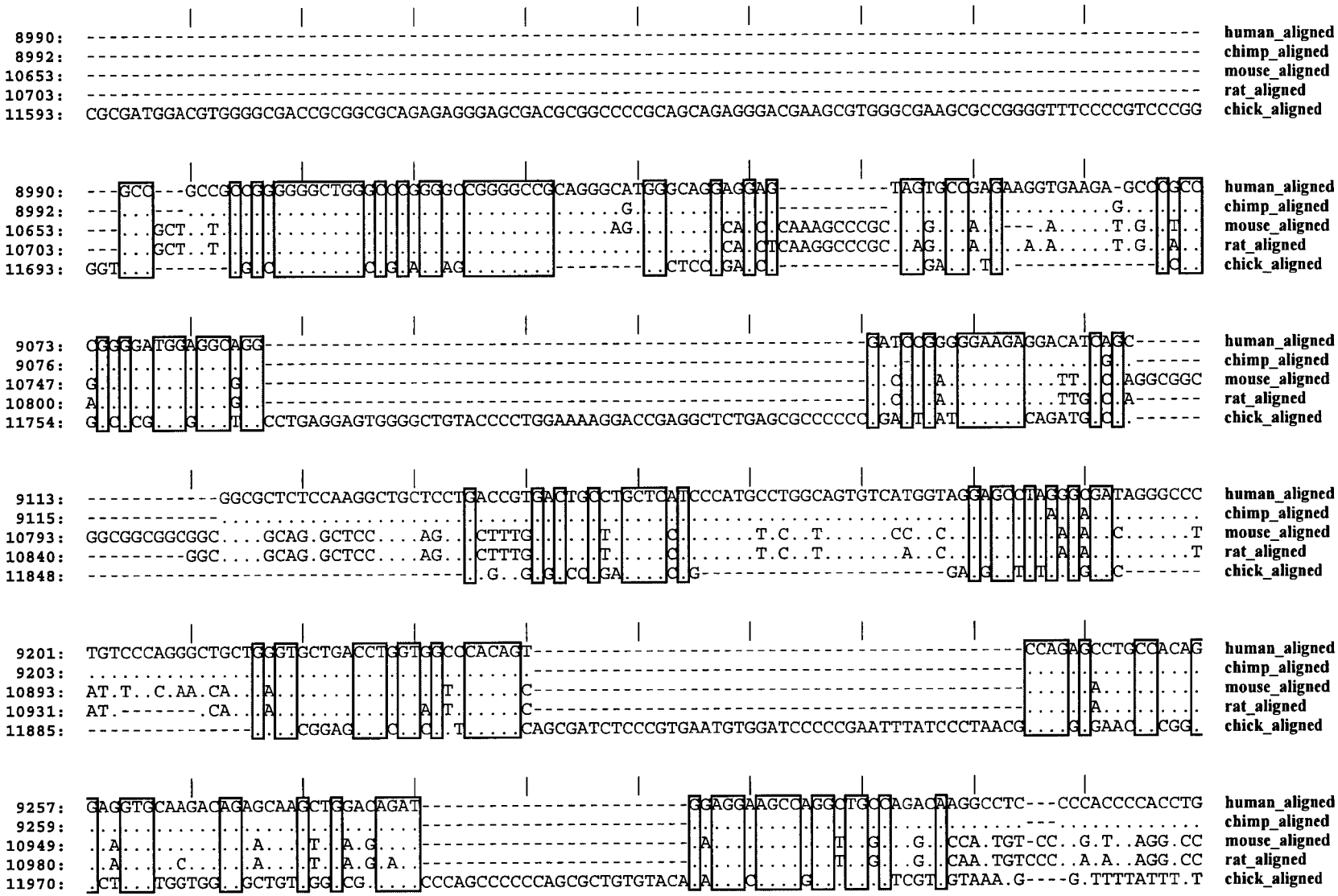

8266: human_aligned
8268: chimp_aligned
9958: mouse_aligned
9988: rat_aligned
8653: ACCTGGGGGCGGATGTAGGGTGACCGTTCTTCTCCTGTCCGTGCTCTGTGCTCAGGGACAGCTGTGGCCCATTGCACCCCAGTGCTGAGGCGATGGGCAC chick_aligned
8753: AGCGGAGCGTCCCGCGTTTCCCGCTGGCTCTGCCCTACCCCGGAGCTGTGTCATTCCCCCAGTCCTTCCCAAAACGTCTGACTTCTGCTACGGGACGCGG chick_aligned
8853: TCCTGCCCACAGTCTGGGCATGAGAAAGGAGTGATGCCAGTGCCGAGTGGCAGAGAGCAGCTCTCCGCTCAGCCCAGCCTGACCTCCCCAGGGTCTTCTC chick_aligned
8953: TTGTCTTGCCCGTCCCTGCTCGCA chick_aligned
9053: GCCCGCCGGTCCCCTCCTGCTCCCTA chick_aligned
9147: CG chick_aligned
CAAGGTGGCCCGTCCCTGGGGTGATGGCAGGCGGCGCCCCGGCGGGGGACTGAAGTTGGAGGGTCTCGGTGTGCGCGTGCGTCCCCTGCAG

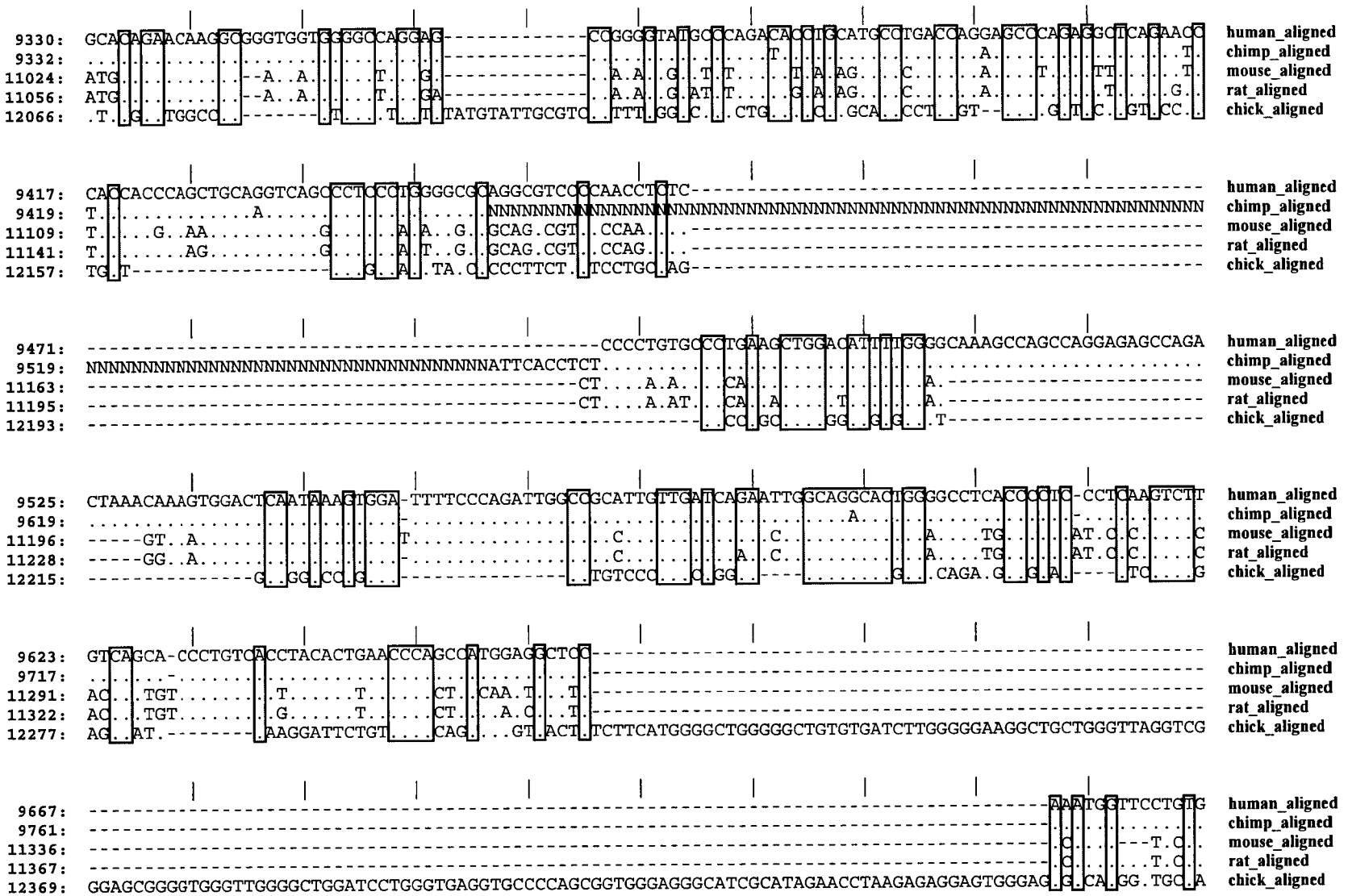

8400: ----------GATGGATACCGGAGCTGAAGACAGCACAGGAGAGAGGGA---------- human_aligned
8402: ----------..---------- chimp_aligned
10086: ----------ATCCA.ACG..AA----.C.AGTGA..GG.AGCCA..T...---------- mouse_aligned
10123: ----------.TCC..A.G..CA----.CGAGTGAG.GG.AGCCA..C...---------- rat_aligned
9247: GGAGCACCGGGAGGGTGGGGGGAAC.G...GGT...A.----------..GGG.TG.T.....CACCTGGAACAACGCAGTGCGGCACCGGGGGCTC chick_aligned
8441: ---------- human_aligned
8443: ---------- chimp_aligned
10123: ---------- mouse_aligned
10160: ---------- rat_aligned
9337: CTCACTTCTGCGGACACGGCCCCGAACCCTCTCACTGGGGTCCTTGGAGTCCTCGTGCCTGCGCGGAGTGGCACCGCGGTGGTGCTTGCTGAGGACAAAA chick_aligned
8441: ---------- human_aligned
8443: ---------- chimp_aligned
10123: ---------- mouse_aligned
10160: ---------- rat_aligned
9437: GTCGCTCCCCTCCCTTCTCCTCCCCTCCCCTGGGCACTGTTTGTCCAACACCTTCCTGCCTTATCTCCCCACGCCCAGGGAGCCACGAGACGAGGGTAGA chick_aligned
8441: ---------CCTCCACCTACCAGTGAAGGCGCCGCC-AGGGCGCTTGCCTCC---------- human_aligned
8443: ---------.....A.....-.....---------- chimp_aligned
10123: ---------T.....AG..TG.A...AGT..C.TG..A.TT.A..A....---------- mouse_aligned
10160: ---------T.....AG..TG.A...AGT..C.TG..A.T..G..A....---------- rat_aligned
9537: GGCGCTCCC.TG..C.A.GTGGAGC..A.GC.T..--....A..GC.....CCGCGGCTCACACCGTGCCAAGGGCGAGCGGAGTTCCCCTGGAGCAGC chick_aligned
8483: ---------- human_aligned
8485: ---------- chimp_aligned
10166: ---------- mouse_aligned
10203: ---------- rat_aligned
9635: ACCCCAGGATAGGGACGGGTGATGGGGGACAGGGAGAGGGCAAACCTACTGGGGCAGCTCCCTGCAGGAGCAGTGCGAGGGCCTTTCTGAGCCTCCGGGC chick_aligned
8483: ---------- human_aligned
8485: ---------- chimp_aligned
10166: ---------- mouse_aligned
10203: ---------- rat_aligned
9735: GCCCAGCAGCAGGTCTATAGCCCCGCTTCACGCTCTCTGTGGTGCTGTCTCATGACCTGCGTGTGCTGCTGCCAGCGTGACAGTCTGCAGCAGGAGATCC chick_aligned

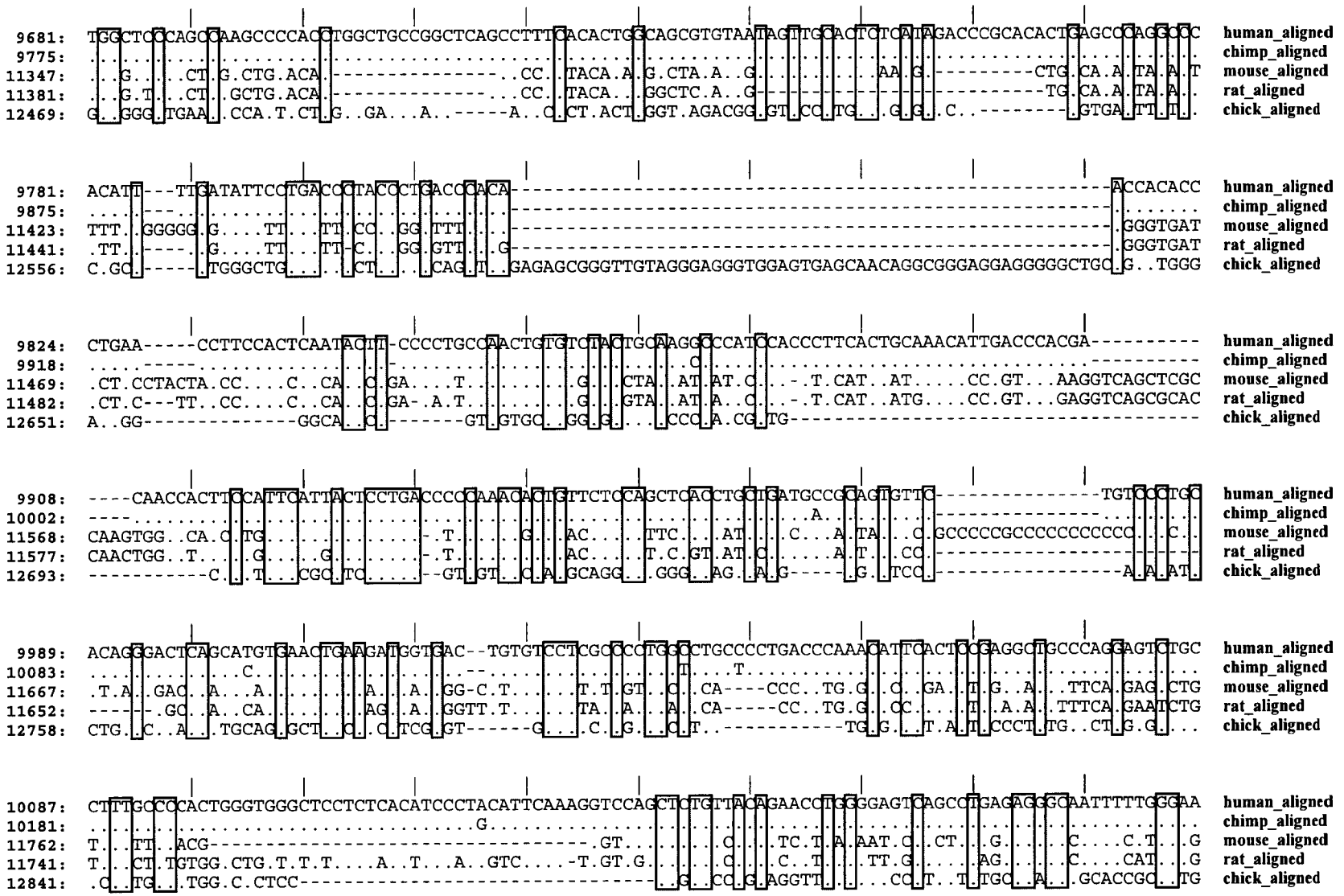

8483: ATGCCTCCTCCAGA human_aligned
8485: chimp_aligned
10166: mouse_aligned
10203: rat_aligned
9835: AGCACCGGGCTGTGACTGCATG GTATCTGCTTTTAGCCTGCTGCCCCGTTGTGTCATTTCATGCTTCCCAACACTTACCCCGTATG chick_aligned
8497: TTTGTGATACCGCCGTGCGGTCCAGATCC human_aligned
8499: chimp_aligned
10180: mouse_aligned
10217: rat_aligned
9935: GGACCACAACTGGTCCTTTTTTTTTTTTGTTCC ACCCACCCTGTTTTATCCAGAAGCACTTCAGTACCATCCG chick_aligned
8526: human_aligned
8528: chimp_aligned
10209: mouse_aligned
10246: rat_aligned
10035: TATCACTTTCTGTGCACCTTTTCCATCCGACCTCCATCCCTCTTACAGGCAGAGAAGGGACCCATAGCGCAGGCTGTACTTGCGGAGCTGGCCCAGGTTT chick_aligned
8526: CCGGGGGCGGCTGTATGCTAGGGAGGGAGGAC TTCAGGGCCCTCTGTTCCTAAGG GAGAG human_aligned
8528: chimp_aligned
10209: mouse_aligned
10246: rat_aligned
10135: GGTCCCCACCATCACGCAGCTG GGTGCTTTGGGGTGA chick_aligned
8586: GGACAGCA CCCTCCCGGGCTTCCAGCTCCCTTGGGT human_aligned
8588: chimp_aligned
10272: mouse_aligned
10309: rat_aligned
10227: GGCCCCTATGGCCACCAGCTGTGGGGCTTTGAGGACCTGCTGTTGCTCCTCGGTCCTTGTGT chick_aligned
8622: GTGGCGGGAGAG TTCTGGGCGA human_aligned
8624: chimp_aligned
10310: GTC.AATCT mouse_aligned
10347: TCAGTC.AATCT rat_aligned
10321: CAGGGACATGAGGAAGTGGGCTCACTCCATGTGGATAGGACGGATGAGGTTGTGAGGCCGCTGTGCCCCTGGGGCA CA------CA chick_aligned

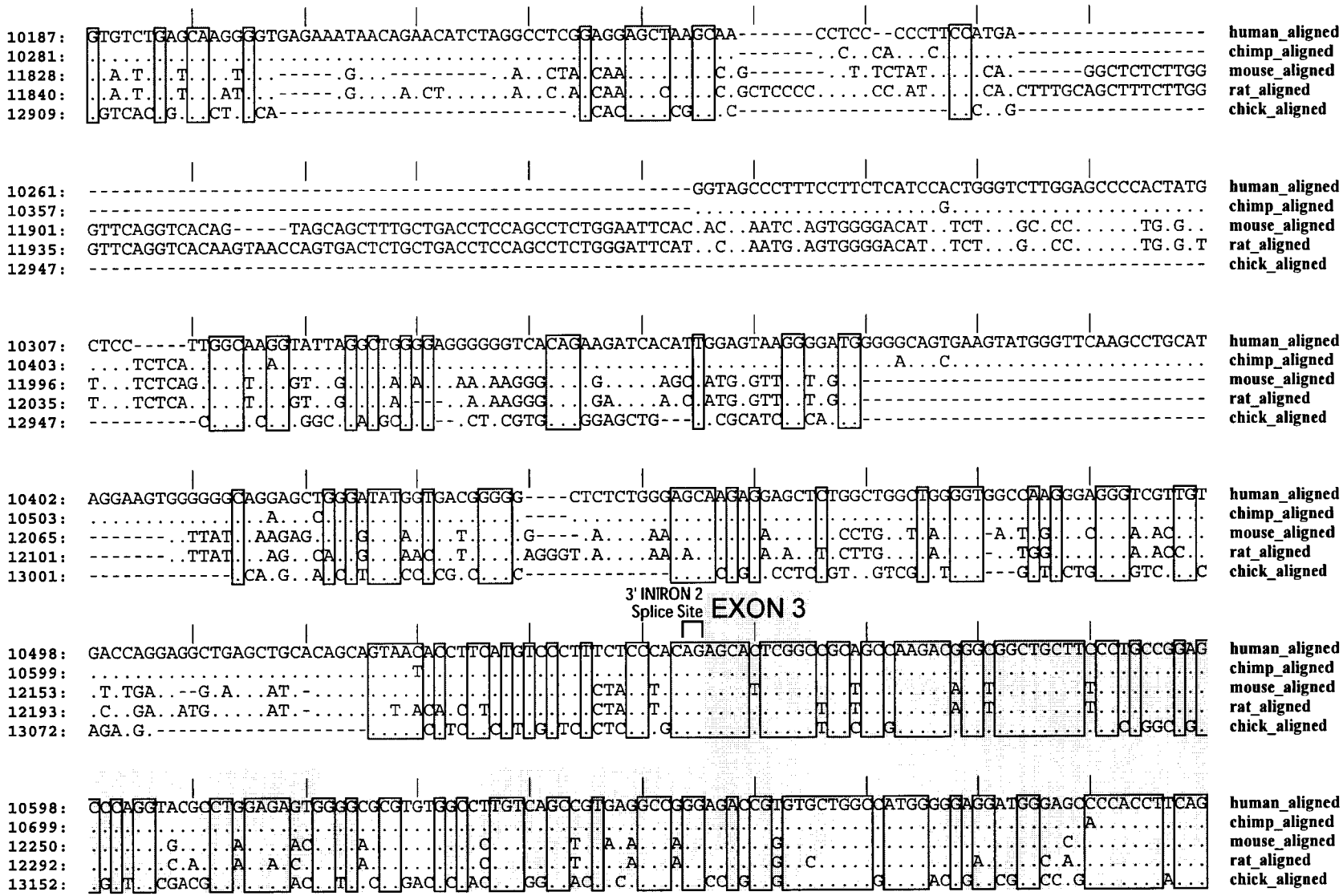

3' INTRON 1 Splice Site
EXON 2
5' INTRON 2 Splice Site
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned

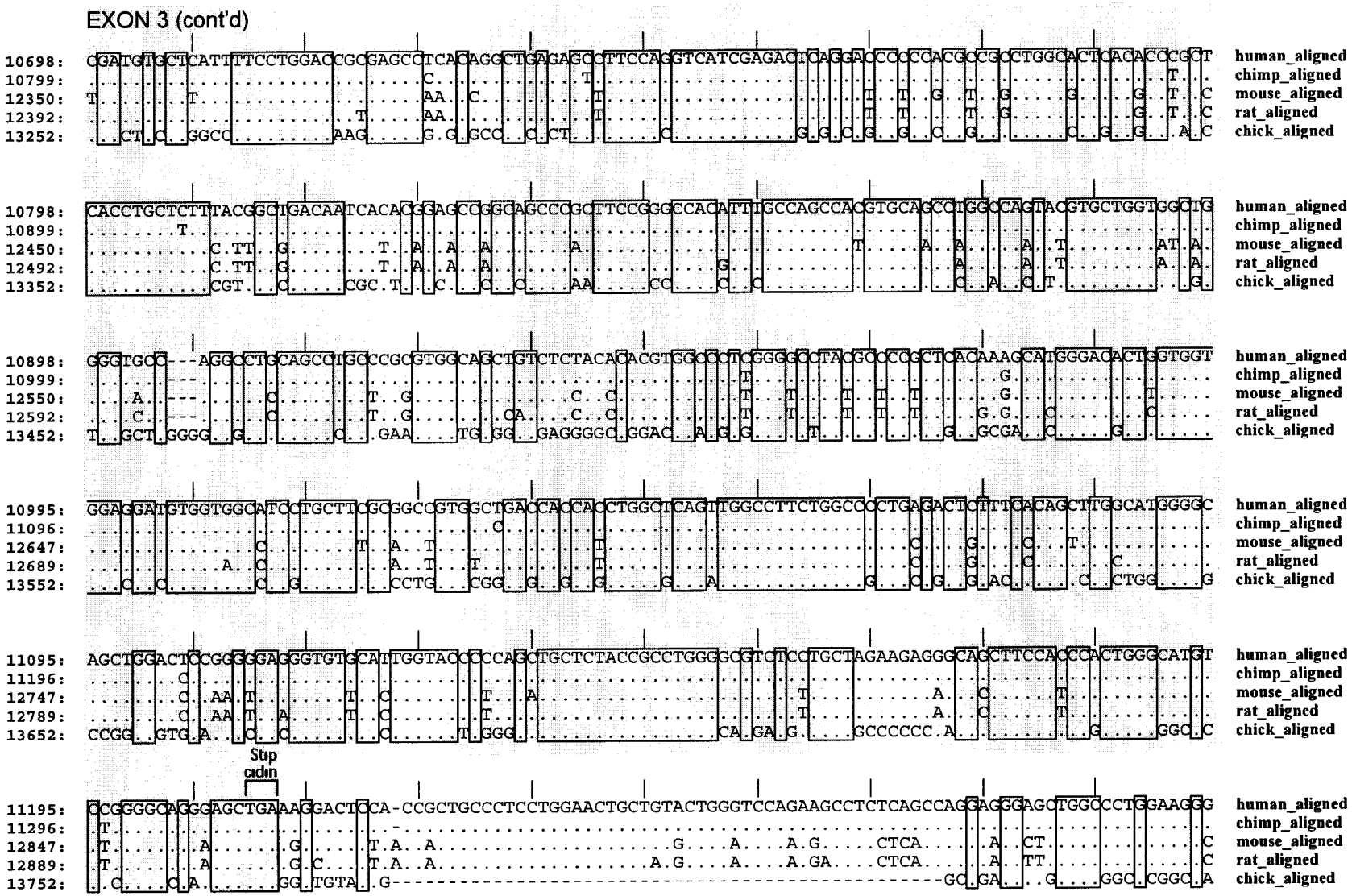

8990: ---------- human_aligned
8992: ---------- chimp_aligned
10653: ---------- mouse_aligned
10703: ---------- rat_aligned
10993: GCAGTCCCTTAGCCCGCGTCCAGCCCTGTCCTCCACAGCAGCACCCAGCCAACCCCAGGGTGTCCTATGGCCATTGCGGGACAAGGACGCGCCAGAAAGG chick_aligned
8990: ---------- human_aligned
8992: ---------- chimp_aligned
10653: ---------- mouse_aligned
10703: ---------- rat_aligned
11093: CGTCCTGCAGTCAGGCTCCTGCTGTGGAGGGGTGCTGGGGGCTGCGGGAAGCTGGGAACAGGGCTGGGTGCTGCGCTCGCTGGCGTTGCTCGCTTCTCGG chick_aligned
8990: ---------- human_aligned
8992: ---------- chimp_aligned
10653: ---------- mouse_aligned
10703: ---------- rat_aligned
11193: GCGGTGCCGTGAGGGGCGATGCCGTAGGGTGTGACGGTGCGGGGTGCGCCGCCGGGGAGACGGGAGCCGTTCTCGGGGTCCCGCGCAACCCCCGGCCCCT chick_aligned
8990: ---------- human_aligned
8992: ---------- chimp_aligned
10653: ---------- mouse_aligned
10703: ---------- rat_aligned
11293: CCCGGCCTTTCCCGGCCGCGCTGGCTGCGTGCGGCAGAGGGCAGCAGCGGCCCACGCACGGCGCGGGGCACCCGTGCGTGCTGCCGGGACGCGTCGGGAG chick_aligned
8990: ---------- human_aligned
8992: ---------- chimp_aligned
10653: ---------- mouse_aligned
10703: ---------- rat_aligned
11393: GGCTCCCCGCCGCTCGCGGGACGGCGTGTGCGTGGGGGCAGCCGCGGCGCCGGGCCCGCCCACGGCCGCGCACACCCACGGAGCCGGGCTCCCAGCGCAC chick_aligned
8990: ---------- human_aligned
8992: ---------- chimp_aligned
10653: ---------- mouse_aligned
10703: ---------- rat_aligned
11493: TGTGCCTGTCGTACCGCCCTAGCCCGCGCCCAGCGGCCCGACGCCCCCCGACGGCAGCCGGTCCCCCCGCGCTGATCCCATCCCCGATGCCAGGACGGCC chick_aligned

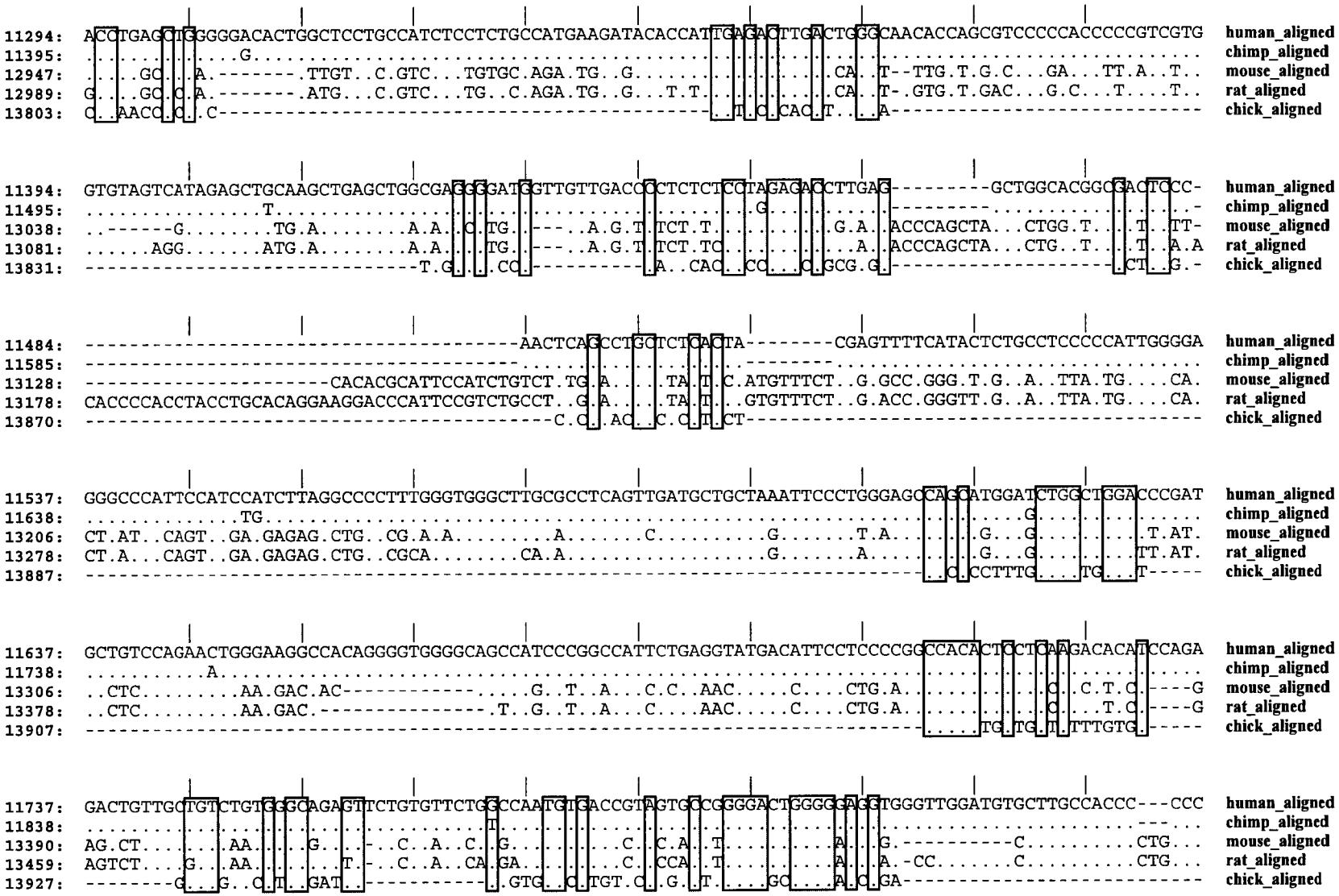

8990: human_aligned
8992: chimp_aligned
10653: mouse_aligned
10703: rat_aligned
11593: chick_aligned
CGCGATGGACGTGGGGCGACCGCGGCGCAGAGAGGGAGCGACGCGGCCCCGCAGCAGAGGGACGAAGCGTGGGCGAAGCGCCGGGGTTTCCCCGTCCCGG
8990: human_aligned
8992: chimp_aligned
10653: mouse_aligned
10703: rat_aligned
11693: chick_aligned
9073: human_aligned
9076: chimp_aligned
10747: mouse_aligned
10800: rat_aligned
11754: chick_aligned
9113: human_aligned
9115: chimp_aligned
10793: mouse_aligned
10840: rat_aligned
11848: chick_aligned
9201: human_aligned
9203: chimp_aligned
10893: mouse_aligned
10931: rat_aligned
11885: chick_aligned
9257: human_aligned
9259: chimp_aligned
10949: mouse_aligned
10980: rat_aligned
11970: chick_aligned


human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned


9681: human_aligned
9775: chimp_aligned
11347: mouse_aligned
11381: rat_aligned
12469: chick_aligned
9781: human_aligned
9875: chimp_aligned
11423: mouse_aligned
11441: rat_aligned
12556: chick_aligned
9824: human_aligned
9918: chimp_aligned
11469: mouse_aligned
11482: rat_aligned
12651: chick_aligned
9908: human_aligned
10002: chimp_aligned
11568: mouse_aligned
11577: rat_aligned
12693: chick_aligned
9989: human_aligned
10083: chimp_aligned
11667: mouse_aligned
11652: rat_aligned
12758: chick_aligned
10087: human_aligned
10181: chimp_aligned
11762: mouse_aligned
11741: rat_aligned
12841: chick_aligned


10187: human_aligned
10281: chimp_aligned
11828: mouse_aligned
11840: rat_aligned
12909: chick_aligned
10261: human_aligned
10357: chimp_aligned
11901: mouse_aligned
11935: rat_aligned
12947: chick_aligned
10307: human_aligned
10403: chimp_aligned
11996: mouse_aligned
12035: rat_aligned
12947: chick_aligned
10402: human_aligned
10503: chimp_aligned
12065: mouse_aligned
12101: rat_aligned
13001: chick_aligned
3' INIRON 2
Splice Site
EXON 3
10498: human_aligned
10599: chimp_aligned
12153: mouse_aligned
12193: rat_aligned
13072: chick_aligned
10598: human_aligned
10699: chimp_aligned
12250: mouse_aligned
12292: rat_aligned
13152: chick_aligned


EXON 3 (cont'd)
10698:
10799:
12350:
12392:
13252:
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned
10798:
10899:
12450:
12492:
13352:
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned
10898:
10999:
12550:
12592:
13452:
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned
10995:
11096:
12647:
12689:
13552:
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned
11095:
11196:
12747:
12789:
13652:
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned
Stop codon
11195:
11296:
12847:
12889:
13752:
human_aligned
chimp_aligned
mouse_aligned
rat_aligned
chick_aligned


11294: ACCTGAGCTGGGGGACACTGGCTCCTGCCATCTCCTCTGCCATGAAGATACACCATTGAGACTTGACTGGCAACACCAGCGTCCCCCACCCCCGTCGTG human_aligned
11395: .....G..... chimp_aligned
12947: ....GC...A.-------.TTGT..C.GTC...TGTGC.AGA.TG..G.....CA..T--TTG.T.G.C...GA...TT.A..T.. mouse_aligned
12989: G....GC.C.A.-------.ATG...C.GTC...TG..C.AGA.TG..G...T.T.....CA..T-.GTG.T.GAC...G.C...T....T.. rat_aligned
13803: C..AACC.C..C---------- chick_aligned
11394: GTGTAGTCATAGAGCTGCAAGCTGAGCTGGCGAGGGGATGGTTGTTGACCCCTCTCTCCTAGAGACCTTGAG---------GCTGGCACGGCGACTCCC- human_aligned
11495: ..............T.....G.....---------- chimp_aligned
13038: ..-------G........TG.A........A.A..C.TG..----..A.G.T.TCT.T.....G.A..ACCCAGCTA...CTGG.T....T..TT- mouse_aligned
13081: ......AGG.......ATG.A........A.A....TG..----..A.G.T.TCT.TC.....A..ACCCAGCTA...CTG..T....T..A.A rat_aligned
13831: ----------T.G...CC.----------.A..CAC..CC...C.GCG.G.----------.CT..G.- chick_aligned
11484: ----------AACTCAGCCTGCTCTCACTA--------CGAGTTTTCATACTCTGCCTCCCCCATTGGGGA human_aligned
11585: ----------..... chimp_aligned
13128: ----------CACACGCATTCCATCTGTCT.TG.A....TA.T.C.ATGTTTCT..G.GCC.GGG.T.G..A..TTA.TG....CA. mouse_aligned
13178: CACCCCACCTACCTGCACAGGAAGGACCCATTCCGTCTGCCT..G.A....TA.T...GTGTTTCT..G.ACC.GGGTT.G..A..TTA.TG....CA. rat_aligned
13870: ----------C.C..AC..C.C.T.CT---------- chick_aligned
11537: GGGCCCATTCCATCCATCTTAGGCCCCTTTGGGTGGGCTTGCGCCTCAGTTGATGCTGCTAAATTCCCTGGGAGCCAGCATGGATCTGGCTGGACCCGAT human_aligned
11638: .....TG.....G..... chimp_aligned
13206: CT.AT..CAGT..GA.GAGAG.CTG..CG.A.A.....A.....C.....G.....T.A.....G...G.....T.AT. mouse_aligned
13278: CT.A...CAGT..GA.GAGAG.CTG..CGCA.....CA.A.....G.....A.....G...G.....TT.AT. rat_aligned
13887: ----------..C.CCTTTG....TG...T----- chick_aligned
11637: GCTGTCCAGAACTGGGAAGGCCACAGGGGTGGGGCAGCCATCCCGGCCATTCTGAGGTATGACATTCCTCCCCGGCCACACTCCTCAAGACACATCCAGA human_aligned
11738: .....A..... chimp_aligned
13306: ..CTC.....AA.GAC.AC----------....G..T..A...C.C..AAC.....C....CTG.A.....C..C.T.C.----G mouse_aligned
13378: ..CTC.....AA.GAC.----------.T..G..T..A...C....AAC.....C....CTG.A.....C...T.C.----G rat_aligned
13907: ----------.....TG.TG.T.TTTGTG.----- chick_aligned
11737: GACTGTTGCTGTCTGTGGGCAGAGTTCTGTGTTCTGGCCAATGTGACCGTAGTGCCGGGGACTGGGGGAGGTGGGTTGGATGTGCTTGCCACCC---CCC human_aligned
11838: .....T..... chimp_aligned
13390: AG.CT.....AA.....G...-..C..A..C..G.....C..C.A..T.....A....G.----------C.....CTG... mouse_aligned
13459: AGTCT....G..AA.....T.-..C..A..CA.GA.....C..CCA..T.....A....A.-CC.....C.....CTG... rat_aligned
13927: --------G...G..C.T..GAT..----------..GTG..C.TGT.C..G..T....GC....A.C.GA---------- chick_aligned

Appendix B

The Unreliability behind the ATDC5 Transient Transfection Model

Table 1: The comparison between the results of Lipofectamine and FuGENE6 transfections in ATDC5 cells.

In the ATDC5 transient transfection assays, the pGL3-Basic vector is used as a negative control plasmid in order to obtain a background reading for the *firefly luciferase* (**Firefly**) gene. ATDC5 cells are co-transfected with either the pGL3-Basic vector or the p1879B construct and the internal control pRL-SV40 plasmid that constitutively expresses the *renilla luciferase* (**Renilla**) gene. The pRL-SV40 plasmid serves to quantify the transfection efficiency for a given assay such as to normalize the firefly luciferase readings for all samples. This is accomplished by dividing the RLU of *firefly luciferase* by *renilla luciferase* (**F/R**).

The F/R ratios obtained from transfecting the pGL3-Basic vector should remain constant during ATDC5 cell differentiation since the plasmid lacks a promoter for the *firefly luciferase* gene. Therefore, the induction of *luciferase* expression following the transfection of the p1879B construct into ATDC5 d21 cells should directly correlate with induction of *Ihh* expression. The amount of *luciferase* expression above background, expressed in relative units (**RU**), is determined by dividing the average F/R ratio (**AVG F/R**) of p1879B by that of pGL3-Basic. Moreover, the experimental value for induction of *luciferase* expression for the p1879B construct is obtained by dividing the RU of p1879B at d21 by that at d0.

In the case of Lipofectamine, the major complication was a low RLU yield for *firefly* and *renilla luciferase* at d21 (see boxes A-D in table 1). In comparison to d0, the *renilla* values were drastically reduced by at least 10-fold for both plasmids in d21 cells, indicating poor transfection efficiencies. Consequently, this explains why the *firefly* readings were actually lower in d21 cells than the baseline d0 cells. Nonetheless, the

luciferase expression for p1879B was 3.7-fold above background in d21 cells, which approximately translates to an overall 2-fold induction of *luciferase* expression from d0. Despite this induction, the absolute RLU in the d21 cell transfections are tremendously low in comparison to those in d0 cells for both *firefly* and *renilla luciferase*. Thus, the observed induction of *luciferase* expression in ATDC5 d21 cells can not be deemed an accurate representation of the gene activity of *Ihh*.

On the other hand, the use of FuGENE6 for ATDC5 transient transfections overcame the difficulties experienced with Lipofectamine. The robust *renilla* values indicated that the plasmids were transfecting efficiently and, as a result, the *firefly* values were proportionally elevated. However, *luciferase* expression in d21 cells fell 2.5-fold lower than d0 cells, from 4.3 to 1.8, even though the average F/R ratio for p1879B was 5-fold higher (see boxes H and J in table 1). The factor responsible for offsetting this induction is the 10-fold difference in the *renilla* values for the transfected pGL3-Basic vector at d0 and d21 (see box E and F in table 1). This would not have mattered so much had the pGL3-Basic *firefly* values dropped in proportion to the *renilla* values at d21, thus keeping the F/R ratio constant. Instead, the resulting average F/R ratio for pGL3-Basic at d21 is about 10-fold higher than that at d0 (see boxes G and I in table 1) which artificially diminishes the relative *luciferase* expression for p1879B at d21. Correcting for this artefact, a restored d21 baseline would result in a p1879B *luciferase* expression that is 21-fold above background, which would translate into an overall 5-fold induction over the d0 cells.

In conclusion, the ATDC5 21-day differentiation transient transfection model was replaced by ATDC5 stable transfections and COS7 transient transfections in order to better study the regulation of *Ihh* gene expression.

Table 1.

		Plasmid	Firefly	Renilla	F/R	AVG F/R	RU
LIPOFECTAMINE	ATDC5 d0	pGL3-Basic	1094	295416	0.0037		
			2377	478697	0.0050	0.0044	1.0
			2003	429333	0.0047		
		p1879B	11172	1221105	0.0091		
			10642	1437924	0.0074	0.0089	**2.0**
			10899	1064823	0.0102		
	ATDC5 d21	pGL3-Basic	A 411	C 45294	0.0091		
			593	50072	0.0118	0.0100	1.0
			409	44565	0.0092		
		p1879B	B 404	D 10590	0.0381		
			578	13926	0.0415	0.0367	**3.7**
			661	21663	0.0305		
FuGENE6	ATDC5 d0	pGL3-Basic	23298	E 8203297	0.0028	G	
			18351	7433904	0.0025	0.0028	1.0
			26422	8777368	0.0030		
		p1879B	5979	477185	0.0125	H	
			9753	895900	0.0109	0.0118	**4.3**
			12119	1006160	0.0120		
	ATDC5 d21	pGL3-Basic	15428	F 442343	0.0349	I	
			18237	584172	0.0312	0.0331	1.0
			17396	522402	0.0333		
		p1879B	29905	483715	0.0618	J	
			29982	510164	0.0588	0.0586	**1.8**
			41417	751212	0.0551		

Appendix C

Nuclear Substances and Radiation Devices Licence

Canadian Nuclear Safety Commission | Commission canadienne de sûreté nucléaire

05439-2-08.0

NUCLEAR SUBSTANCES AND RADIATION DEVICES LICENCE | PERMIS PORTANT SUR LES SUBSTANCES NUCLÉAIRES ET LES APPAREILS À RAYONNEMENT

Licence Number / Numéro de permis

I) LICENSEE

Pursuant to section 24(2) of the Nuclear Safety and Control Act, this licence is issued to:

Hôpital Shriners pour l'Enfant/Shriners Hospitals for Children
1529 Cedar Avenue
Montréal, QC
H3G 1A6
Canada

hereinafter «the licensee».

II) PERIOD

This licence is valid from: June 1 2003 to May 31 2008.

III) LICENSED ACTIVITIES

This licence authorizes the licensee to possess, transfer, import, export, use and store the nuclear substances listed in section IV) of this licence.

This licence is issued for: laboratory studies: 10 or more laboratories where radioisotopes are used or handled (836)

IV) NUCLEAR SUBSTANCES AND PRESCRIBED EQUIPMENT

ITEM	NUCLEAR SUBSTANCE	UNSEALED SOURCE MAXIMUM QUANTITY	SEALED SOURCE MAXIMUM QUANTITY	EQUIPMENT MAKE AND MODEL
1	Hydrogen 3	10 GBq	n/a	n/a
2	Carbon 14	200 MBq	n/a	n/a
3	Phosphorus 32	1 GBq	n/a	n/a
4	Sulfur 35	1 GBq	n/a	n/a
5	Calcium 45	200 MBq	n/a	n/a
6	Chromium 51	200 MBq	n/a	n/a
7	Iodine 125	1 GBq	n/a	n/a
8	Phosphorus 33	200 MBq	n/a	n/a
9	Technetium 99m	100 MBq	n/a	n/a

The total quantity of an unsealed nuclear substance in possession shall not exceed the corresponding listed unsealed source maximum quantity. The total quantity of nuclear substance per sealed source shall not exceed its corresponding listed sealed source maximum quantity. Sealed sources shall only be used in the corresponding listed equipment.

V) LOCATION(S) OF LICENSED ACTIVITIES

1529 Cedar Avenue
Montréal, QC

VI) CONDITIONS

1. Prohibition of Human Use
   This licence does not authorize the use of nuclear substances in or on human beings.
   (2696-0)

References

1. Ingham, P.W. and A.P. McMahon, *Hedgehog signaling in animal development: paradigms and principles.* Genes Dev, 2001. **15**(23): p. 3059-87.

2. Nusslein-Volhard, C. and E. Wieschaus, *Mutations affecting segment number and polarity in Drosophila.* Nature, 1980. **287**(5785): p. 795-801.

3. Mohler, J., *Requirements for hedgehod, a segmental polarity gene, in patterning larval and adult cuticle of Drosophila.* Genetics, 1988. **120**(4): p. 1061-72.

4. Bitgood, M.J., L. Shen, and A.P. McMahon, *Sertoli cell signaling by Desert hedgehog regulates the male germline.* Curr Biol, 1996. **6**(3): p. 298-304.

5. Clark, A.M., K.K. Garland, and L.D. Russell, *Desert hedgehog (Dhh) gene is required in the mouse testis for formation of adult-type Leydig cells and normal development of peritubular cells and seminiferous tubules.* Biol Reprod, 2000. **63**(6): p. 1825-38.

6. Umehara, F., et al., *A novel mutation of desert hedgehog in a patient with 46,XY partial gonadal dysgenesis accompanied by minifascicular neuropathy.* Am J Hum Genet, 2000. **67**(5): p. 1302-5.

7. Marigo, V., et al., *Cloning, expression, and chromosomal location of SHH and IHH: two human homologues of the Drosophila segment polarity gene hedgehog.* Genomics, 1995. **28**(1): p. 44-51.

8. Levine, E.M., et al., *Sonic hedgehog promotes rod photoreceptor differentiation in mammalian retinal cells in vitro.* J Neurosci, 1997. **17**(16): p. 6277-88.

9. Byrd, N., et al., *Hedgehog is required for murine yolk sac angiogenesis.* Development, 2002. **129**(2): p. 361-72.

10. Takamoto, N., et al., *Identification of Indian hedgehog as a progesterone-responsive gene in the murine uterus.* Mol Endocrinol, 2002. **16**(10): p. 2338-48.

11. Valentini, R.P., et al., *Post-translational processing and renal expression of mouse Indian hedgehog.* J Biol Chem, 1997. **272**(13): p. 8466-73.

12. Hebrok, M., et al., *Regulation of pancreas development by hedgehog signaling.* Development, 2000. **127**(22): p. 4905-13.

13. Bitgood, M.J. and A.P. McMahon, *Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo.* Dev Biol, 1995. **172**(1): p. 126-38.

14. Porter, J.A., et al., *The product of hedgehog autoproteolytic cleavage active in local and long-range signalling.* Nature, 1995. **374**(6520): p. 363-6.

15. Hall, T.M., et al., *Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins.* Cell, 1997. **91**(1): p. 85-97.

16. Porter, J.A., et al., *Hedgehog patterning activity: role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain.* Cell, 1996. **86**(1): p. 21-34.

17. Pepinsky, R.B., et al., *Identification of a palmitic acid-modified form of human Sonic hedgehog.* J Biol Chem, 1998. **273**(22): p. 14037-45.

18. Lewis, P.M., et al., *Cholesterol modification of sonic hedgehog is required for long-range signaling activity and effective modulation of signaling by Ptc1.* Cell, 2001. **105**(5): p. 599-612.

19. Ingham, P.W., *Transducing Hedgehog: the story so far.* Embo J, 1998. **17**(13): p. 3505-11.

20. Marigo, V., et al., *Conservation in hedgehog signaling: induction of a chicken patched homolog by Sonic hedgehog in the developing limb.* Development, 1996. **122**(4): p. 1225-33.

21. Marigo, V., et al., *Biochemical evidence that patched is the Hedgehog receptor.* Nature, 1996. **384**(6605): p. 176-9.

22. Chen, Y. and G. Struhl, *Dual roles for patched in sequestering and transducing Hedgehog.* Cell, 1996. **87**(3): p. 553-63.

23. Briscoe, J., et al., *A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube.* Mol Cell, 2001. **7**(6): p. 1279-91.

24. Quirk, J., et al., *The smoothened gene and hedgehog signal transduction in Drosophila and vertebrate development.* Cold Spring Harb Symp Quant Biol, 1997. **62**: p. 217-26.

25. Chen, W., S. Burgess, and N. Hopkins, *Analysis of the zebrafish smoothened mutant reveals conserved and divergent functions of hedgehog activity.* Development, 2001. **128**(12): p. 2385-96.

26. Zhang, X.M., M. Ramalho-Santos, and A.P. McMahon, *Smoothened mutants reveal redundant roles for Shh and Ihh signaling including regulation of L/R symmetry by the mouse node.* Cell, 2001. **106**(2): p. 781-92.

27. Sasaki, H., et al., *A binding site for Gli proteins is essential for HNF-3beta floor plate enhancer activity in transgenics and can respond to Shh in vitro.* Development, 1997. **124**(7): p. 1313-22.

28. Hynes, M., et al., *Control of cell pattern in the neural tube by the zinc finger transcription factor and oncogene Gli-1.* Neuron, 1997. **19**(1): p. 15-26.

29. Marigo, V., et al., *Sonic hedgehog differentially regulates expression of GLI and GLI3 during limb development.* Dev Biol, 1996. **180**(1): p. 273-83.

30. Lee, J., et al., *Gli1 is a target of Sonic hedgehog that induces ventral neural tube development.* Development, 1997. **124**(13): p. 2537-52.

31. Hall, B.K. and T. Miyake, *All for one and one for all: condensations and the initiation of skeletal development.* Bioessays, 2000. **22**(2): p. 138-47.

32. Hall, B.K. and T. Miyake, *Divide, accumulate, differentiate: cell condensation in skeletal development revisited.* Int J Dev Biol, 1995. **39**(6): p. 881-93.

33. Zelzer, E. and B.R. Olsen, *The genetic basis for skeletal diseases.* Nature, 2003. **423**(6937): p. 343-8.

34. Gilbert, S.F., *Developmental Biology*. 7 ed. 2003, Sunderland, MA: Sinauer Associates, Inc.

35. *Pediatric Bone: Biology and Diseases*, ed. F.H. Glorieux. 2003, San Diego, CA: Academic Press.

36. Bi, W., et al., *Sox9 is required for cartilage formation.* Nat Genet, 1999. **22**(1): p. 85-9.

37. Poole, A.R., *The growth plate: Cellular physiology, cartilage assembly and mineralization.*, in *Cartilage: molecular aspects*, B. Hall, Newman, S., Editor. 1991, CRC Press: Boca Raton, FL. p. 179-211.

38. Hill, P.A., *Bone remodelling.* Br J Orthod, 1998. **25**(2): p. 101-7.

39. Russell, G., et al., *Clinical disorders of bone resorption.* Novartis Found Symp, 2001. **232**: p. 251-67; discussion 267-71.

40. Karaplis, A.C., et al., *Lethal skeletal dysplasia from targeted disruption of the parathyroid hormone-related peptide gene.* Genes Dev, 1994. **8**(3): p. 277-89.

41. Lanske, B., et al., *PTH/PTHrP receptor in early development and Indian hedgehog-regulated bone growth.* Science, 1996. **273**(5275): p. 663-6.

42. St-Jacques, B., M. Hammerschmidt, and A.P. McMahon, *Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation.* Genes Dev, 1999. **13**(16): p. 2072-86.

43. Vortkamp, A., et al., *Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein.* Science, 1996. **273**(5275): p. 613-22.

44. Weir, E.C., et al., *Targeted overexpression of parathyroid hormone-related peptide in chondrocytes causes chondrodysplasia and delayed endochondral bone formation.* Proc Natl Acad Sci U S A, 1996. **93**(19): p. 10240-5.

45. Kindblom, J.M., et al., *Expression and localization of Indian hedgehog (Ihh) and parathyroid hormone related protein (PTHrP) in the human growth plate during pubertal development.* J Endocrinol, 2002. **174**(2): p. R1-6.

46. Alvarez, J., et al., *TGFbeta2 mediates the effects of hedgehog on hypertrophic differentiation and PTHrP expression.* Development, 2002. **129**(8): p. 1913-24.

47. Long, F., et al., *Genetic manipulation of hedgehog signaling in the endochondral skeleton reveals a direct role in the regulation of chondrocyte proliferation.* Development, 2001. **128**(24): p. 5099-108.

48. Naski, M.C., et al., *Repression of hedgehog signaling and BMP4 expression in growth plate cartilage by fibroblast growth factor receptor 3.* Development, 1998. **125**(24): p. 4977-88.

49. Seki, K. and A. Hata, *Indian hedgehog gene is a target of the bone morphogenetic protein signaling pathway.* J Biol Chem, 2004. **279**(18): p. 18544-9.

50. Yoshida, C.A., et al., *Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog.* Genes Dev, 2004. **18**(8): p. 952-63.

51. Myers, R.M., K. Tilly, and T. Maniatis, *Fine structure genetic analysis of a beta-globin promoter.* Science, 1986. **232**(4750): p. 613-8.

52. McKnight, S. and R. Tjian, *Transcriptional selectivity of viral genes in mammalian cells.* Cell, 1986. **46**(6): p. 795-805.

53. Grosschedl, R. and M.L. Birnstiel, *Spacer DNA sequences upstream of the T-A-T-A-A-A-T-A sequence are essential for promotion of H2A histone gene transcription in vivo.* Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7102-6.

54. Tamayo, E., et al., *Mediator is required for activated transcription in a Schizosaccharomyces pombe in vitro system.* Eur J Biochem, 2004. **271**(12): p. 2561-72.

55. Kadonaga, J.T., *Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors.* Cell, 2004. **116**(2): p. 247-57.

56. Sawadogo, M. and R.G. Roeder, *Energy requirement for specific transcription initiation by the human RNA polymerase II system.* J Biol Chem, 1984. **259**(8): p. 5321-6.

57. Bunick, D., et al., *Mechanism of RNA polymerase II--specific initiation of transcription in vitro: ATP requirement and uncapped runoff transcripts.* Cell, 1982. **29**(3): p. 877-86.

58. Koleske, A.J., et al., *A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID.* Cell, 1992. **69**(5): p. 883-94.

59. Usheva, A., et al., *Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein.* Cell, 1992. **69**(5): p. 871-81.

60. Dynlacht, B.D., T. Hoey, and R. Tjian, *Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation.* Cell, 1991. **66**(3): p. 563-76.

61. Kadonaga, J.T., et al., *Distinct regions of Sp1 modulate DNA binding and transcriptional activation.* Science, 1988. **242**(4885): p. 1566-70.

62. Pugh, B.F. and R. Tjian, *Transcription from a TATA-less promoter requires a multisubunit TFIID complex.* Genes Dev, 1991. **5**(11): p. 1935-45.

63. Rigby, P.W., *Three in one and one in three: it all depends on TBP.* Cell, 1993. **72**(1): p. 7-10.

64. Green, S. and P. Chambon, *Nuclear receptors enhance our understanding of transcription regulation.* Trends Genet, 1988. **4**(11): p. 309-14.

65. Amouyal, M., *The remote control of transcription, DNA looping and DNA compaction.* Biochimie, 1991. **73**(10): p. 1261-8.

66. Yuh, C.H., H. Bolouri, and E.H. Davidson, *Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene.* Science, 1998. **279**(5358): p. 1896-902.

67. Arnone, M.I. and E.H. Davidson, *The hardwiring of development: organization and function of genomic regulatory systems.* Development, 1997. **124**(10): p. 1851-64.

68. Nemeth, A. and G. Langst, *Chromatin higher order structure: opening up chromatin for transcription.* Brief Funct Genomic Proteomic, 2004. **2**(4): p. 334-43.

69. Marcand, S., S.M. Gasser, and E. Gilson, *Chromatin: a sticky silence.* Curr Biol, 1996. **6**(10): p. 1222-5.

70. Workman, J.L. and R.E. Kingston, *Alteration of nucleosome structure as a mechanism of transcriptional regulation.* Annu Rev Biochem, 1998. **67**: p. 545-79.

71. Grunstein, M., *Histone acetylation in chromatin structure and transcription.* Nature, 1997. **389**(6649): p. 349-52.

72. Bestor, T.H., *The DNA methyltransferases of mammals.* Hum Mol Genet, 2000. **9**(16): p. 2395-402.

73. Kelly, T.L. and J.M. Trasler, *Reproductive epigenetics.* Clin Genet, 2004. **65**(4): p. 247-60.

74. Hogan, B.L., *Bone morphogenetic proteins: multifunctional regulators of vertebrate development.* Genes Dev, 1996. **10**(13): p. 1580-94.

75. Wozney, J.M., et al., *Novel regulators of bone formation: molecular clones and activities.* Science, 1988. **242**(4885): p. 1528-34.

76. Tsumaki, N., et al., *Bone morphogenetic protein signals are required for cartilage formation and differently regulate joint development during skeletogenesis.* J Bone Miner Res, 2002. **17**(5): p. 898-906.

77. Zhang, D., et al., *ALK2 functions as a BMP type I receptor and induces Indian hedgehog in chondrocytes during skeletal development.* J Bone Miner Res, 2003. **18**(9): p. 1593-604.

78. Duprez, D., et al., *Overexpression of BMP-2 and BMP-4 alters the size and shape of developing skeletal elements in the chick limb.* Mech Dev, 1996. **57**(2): p. 145-57.

79. Pizette, S. and L. Niswander, *BMPs are required at two steps of limb chondrogenesis: formation of prechondrogenic condensations and their differentiation into chondrocytes.* Dev Biol, 2000. **219**(2): p. 237-49.

80. Pizette, S. and L. Niswander, *BMPs negatively regulate structure and function of the limb apical ectodermal ridge.* Development, 1999. **126**(5): p. 883-94.

81. Zou, H., et al., *Distinct roles of type I bone morphogenetic protein receptors in the formation and differentiation of cartilage.* Genes Dev, 1997. **11**(17): p. 2191-203.

82. Grimsrud, C.D., et al., *BMP signaling stimulates chondrocyte maturation and the expression of Indian hedgehog.* J Orthop Res, 2001. **19**(1): p. 18-25.

83. Pathi, S., et al., *Interaction of Ihh and BMP/Noggin signaling during cartilage differentiation.* Dev Biol, 1999. **209**(2): p. 239-53.

84. Brunet, L.J., et al., *Noggin, cartilage morphogenesis, and joint formation in the mammalian skeleton.* Science, 1998. **280**(5368): p. 1455-7.

85. Chen, D., et al., *Signal transduction and biological functions of bone morphogenetic proteins.* Front Biosci, 2004. **9**: p. 349-58.

86. Haaijman, A., et al., *Correlation between ALK-6 (BMPR-IB) distribution and responsiveness to osteogenic protein-1 (BMP-7) in embryonic mouse bone rudiments.* Growth Factors, 2000. **17**(3): p. 177-92.

87. Sakou, T., et al., *Localization of Smads, the TGF-beta family intracellular signaling components during endochondral ossification.* J Bone Miner Res, 1999. **14**(7): p. 1145-52.

88. Yi, S.E., et al., *The type I BMP receptor BMPRIB is required for chondrogenesis in the mouse limb.* Development, 2000. **127**(3): p. 621-30.

89. Minina, E., et al., *BMP and Ihh/PTHrP signaling interact to coordinate chondrocyte proliferation and differentiation.* Development, 2001. **128**(22): p. 4523-34.

90. Ducy, P., et al., *Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation.* Cell, 1997. **89**(5): p. 747-54.

91. Ducy, P., et al., *A Cbfa1-dependent genetic pathway controls bone formation beyond embryonic development.* Genes Dev, 1999. **13**(8): p. 1025-36.

92. Kim, I.S., et al., *Regulation of chondrocyte differentiation by Cbfa1.* Mech Dev, 1999. **80**(2): p. 159-70.

93. Inada, M., et al., *Maturational disturbance of chondrocytes in Cbfa1-deficient mice.* Dev Dyn, 1999. **214**(4): p. 279-90.

94. Stricker, S., et al., *Role of Runx genes in chondrocyte differentiation.* Dev Biol, 2002. **245**(1): p. 95-108.

95. Klint, P. and L. Claesson-Welsh, *Signal transduction by fibroblast growth factor receptors.* Front Biosci, 1999. **4**: p. D165-77.

96. Ornitz, D.M. and P.J. Marie, *FGF signaling pathways in endochondral and intramembranous bone development and human genetic disease.* Genes Dev, 2002. **16**(12): p. 1446-65.

97. Ornitz, D.M. and N. Itoh, *Fibroblast growth factors.* Genome Biol, 2001. **2**(3): p. REVIEWS3005.

98. Heldin, C.H., *Dimerization of cell surface receptors in signal transduction.* Cell, 1995. **80**(2): p. 213-23.

99. Pawson, T., *Protein modules and signalling networks.* Nature, 1995. **373**(6515): p. 573-80.

100. Colvin, J.S., et al., *Skeletal overgrowth and deafness in mice lacking fibroblast growth factor receptor 3.* Nat Genet, 1996. **12**(4): p. 390-7.

101. Deng, C., et al., *Fibroblast growth factor receptor 3 is a negative regulator of bone growth.* Cell, 1996. **84**(6): p. 911-21.

102. Webster, M.K. and D.J. Donoghue, *Constitutive activation of fibroblast growth factor receptor 3 by the transmembrane domain point mutation found in achondroplasia.* Embo J, 1996. **15**(3): p. 520-7.

103. Naski, M.C., et al., *Graded activation of fibroblast growth factor receptor 3 by mutations causing achondroplasia and thanatophoric dysplasia.* Nat Genet, 1996. **13**(2): p. 233-7.

104. Li, C., et al., *A Lys644Glu substitution in fibroblast growth factor receptor 3 (FGFR3) causes dwarfism in mice by activation of STATs and ink4 cell cycle inhibitors.* Hum Mol Genet, 1999. **8**(1): p. 35-44.

105. Su, W.C., et al., *Activation of Stat1 by mutant fibroblast growth-factor receptor in thanatophoric dysplasia type II dwarfism.* Nature, 1997. **386**(6622): p. 288-92.

106. Sambrook, J., Russell, D. W., *Molecular Cloning: A Laboratory Manual.* 3rd ed. 2001, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

107. Brudno, M., et al., *LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA.* Genome Res, 2003. **13**(4): p. 721-31.

108. Stojanovic, N. and K. Dewar, *Identifying multiple alignment regions satisfying simple formulas and patterns.* Bioinformatics, 2004.

109. Atsumi, T., et al., *A chondrogenic cell line derived from a differentiating culture of AT805 teratocarcinoma cells.* Cell Differ Dev, 1990. **30**(2): p. 109-16.

110. Shukunami, C., et al., *Cellular hypertrophy and calcification of embryonal carcinoma-derived chondrogenic cell line ATDC5 in vitro.* J Bone Miner Res, 1997. **12**(8): p. 1174-88.

111. Courey, A.J. and R. Tjian, *Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif.* Cell, 1988. **55**(5): p. 887-98.

112. Weksler, N.B., et al., *Differential effects of fibroblast growth factor (FGF) 9 and FGF2 on proliferation, differentiation and terminal differentiation of chondrocytic cells in vitro.* Biochem J, 1999. **342 Pt 3**: p. 677-82.

113. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

114. Trippel, S.B., et al., *Characterization of chondrocytes from bovine articular cartilage: I. Metabolic and morphological experimental studies.* J Bone Joint Surg Am, 1980. **62**(5): p. 816-20.

115. Gustafson, K.S. and G.D. Ginder, *Interferon-gamma induction of the human leukocyte antigen-E gene is mediated through binding of a complex containing STAT1alpha to a distinct interferon-gamma-responsive element.* J Biol Chem, 1996. **271**(33): p. 20035-46.

116. Bugno, M., et al., *Identification of the interleukin-6/oncostatin M response element in the rat tissue inhibitor of metalloproteinases-1 (TIMP-1) promoter.* Nucleic Acids Res, 1995. **23**(24): p. 5041-7.

117. Nakamasu, K., et al., *Structure and promoter analysis of the mouse membrane-bound transferrin-like protein (MTf) gene.* Eur J Biochem, 2001. **268**(5): p. 1468-76.

118. Kozak, M., *An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs.* Nucleic Acids Res, 1987. **15**(20): p. 8125-48.

119. Echelard, Y., et al., *Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity.* Cell, 1993. **75**(7): p. 1417-30.

120. Kozak, M., *Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes.* Cell, 1986. **44**(2): p. 283-92.

121. Liu, C.C., C.C. Simonsen, and A.D. Levinson, *Initiation of translation at internal AUG codons in mammalian cells.* Nature, 1984. **309**(5963): p. 82-5.

122. Gluzman, Y., *SV40-transformed simian cells support the replication of early SV40 mutants.* Cell, 1981. **23**(1): p. 175-82.

123. Iwata, T., et al., *A neonatal lethal mutation in FGFR3 uncouples proliferation and differentiation of growth plate chondrocytes in embryos.* Hum Mol Genet, 2000. **9**(11): p. 1603-13.