NOTE TO USERS

This reproduction is the best copy available.

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHARACTERIZATION OF UPSTREAM SEQUENCES REQUIRED FOR TRANSCRIPTIONAL CONTROL OF THE *als* GENE ENCODING A SUBUNIT OF A NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR FROM *DROSOPHILA*

by

DONGDONG MA

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2004

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 3144117

Copyright 2004 by
Ma, Dongdong

All rights reserved.

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3144117
Copyright 2004 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

©2004

DONGDONG MA

All Rights Reserved

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This manuscript has been read and accepted for the Graduate Faculty in Biology in Satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

9/14/04
Date

Chair of Examining Committee
Dr. Thomas Schmidt-Glenewinkel, Hunter College

9/20/04
Date

Executive Officer
Dr. Richard L. Chappell

Dr. Rivka Rudner, Hunter College

Dr. Maria Figueiredo-Pereira, Hunter College

Dr. Shubha Govind, City College

Dr. Mindong Ren
New York University School of Medicine

Supervising Committee

The City University of New York

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

Characterization of Upstream Sequences Required for Transcriptional Control of the *als* Gene Encoding A Subunit of A Neuronal Nicotinic Acetylcholine Receptor from *Drosophila*

by

Dongdong Ma

Adviser: Professor Thomas Schmidt-Glenewinkel

The *als* gene (nAcRα-96Aa) encodes an α-like subunit of a neuronal nicotinic acetylcholine receptor (nAChR) from *Drosophila*. To understand how *als* expression is regulated, we report here the molecular characterization of the *als* gene and *in vivo* analysis of upstream regulatory sequences which confer the correct temporal and spatial expression. Primer Extension and RACE-PCR analysis in both embryo and adult flies allowed us to establish the transcription start site of *als*. Upstream fragments of 2.9 kb and 1.25 kb including the transcription start site and partial noncoding exon were cloned into the P-element vector pCaSpeR-AUG-βgal generating promoter fusions with *lacZ* reporter gene and used to transform *Drosophila*. Both upstream fragments were shown to drive *als* expression in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

central nervous system and reproduce the tissue and developmental expression pattern as comparable to that of the endogenous gene *als* (in embryo) and *ard* (in adult).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGMENTS

To Dr. Thomas Schmidt-Glenewinkel, the multi-talented, resourceful, sophisticated and understanding professor, the nicest person I have ever met. My gratitude to you is beyond words.

To my dear parents, for your unconditional love.

Special thanks to Dr. Maria Figueiredo-Pereira, and Dr. Rivka Rudner, it is with your encouragement, that I made it through the hardest period of my thesis research.

To my colleagues Marek Pierzchala, Jingguo Ma and all my good friends for their morale support.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

TITLE i
COPYRIGHT ii
APPROVAL PAGE iii
ABSTRACT iv
ACKNOWLEDGMENTS vi
TABLE OF CONTENTS vii
LIST OF FIGURES xi
LIST OF TABLES xii
LIST OF ABBREVIATIONS xiii

INTRODUCTION 1
General Background 1
Diversity of Neuronal nAChR Subtypes 5
Insect nAChR Subunit Genes 9
Molecular Mechanisms That Regulate Neuronal nAChR Expression 16

CHAPTER 1. 25
IN SEARCH OF GENOMIC DNA CLONES CARRYING THE *ALS* GENE AND ITS UPSTREAM REGULATORY SEQUENCE 25
I. Experimental Approach I
Early Strategy: Conventional Genomic Library Screening with *als*-specific probe 26
1. Material and Method 26
2. Result 30
II. Experimental Approach II
The Post-Genome Era and the Bacterial Artificial Chromosomes (BAC) – An Alternative strategy to identify and manipulate the *als* locus 32
1. Material and Method 32
2. Result 35
Discussion 37

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 2. 41
IDENTIFYING THE TRANSCRIPTION START SITE OF THE *als* GENE 41
MATERIAL AND METHOD 44
1. Fly Culture 44
2. Preparation of Total and Poly(A)$^+$ RNA 44
3. RNA gel electrophoresis 45
4. RT-PCR 46
5. Primer extension analysis 48
6. RACE-PCR 50
RESULTS 56
RT-PCR 56
Primer extension analysis 58
RACE-PCR 59

CHAPTER 3. 65
BIOINFORMATIC ANALYSIS OF THE *als* GENE AND ITS ENCODING PROTEIN 65

PART 1. The *als* Gene and the ALS Protein 65
I. MATERIAL AND METHOD 65
1. Gene Structure Analysis 65
2. Amino Acid Sequence Analysis and Alignment 66
3. Computer-assisted analysis of promoter regions and consensus transcription factor binding sites 67
II. RESULTS 68
III. DISCUSSION 80

PART 2. Phylogenetic Study of the nAChR Subunit Genes 87
I. MATERIAL AND METHOD 87
1. Sequence Collection 87
2. Phylogenetic reconstruction 94
II. RESULT 94
III. DISCUSSION 99

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 4. 104
***In Vivo* Study of the Upstream Regulatory Sequence of the *als* Gene** 104
I. MATERIAL AND METHOD 107
1. Making Fusion Construct: 107
2. Microinjection 110
(1) Supplies: 110
(2) Procedures: 111
i) Preparation of egg-laying plates 111
ii) Collection of embryos 112
iii) Lining up of embryos 113
iv) Preparation of glass needle for injections 114
v) Preparation of DNA 115
vi) Injection 116
vii) Post-injection 116
viii) Collecting G_0s and making genetics crosses of transformants 117
3. Analysing Transgenic Flies 118
(1) β-galactosidase activity assay 118
(2) β-galactosidase Histochemistry 119
i) Frozen Sections 119
ii) X-gal Staining in Frozen Sections 120
iii) X-gal Staining in Whole-mount Embryos 121
iv) β–Galactosidase Histochemistry in Larvae 122
(3) Immunohistochemistry – Antibody stainings of frozen sections 123
RESULTS 125
DISCUSSION 140

REFERENCES 144

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES

Figure 1. Structure of A Neuronal Nicotinic Acetylcholine Receptor 2
Figure 2A. The Design of *als*-specific Oligonucleotide Probe for Genomic Library Screening (1) 27
Figure 2B. The Design of *als*-specific Oligonucleotide Probe for Genomic Library Screening (2) 28
Figure 2C. XhoI Digestion of Four Positive Clones from Genomic Library Screening 30
Figure 2D. Southern Blotting of Four Genomic Clones 31
Figure 3. The *als* Gene Versus BACR 13M04 36
Figure 4A. The Design of *als*-specific primers for exon-walking. 43
Figure 5A. The Design of Oligonucleotide Probes for Primer Extension Analysis. 49
Figure 6A. Mechanism of SMART cDNA synthesis. 52
Figure 6B. The Design of primers for RACE-PCR. 54
Figure 4B. RT-PCR Products. 57
Figure 5B. Autoradiography of Primer Extension Products. 59
Figure 6C. Results of 5'RACE-PCR 60
Figure 6D. Results of 3'RACE-PCR 61
Figure 7. Full-length als cDNA. 62
Figure 8. Transcription Unit of the *als* cDNA. 68
Figure 9. Nucleotide sequences and putative amino acid sequences of the *als* gene. 71

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 10. Protein Sequence Pile-Up of all known *Drosophila* nAChR subunits. 75
Figure 11A. Molecular Map of the 1.2 kb fragment. 78
Figure 11B. Analysis of DNA sequence upstream of *als* transcription start site. 79
Figure 12. Phylogenetic relationships among 107 nAChR subunit genes. 94
Figure 13. Molecular Map of the P-element vector pCaSpeR-AUG-βgal. 108
Figure 14. The 2.9 kb *als* upstream fragment. 110
Figure 15. β-galactosidase Activity Assay. 128
Figure 16. β-galactosidase distribution pattern in embryos. .. 131
Figure 17. Distribution of β-galactosidase activity in one-day old head section of RE-3 transformed isolines. 133
Figure 18. Distribution of β-galactosidase activity in one-day old head section of RB-2 transformed isolines. 134
Figure 19. Distribution of β-galactosidase activity in one-day old head section of RE-11 transformed isolines 135
Figure 20. Distribution of β-galactosidase activity in thoracic region of RE-13 transformed isolines 136
Figure 21. Distribution of β-galactosidase activity in larvae of RE-18 transformed isolines. 137
Figure 22. Distribution of β-galactosidase activity in prepupae of RB-9 transformed isolines. 137

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Table 1. List of Identified Nicotinic Acetylcholine Receptor Subunit Genes in *Drosophila melanogaster* 15

Table 2. List of Matching Clones with 24 mer *als*-specific Oligonucleotide Probe – A Hindsight 40

Table 3. Primers Used for Various Applications as Described in the Text 63

Table 4. Exon/Intron Boundaries of the *als* Gene. 70

Table 5. Amino Acid Sequence Identity Score of Pileuped nAChR Subunits. 76

Table 6. nAChR Subunit Nucleotide Sequences Used in Phylogenetic Study 88

Table 7. Summary of X-gal Staining in Isolines of Transgenic Flies. 129

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF ABBREVIATIONS

ACh	acetylcholine
AChR	acetylcholine receptor
nAChR	nicotinic acetylcholine receptor
LGIC	ligand-gated ion channels
VGIC	voltage-gated ion channels
α-BgtX	α-Bungarotoxin
CNS	central nervous system
PNS	peripheral nervous system
ADAR	Adenosine Deaminase Act on RNA
UTR	untranslated region
Tm	melting temperature
BAC	Bacterial Artificial Chromosome
RACE	Rapid Amplification of cDNA Ends
SMART	Switching Mechanism At 5′ end of RNA Transcript
DTT	Dithiothreitol
SDS	sodium dodecyl sulfate
EDTA	ethylenediamine tetraacetic acid

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

INTRODUCTION

General Background

The nervous system requires fast and efficient intercellular communications among neurons. Rapid responses of a neuron to neurotransmitter release into the synaptic cleft are accomplished through the action of ligand-gated ion channels (LGICs) that either excite or inhibit postsynaptic neuronal response. Nicotinic acetylcholine receptors (nAChRs) belong to the supergene family of LGICs and can be divided into two types: the muscle type with stoichiometry $(\alpha)_2\beta_1\gamma_1\delta_1$ which is found at the skeletal neuromuscular junction mediating neuromuscular transmission; and the neuronal type, distributed throughout the peripheral and central nervous system where it is involved in fast excitatory synaptic transmission. The neuronal nAChRs can be subdivided into two classes based on their subunit composition, either as heteropentamers with general composition of two α-type subunits and three β-type subunits (Anand *et al.* 1991; Cooper *et al.* 1991), or as homopentamers which is composed only of α subunits (Chen and Patrick 1997; Drisdel and Green 2000).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

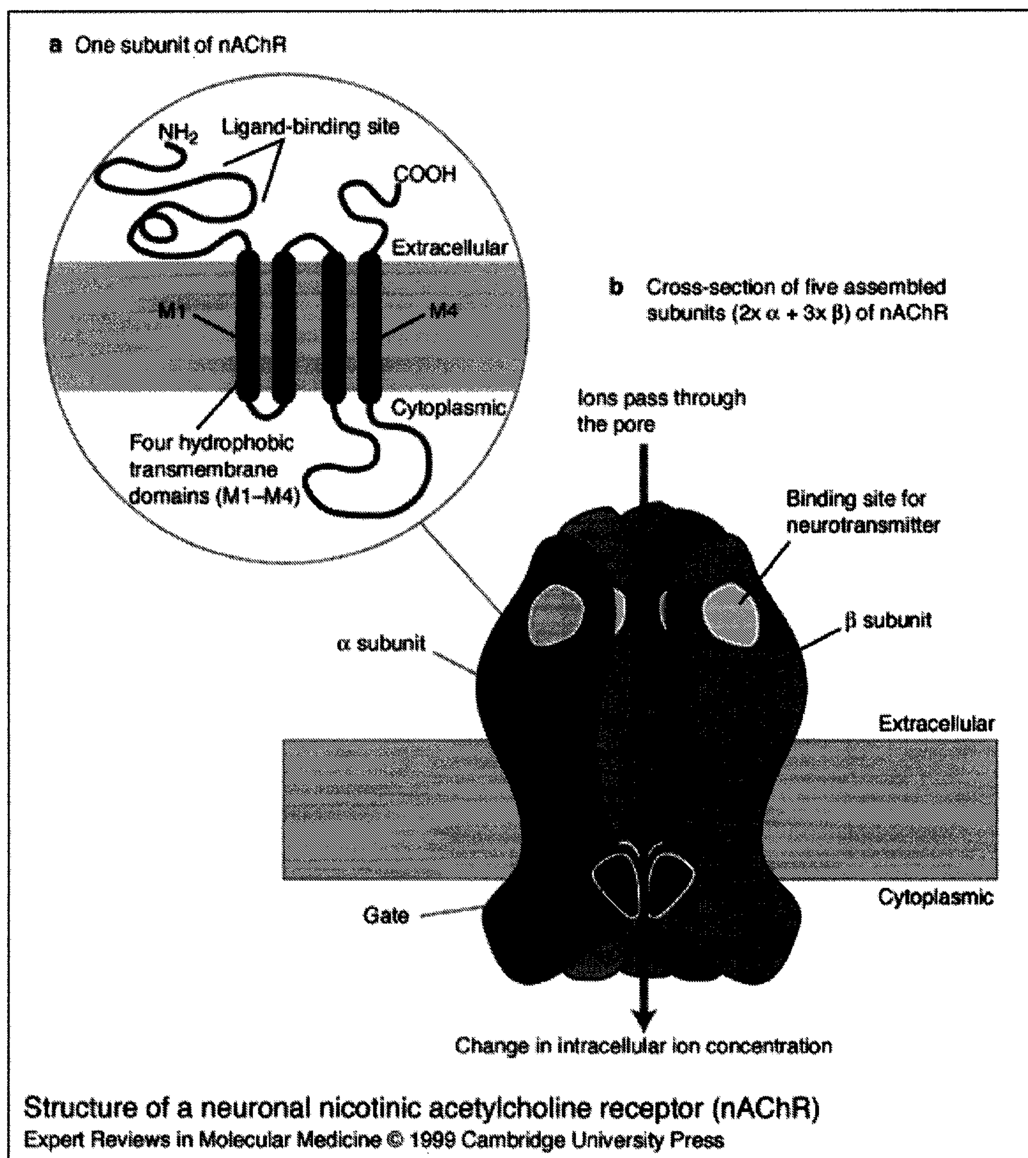


Figure 1. Structure of A Neuronal Nicotinic Acetylcholine Receptor.

Pentameric structure of a neuronal heteromeric nicotinic acetylcholine receptor indicating α and β type subunits, binding sites for neurotransmitter, the ion channel and topology of a ligand-binding subunit. (adopted from Louise Bate and Mark Gardiner, Expert Reviews in Molecular Medicine, Cambridge University Press 1999).

nAChRs are pentameric structures that are made up of a combination of individual subunits. The central ion channel delineated by five

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

subunits functions as an allosteric protein in which effects arising from the binding of a ligand to a site on the protein can lead to changes in another part of the molecule. Each individual subunit consists of an N-terminal extracellular domain; four putative transmembrane domains denoted M1 through M4, among which M2 is believed to be the pore-lining domain; a cytoplasmic loop between M3 and M4; and other shorter loops connecting the domains (**Figure 1**). A ligand-binding subunit, designated as the α-type (characterized by neighboring cysteine residues in the vicinity of neurotransmitter binding sites), contains binding sites for numerous agonists and antagonists on their extracellular domain. While structural subunits, missing the binding sites, are refered to as the β-type subunits.

Functional studies have revealed that nAChRs contribute to the control of resting membrane potential, modulation of synaptic transmission and mediation of fast excitatory transmission. A variety of structural and genetic experiments have suggested an all-β sheet folding of the N-terminal extracellular domain, with the connecting loops contributing to the acetylcholine (ACh) binding pocket and to the subunit interfaces that mediate the allosteric transitions between conformational states. The allosteric transitions of the protein underlying the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

physiological ACh-evoked activation and desensitization possibly involve rigid body motion of the extracellular domain of each subunit, linked to a global reorganization of the transmembrane domain responsible for channel gating (Corringer and Changeux *et al*, 2000).

Neuronal nAChRs are widely distributed in the nervous system. The diverse ion channels assembled from multiple subunits involved in synaptic signaling, and participate in key developmental processes (Role and Berg 1996; Dani *et al.* 2001; Rezvani and Levin 2001; Nai *et al.* 2003). Recent data from knockout animals has extended the understanding of nAChR function. The activity of these ionotropic receptors contributes to a wide range of brain processes from cognitive functions such as learning, memory formation or reward (Jones *et al.* 1999), to cellular events such as degeneration (Zoli *et al.* 1999) and neural development (Jones *et al* 1999; Lauder 1993; reviewed by Linstrom 1997). Abnormalities in nAChR number and function have been implicated in a number of human neuropathologies such as schizophrenic syndrome (Freedman *et al.* 1995; Gault *et al.* 2003; Faraone and Tsuang 2004), Alzheimer's (Wilson *et al.* 1995; Robbins *et al.* 1997; Zoli and Changeux *et al.* 1999; Narahashi and Zhao *et al.* 2003; Dougherty and Nichols 2003; Cook and Rubinsztein *et al.* 2004), *Gilles*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

de la Tourette's syndrome (Sandyk 1995; McEvoy and Allen 2002), Autism (Martin-Ruiz and Perry *et al.* 2004) and Parkinson's diseases (Whitehouse *et al.* 1983; Aubert *et al.* 1992). nAChRs also play a significant role in nicotine addiction (Marks and Collins *et al.* 1992; reviewed by Larsson and Engel 2004), which is another major public health concern. A genetically transmissible epilepsy, ADNFLE (Autosomal Dominant Nocturnal Frontal Lobe Epilepsy), has been associated with specific mutations in the gene coding for the $\alpha4$ or $\beta2$ subunits, which lead to altered receptor properties (Kuryatov and Lindstrom *et al.* 1997; Meisler and Escayg *et al.* 2001; Hirose and Mitsudome *et al.* 2002, Loscher *et al.* 2003). The essential physiological role of neuronal nAChRs in synaptic signaling and the dire consequences of nicotinic synapse-related pathology underscore the importance of defining molecular mechanisms that regulate nAChR expression in the vertebrate system.

Diversity of Neuronal nAChR Subtypes

Perhaps the most striking finding is the considerable diversity of neuronal nAChRs. This diversity takes several forms: open-channel conductance, gating, agonist and antagonist potencies. At least one source of this diversity is the multiplicity of nAChR genes (Deneris *et al.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1991).

There is a great diversity of nAChR subtypes expressed in vertebrates. To date, 12 neuronal nAChR subunit genes have been cloned from vertebrates (Le Novere and Changeux 1995; Elgoyhen *et al.* 2001). In the CNS and PNS, different cholinoceptive neuron populations express distinct combinations of neuronal nAChR subunit genes. Within a single cholinoceptive neuronal population, the different subunits segregate to assemble into two homomeric- or heteromeric-receptors that differ in subunit composition, functional properties, and spatial distribution (Gundelfinger 1992). There is great interest in defining, within one neuronal population, the diversity of nAChR subtypes that are expressed and their subunit composition. This interest stems from the observation that nAChR functional properties are exquisitely sensitive to subunit composition (Sargent 1993; McGehee and Role 1995). The two general classes of neuronal nAChRs have different biophysical properties. In particular, the α-BgtX-sensitive receptors have higher level of calcium permeability, and in some neuronal populations faster kinetics of activation and desensitization, but a slowly desensitizing response in others (Seguela *et al.* 1993; Ullian *et al.* 1997; Chang and Berg 1999; Cuevas *et al.* 2000). Within one neuron, the diverse nAChR subtypes are spatially segregated relative to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

one another and target to discrete synapse-associated sites: the presynaptic terminal, the specialized postsynaptic membrane, and the perisynaptic dendritic surface membrane (Jacob and Berg 1983; Jacob *et al.* 1984, 1986; Moss and Role 1993; Horch and Sargent 1995; McGehee and Role 1995; Gray *et al.* 1996; Shoop *et al.* 1999). The spatial segregation and distinct biophysical properties of the diverse nAChR subtypes are likely to create functionally specialized synapse-associated microregions and establish distinct spatial and temporal patterns of calcium influx that locally target different downstream signaling events (reviewed by Berg and Conroy 2002).

It is increasingly clear that neuronal function is affected not only by ion channel properties but also by receptor subtype heterogeneity. The enormous diversity of nAChR subunit gene expression has raised many questions: How many α and β subunit genes are expressed in one individual neuron? What is the functional specificity of different nAChR subtypes? How do neurons govern the expression and distribution of various receptor subtypes? What is the control mechanism for intracellular trafficking and surface expression of the different nAChR subtypes? In this laboratory, we choose to address the important issue of transcriptional regulation of nAChR subunit genes. We hypothesize

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that co-expression of a ligand-binding subunit and a structural subunit is sufficient for receptor assembly of a representative heteromeric receptor subtype. The coordinate regulation of ligand-binding and structural subunits at the level of transcriptional regulation is testable with our experimental approaches using molecular cloning and *in vivo* studies in *Drosophila*.

Extensive genetic and biochemical analysis of *Drosophila melanogaster* has made this system an important model for characterization of transcriptional regulatory elements and factors. Many fundamental ideas – such as *cis*-acting control elements that act over long distances, the regulation of development by hierarchical cascades of transcription factors, dosage compensation, and position effect variegation- originated from studies of the fruit fly. Besides being a most convenient, economical animal model for its small size, short life cycle, fast reproduction rate, the entire *Drosophila* genome has recently been sequenced. The *Drosophila* system also carries the "magic markers" of polytene chromosomes, enchants biologists for several decades with the powerful tool of P-element transformation and availability of balancer chromosomes for genetic manipulation. Given the striking conservation of transcriptional controls in metazoans,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

general principles derived from studies of *Drosophila* are expected to continue to illuminate transcriptional regulation in other systems, including vertebrates. Therefore *Drosophila* has been selected as the animal model for various research projects in our laboratory. I would like to dedicate the next section for a discussion on insect nAChR subunit genes before I lay out specific aims for my thesis research topic.

Insect nAChR Subunit Genes

The insect nervous system is an unusually rich source of nicotinic-type receptors. Most sensory input pathways as well as many CNS interneurons and projection neurons use ACh as the excitatory transmitter, while in contrast with vertebrates, neuromuscular transmission appears to be glutaminergic (Sattelle 1992; Tornoe *et al.* 1995).

Molecular probes from vertebrate receptors have opened up the way for the molecular cloning of insect nAChR DNAs. The cDNA encoding the γ subunit of the nAChR from the electroplax of the fish *Torpedo californica* was used by two groups to isolate the cDNA and the gene for the first non-α subunit from *Drosophila*: the AChR64B gene

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Hermans-Borgmeyer 1986; Sawruk and Gundelfinger 1988; Wadsworth and Nelson 1988). This gene and the cDNA of the first α-like subunit (ALS) were identified by their cross-hybridization with a fragment of the chicken neuronal α2 subunit gene (Bossy and Spierer 1988). In general the sequence similarity of insect nAChR protein subunits appears to be higher when compared with vertebrate neuronal receptors than with the nAChR subunits expressed in vertebrate muscles.

Like their vertebrate counterparts, insect nAChRs display a substantial diversity, which becomes apparent at the pharmacological and physiological level, as well as by the number of different genes encoding nAChR proteins. To date, there are ten identifiable nAChR genes in *Drosophila* (see Table 1). Among the ten nAChR subunits so far cloned from the fruit fly *Drosophila melanogaster*, seven have been classified as α subunits (ALS, Dα2, Dα3, Dα4; Dα5; Dα6 and Dα7) (Bossy *et al.* 1988; Sawruk *et al.* 1990; Schulz *et al.* 1998; Grauso *et al.* 2002; Lansdell and Millar 2000, 2004) and three as non-α or β subunits (ARD, SBD and Dβ3) (Hermans-Borgmeyer *et al.* 1986; Sawruk *et al.* 1990; Lansdell and Millar 2002).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As in the vertebrate nervous system, where the majority of nAChRs are not affected by snake α toxins, receptor subtypes insensitive to α-BgtX were disclosed by elecrophysiological and pharmacological studies on grasshopper (Goodman and Spitzer, *et al.* 1979), cockroach (Lees and Botham *et al.* 1983), locust (Macallan *et al.* 1988) and *Drosophila* (Gorczyca and Wu *et al.* 1991) neurons. Furthermore, a detailed examination of the association, dissociation and equilibrium binding of $[^{125}I]\alpha$-BgtX to *Drosophila* head membrane extracts has revealed a heterogeneity within the toxin-binding receptors. At least two physically distinct classes of high-affinity binding sites were observed with K_D values of ~0.1 nM (Class 1) and ~4 nM (Class 2) (Schloss and Gundelfinger *et al.* 1988). Class 1 represents about 25% and Class 2 about 75% of all high affinity α-BgtX binding sites found in the heads of flies.

Since identification of insect nAChR subunits was initially accomplished by cloning the genes and cDNAs via cross-hybridization with vertebrate probes, a correlation between the encoded proteins and the pharmacologically and electrophysiologically defined receptors is required. Two different approaches are in progress to address this

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

problem: the production of subunit-specific antibodies, and the expression of insect nAChR subunits in the *Xenopus* oocyte system.

Antisera against bacterially expressed fusion proteins from two different regions of ARD subunit (see **Table 1**) immunoprecipitated class 1 (but not class 2) high-affinity α-BgtX binding sites in *Drosophila* (Schloss and Gundelfinger 1988). As ARD itself is not expected to bind the toxin, it was proposed that ARD is a structural subunit of class 1 receptors. Based on indirect evidence, i.e. immunoprecipitation of α-BgtX binding sites, it has been suggested that ALS and ARD are components of the same receptor complex (Schloss *et al.* 1991, 1992). Subunit-specific antibodies directed against the cytoplasmic regions of various nAChRs subunits have been used to study the subunit composition of *Drosophila* nicotinic receptors. *In vivo* immunoprecipitation data shows that Dα3 (see **Table 1**) and ARD subunits co-assemble within one type of α-BgtX binding receptor complex (Chamaon and Gundelfinger *et al.* 2000), and that ALS and Dα2 (see **Table 1**) are integral components of another receptor complex (Schulz *et al.* 2000). So far at least three physically distinct receptor complexes in the *Drosophila* synaptic neurophil has been proposed: one including the subunits ALS, Dα2 and SBD (see **Table 1**); the second

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

harbouring Dα3 and ARD – but not the other three subunits; the third with ALS and ARD. All three heteromeric receptor complexes are considered as toxin-binding receptors.

Considerable difficulties have been encountered in attempts to express functional nAChRs by heterologous expression of cloned insect nAChR subunits (reviewed by Millar 1999; Gundelfinger and Schulz 2000). Heterologous expression studies have been reported with six cloned *Drosophila* nAChR α subunits [Dα1/ALS, Dα2/SAD, Dα3, Dα4, Dα6 and Dα7 (see **Table 1** for all *Drosophila* subunit genes), (Bossy *et al.* 1988; Sawruk *et al.* 1990; Schulz *et al.* 1998; Lansdell and Millar 2000, 2004)] and with three β subunits [Dβ1/ARD, Dβ2/SBD and Dβ3 (Hermans-Borgmeyer *et al.* 1986; Sawruk *et al.* 1990; Lansdell *et al.* 1997; Lansdell and Millar 2002)]. Although no combination of these seven cloned nAChR subunits has been shown to be capable of forming functional nAChRs or a high-affinity binding site for nicotinic radioligands, this has been achieved when *Drosophila* α subunits have been co-expressed with vertebrate nAChR non-α subunits (Bertrand *et al.* 1994; Lansdell *et al.* 1997; Schulz *et al.* 1998; Lansdell and Millar 2000); while in the case of Dα6 and Dα7, none forms nicotinic radioligand-binding sites, even when co-expressed as heteromeric

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

complexes with vertebrate non-α subunits, but only to chimeric subunits (Dα6/5HT_{3A} and Dα7/5HT_{3A}) (Landsdell and Millar 2004). The more likely explanation for problems in heterologous expression may be that *Drosophila* nAChRs require host-cell specific factors which are not provided by expression systems such as *Xenopus* oocytes, mammalian kidney cells or *Drosophila* S2 cells (Lansdell and Millar 2004). Also, it is possible that efficient folding and/or assembly requires essential neurone-specific chaperone/interacting proteins (Jeanclos *et al.* 2001). Whilest relatively few proteins have been identified which interact with neuronal nAChRs, there have been recent reports that such proteins can influence nAChR assembly and trafficking, for example the mammalian chaperone protein 14-3-3η (Jeanclos *et al.* 2001) and the *Caenorhabditis elegans* transmembrane protein RIC-3 (Halevi *et al.* 2002).

Theoretically, the varying combination of the ten *Drosophila* nAChR subunit genes could give rise to thousands of receptor subtypes. It is currently unknown how many receptors are actually assembled into various subtypes and in which way these receptor subtypes functionally differ. Regulatory mechanisms that govern the expression and distribution of the distinct nAChR subtypes within one neuron are essential for the proper formation and function of interneuronal nicotinic synapses.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1. List of identified nicotinic acetylcholine receptor subunit genes in *Drosophila melanogaster*

GENES	Synonyms	Chromosomal location	cDNA Accession#	Protein Accession#	Length of Protein (aa)	BAC Clone Match
als	*nAcRα-96Aa*, CG5610	96A1-A2	X07194	AAF56301	567	BACR13M04
Dα2	*nAcRα-96Ab*, CG6844	96A3	X52274	CAA36517	576	BACR47D16
Dα3	*nAcRα-7E* CG2302	7E1-E2	Y15593	AAF46361 CAA75688	783, 795	BAC RP98-4P5 (RPCI)
Dα4	*nAcRα-80E*	80F1-F3	AJ272159	CAB77445 AAL28838	568 384	BAC RP98-22B15 (RPCI)
Dα5	*nAcRα-34E*, CG4498, BG:DS05899.4	34F3-F4	AF272778	AAM13390 AAK67256	807, 482 570	BACR18J08
Dα6	*nAcRα-30D* CG4128	30D1-D3	AF321445	AAM13396 AAF52817 AAM13395 AAM13394	523, 544 509, 494	BACR22H06
Dα7	*nAcRα-18C* CG8109, CG8082	18C2-C3	AY036614	AAK67257 AAF48950	219, 545	BACR02B03
ard	*nAcRβ-64B*, CG12606	64B6	X07956	CAA27641	521	BAC RP98-5M7 (RPCI)
sbd	*nAcRβ-96A*, CG6798, *Dβ2*	96A5	Y14678	CAA39211	519	BACR47D16
Dβ3	*nAcRβ-21C*, CG11822	21D1	AJ318761	AAF51485	541	BACR48E08

Molecular Mechanisms That Regulate Neuronal nAChR Expression

The important issue, how differential expression of the genes encoding the receptor in the temporal and spatial restricted fashion required for the orderly development of the nervous system, has been exploited more extensively in recent years. There is evidence that the expression of nAChR is not only determined by an invariable genetic program but also subject to regulatory signals provided by neurotrophic factors (Yang and Role 1998; Belluardo and Fuxe *et al.* 2000), synaptic partners (Rosenberg and Jacob *et al.* 2002), soluble factors (Harborne and Smith 1988) and electrical activity (Levey and Jacob 1995). That is, the combination of the above factors act on neurons to trigger signaling events that control nAChR level. In this context, several levels of regulation, including transcriptional as well as posttranscriptional control, should be considered.

One level at which nAChR expression can be regulated is the transcriptional control of the subunit genes. The $\beta4$, $\alpha3$ and $\alpha5$ subunit genes from a unique gene cluster in mammals and birds (Boulter *et al.* 1990; Couturier *et al.* 1990; Raimondi *et al.* 1992). These three subunits can co-assemble into the same receptor complex (Conroy and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Berg 1995). The gene cluster is organized with β4 being most 5′, followed by α3 and α5; α5 is transcribed in the opposite direction from β4 and α3. These three genes are often coexpressed in central and peripheral neurons and this, along with their tight linkage, suggests coordinate regulation (Boulter *et al.* 1990; Corriveau and Berg 1993; Mandelzys *et al.* 1994). A neural-specific enhancer element identified in the 3′ untranslated exon of the β4 gene may be important for regulating neuronal expression of these clustered nAChR subunit genes (McDonough and Deneris, 1997, McDonough *et al.* 2000). These data suggest that different transcription factors likely mediate the regulatory effects of the distinct synaptic partners.

The *Drosophila* genes nACRα96Aa, nACRα96Aab and nACRβ96A encoding the nAChR subunits ALS, Dα2 and SBD map to the same chromosomal band in the 96A region of the third chromosome. Alignment of the available cDNA and EST sequences (Flybase at: www.flybase.org/) with the corresponding genomic region revealed that the transcription units for ALS and Dα2 are arranged head-to-head, whereas those encoding Dα2 and SBD are oriented tail-to-tail (Chamaon and Gundelfinger *et al.* 2002). No other transcription unit is predicted within this region, indicating that the three subunits are encoded in a gene cluster that spans about 120kb. In both cases

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

clustering may facilitate coordinate expression of the nAChR genes and regulation of coassembly of the respective subunits (Chamaon and Gundelfinger *et al.* 2002).

In addition to the existence of large numbers of genes encoding many of the subtypes of channel or receptor subunits, diversification of components of these molecules also utilizes many of the other molecular mechanism capable of generating variation. With the completion of genome sequencing, interest has been focused on the diversity in gene products arising from alternative splicing and RNA editing (Caceres and Kornblihtt 2002; Schaub and Keller 2002). Both are important post-transcriptional modifications that play a role in expanding protein diversity in an organism with a small genome such as *Drosophila* while the gene number does not correspond to animal complexity. Comparing with the large families of nAChR genes identified: 11 in mammals (Le Novere and Changeux 1999; Elgoyhen *et al.* 2001), 27 in *C. elegans* (Mongan *et al.* 1998; Sattelle *et al.* 2002) and 28 in *Fugu rubripes*, the pufferfish (Jones and Sattelle *et al.* 2003), the ten numbers of the *Drosophila* nAChR subunit gene so far identified does not seem to make sense in terms of its more complex CNS with more than 200,000 neurons, while *C. elegans* has only 302 neurons, yet a larger number of genes. Therefore, variants of gene products at the genomic and mRNA

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

level must be playing a very important role.

Alternative splicing as a means of generating diversity is particularly prominent in the primary subunits of large-conductance Ca^{2+}-activated K^+ channels. The gene encoding these proteins has at least 4-5 independent alternative splicing sites, with the possibility of generating a large number of isoforms, of which more than 10 have already been found. This alternative splicing serves to produce a large variety of Ca^{2+}-activated K^+ channels differing in Ca^{2+} sensitivity, and may be a key molecular mechanism used to tune individual auditory hair cells to respond to specific sounds. The alternative splicing of these molecules may be modulated under interesting physiological conditions (Xie and McCobb 1998). The *Drosophila* nAChR subunit Dα4 (Lansdell and Millar 2000) and Dα6 have been reported to be subject to alternative splicing diversity. Furthermore, alternative splicing in Dα6 involves exons encoding nAChR functional domain (Grauso 2002).

RNA editing is another mechanism to modify channel components posttrancriptionally (Scott 1997; O'Connell 1997). It is first discovered in AMPA glutamate receptors by Peter Seeburg, and has now been observed in Na^+ channels (Reenan *et al.* 2000), K^+ channels (Patton *et al.* 1997), serotonin receptors (Burns *et al.* 1997) and neuronal nAChRs (Hoopengardner and Reenan *et al.* 2003). RNA editing can profoundly

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

alter ion selectivity and gating properties of ion channels (Sommer *et al.* 1991; Higuchi *et al.* 2000) as well as subunit maturation and transport from the endoplasmic reticulum to the cell membrane (Greger *et al.* 2002). Transcripts that are edited in *Drosophila* also encode ion channel subunits, but the numer of sites per transcript is more than is found in vertebrates. The *cacophony (cac)* transcript encoding the pore-forming $\alpha 1$ subunit of a voltage-gated calcium channel is edited at 12 sites (Kawasaki *et al.* 2002).

Editing of nAChRs in *Drosophila* occurs in several conserved regions including the ligand-binding domain, transmembrane (TM) domains involved in gating and agonist sensitivity and, importantly, the lining of the conducting pore, TM2. Editing sites found in TM2 in ion channels are in hydrophilic or charged residues that have been shown by mutagenesis studies to affect ion selectivity and channel conductance. Given the spatially and temporally hierarchical events that occur in neurons on the time scale of milliseconds, it is proposed that the advantage of regulating VGICs, LGICs, and the synaptic release machinery by RNA editing lies in a finer level of control than discrete genetic change permits (Hoopengardner and Reenan *et al.* 2003). The first case in any organism where a nAChR gene is the target of mRNA editing is Dα6, whose editing is dADAR dependent, seven adenosines could be modified

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in the extracellular ligand-binding domain region of Dα6, four of which are also edited in the Dα6 ortholog in the tobacco budworm Heliothis virescens. The conservation of an editing site between the insect orders Dipter and Lepidoptera makes nAChR editing the most evolutionarily conserved invertebrate RNA editing site so far described (Grauso *et al.* 2002).

As we know, the electrical properties of a cell are produced by the coordinated activity of a suite of channels that conduct ions across the membrane. To manifest the appropriate electrical character, a cell must regulate gene expression so that the proper subset of ion channels is expressed at appropriate levels. The need for controlled and coordinated expression makes it likely that the transcriptional control region of nAChR subunit genes contributes a key role resulting in collections of phenotypically characteristic neurons.

Despite the wealth of information available on the gene expression and physiological profiles of neuronal nAChRs, our knowledge of the molecular mechanisms involved in the transcriptional regulation of nAChRs is still rudimentary. Since neuronal nAChR subunit genes exhibit distinct, yet overlapping temporally and spatially restricted pattern of expression, which are consistent with the proposed heteromeric compositions of neuronal nAChRs in vivo, these patterns of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expression suggest specific molecular mechanisms for generating functional diversity within the nervous system (Hu *et al.*, 1995). The discrete DNA elements that are directly involved in regulating the transcription of genes, and the proteins that interact specifically with these elements are the emphasis of a molecular biologist.

One aspect of the research in our laboratory has focused on transcriptional control of the *ard* gene. Previous thesis project by Lee-Chuan Yang has established the transcription start site and minimal promoter of the *ard* gene. Germ line transformation with fusion constructs of upstream sequences with an *E.coli* β-galactosidase reporter gene was used to isolate a 325 nucleotide fragment which confers spatial and temporal correct expression in embryo, pupae and adult when compared with the known distribution pattern of the *ard* mRNA by *in situ* hybridization (Hermans-Borgmeyer 1989). Deletion of further 156 nucleotides on the 5' side of the upstream sequence abolished *lacZ* expression (Yang 1996).

The *Drosophila* gene encoding a nAChR α-like subunit ALS was isolated by crosshybridization with a vertebrate probe (Bossy *et al.* 1988). The ALS encoding gene exceeds 54 kb in length and the transcript has a very long and unusual 5' leader. The ALS gene product most resembles the neuronal set of vertebrate nAChR α-like subunits

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that do not bind αBgTx. We were tempted to the project of studying the transcriptional regulation of *als*, because it had been suggested that ALS and ARD are components of the same hetero-oligomeric receptor complex based on immunoprecipitation of α-BgtX binding sites (Schloss *et al.* 1991, 1992). The *als* gene has been mapped to chromosomal region 96A (Bossy *et al.* 1988) while *ard* is located at 64B10-B12 (Sawruk *et al.* 1988). We expect not only to make a fundamental contribution to our understanding of the developmental regulation of the cholinergic receptor in general, but also add to our previous knowledge about the structural-subunit-encoding gene *ard*, whose gene product is part of the same receptor complex with ALS subunit, that constitutes one of the major nAChR subtypes (Gundelfinger and Hess 1992). It would be very interesting to investigate if the expression of the *als* and *ard* genes are coordinately regulated; furthermore, if the same molecular switches are used in all neurons, which express various nAChR genes and which other neuronal proteins are under control of the same transcription factors.

The specific aims of this study are: (1) to characterize the molecular structure of the *als* gene, which encodes the ligand-binding subunit of a neuronal nAChR in *Drosophila* central nervous system; (2) to identify the *als* 5′ upstream regulatory sequence which governs its spatial and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

temporal expression. Specifically, the contribution of transcriptional mechanisms in the regulations of *als* expression levels in *Drosophila* CNS was assessed and evaluated by several complementary approaches.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 1.

IN SEARCH OF GENOMIC DNA CLONES CARRYING THE *als* GENE AND ITS UPSTREAM REGULATORY SEQUENCE

The complex temporal and spatial expression patterns of developmentally regulated genes often reflect the action of modular regulatory elements, generally 300-700 bp in size that can be located up to 100 kb 5′- from the transcription start site. To begin with the investigation of the transcriptional regulatory elements controlling the expression of the *Drosophila* neuronal nAChR *als* subunit gene, my first goal was to isolate genomic clones that harbor the *als* gene and its 5′-flanking sequence. My effort to reach this specific goal evolved from the initial attempt to isolate *als*-containing genomic clones from bacteriophage λ library of *Drosophila melanogaster* early in 1997, to screening P1 clones in 1998, until Bacterial Artificial Chromosome (BAC) became available as the Berkeley *Drosophila* Genome Project (BDGP) and Celera Genomics collaborated in 1999-2000 to sequence the *Drosophila melanogaster* genome. Two different experimental approaches are presented in this chapter, both of which aim at the same specific goal. "Material and Method" and "Result" are described for each approach, followed by a "Discussion".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I. Experimental Approach I

Early Strategy: Conventional Genomic Library Screening with *als*-specific probe

1. Material and Method

A genomic DNA library from *Drosophila melanogaster* (Canton S) constructed in the lambda bacteriophage EMBL3 SP6/T7 vector (Clontech Inc.) was screened at low stringency with a ^{32}P-end-labeled oligonucleotide probe derived from the 2576-2599 bp of the *als* gene (based on the published *als* mRNA sequence, GenBank Accession No. X07194, Bossy *et al.* 1988) (**Figure 2A**). The 24 mer oligonucleotide: 5′-cgcagaaccgccgctgcc ctcctc -3′ was verified to be specific to *als* (1) by blast search against the Drosophila genomic DNA database; (2) by sequence comparison with the then-known subunit genes *D*α*2*, *ard* and *sbd* using GCG program (**Figure 2B**). The GCG package is a collection of programs used to analyse or manipulate DNA and protein sequence data. The *als* probe was synthesized by the Sequencing and Synthesis Facility of Hunter College.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

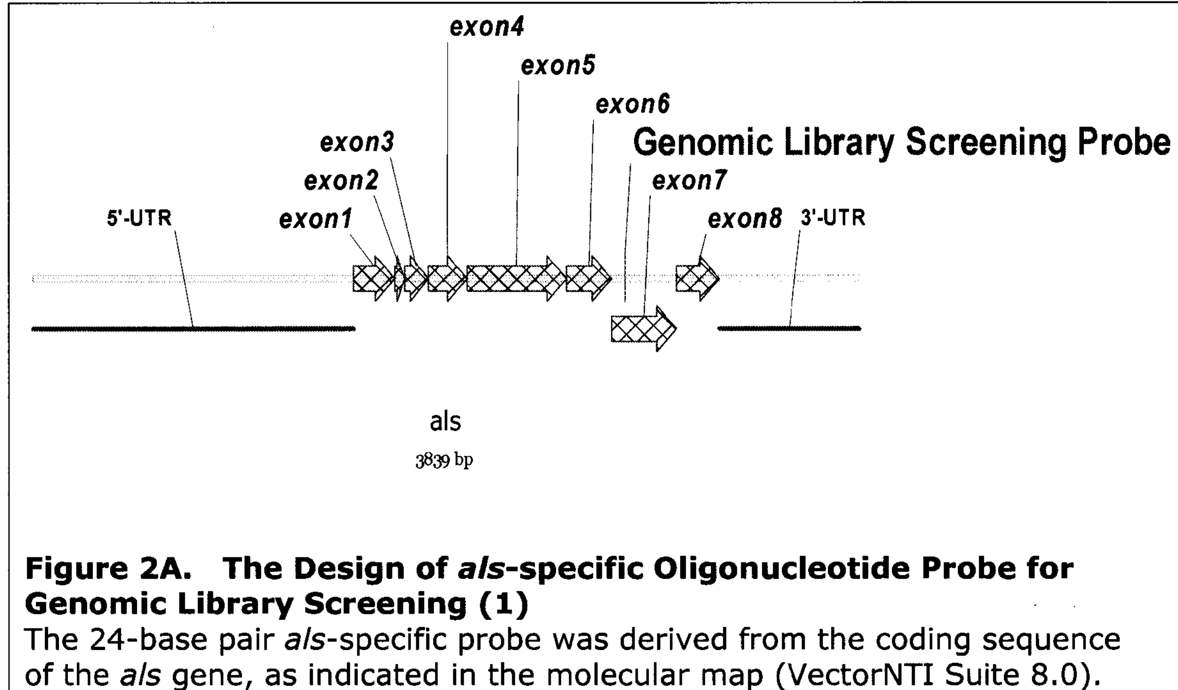


Figure 2A. The Design of *als*-specific Oligonucleotide Probe for Genomic Library Screening (1)
The 24-base pair *als*-specific probe was derived from the coding sequence of the *als* gene, as indicated in the molecular map (VectorNTI Suite 8.0).

Approximately 1.3 x 10^7 phages were first hybridized under low-stringency condition (6X SSPE, 5X Denhardt's, 0.5% sodium pyrophosphate, 100μg/ml sheared and denatured salmon sperm DNA, and 0.25% SDS) at 45°C overnight. The filters were then washed with two changes of 6X SSPE. After washing, autoradiography was carried out and the film was developed. The second and tertiary screening conditions were the same as the first screening except for the high-stringency condition of hybridization temperature at 60°C.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 2B. The Design of *als*-specific Oligonucleotide Probe for Genomic Library Screening (2)
The result of GCG – sequence alignment of the designed probe (in bold letter underlined) with then-known *Drosophila* subunit genes indicated that the designed oligonucleotide probe was unique to the gene of interest, *als*.

```
PileUp of @alpha.list:  Dalpha2, sad, alsspier, sbd, ard
          2501                                              2550
 Dalpha2  cggctccagt ccagactccc tgcggcggat gcaaggtcgt gtgggtgctg
     sad  cggctccagt ccagactccc tgcggcggat gcaaggtcgt gtgggtgctg
alsspier  ctacggcatt ccagcactac ccgcttcgca tcgcttcgat ttggctgcgg
     sbd  caacttaaca cccgaagtgc tgcaggcgtt acg....... ........tg
     ard  tcgcatca.t cgagatttac aggggaaagt aagcgagtga ggacttggag

          2551                                              2600
 Dalpha2  gtgggtgcaa tggcatgcac gtgaccacgg ccacaaacag attcagcggc
     sad  gtgggtgcaa tggcatgcac gtgaccacgg ccacaaacag attcagcggc
alsspier  cgggcgggat tagcgcccac tgtttcgcag aaccgccgct gccctcctca
     sbd  cggtcagatt tattgcgcag catatcaagg atgccgacaa ggataacg..
     ard  cagccggggg ctgattgaaa tggttaacag cattgcaaac gcaacaataa

          2601                                              2650
 Dalpha2  ttg...gtgg gggctttggg cggcggactg agcaccctga gcggctacaa
     sad  ttg...gtgg gggctttggg cggcggactg agcaccctga gcggctacaa
alsspier  ctgccactcc cgggcgccga cgacgatctg ttcagcccgt cgggcctcaa
     sbd  .agattgtgg aggactggaa gttcgtatcg atggtgctgg accgcttctt
     ard  taagcaaacg aggtattgag aca.gaggaa aacaatgaaa attgcaacaa
```

Positive clones after three rounds of genomic library screening were picked using sterile toothpicks and stored in 1ml of SM solution (100mM NaCl, 10mM Mg_2SO_4, 50mM Tris pH 7.5 and 0.1% gelatin). To further prove the identity of the positive clones, southern blotting analysis were carried out according to the procedure described in Current Protocols in Molecular Biology (Ausubel *et al.* 1988).

The λ DNA from positive clones were digested with XhoI (New England Biolab), separated by 0.6% agarose gel electrophoresis (**Figure**

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2C), and then transferred onto a nylon membrane (Zeta-Probe Blotting Membranes) with 0.4N NaOH as transfer buffer via downward capillary transfer. The same 24 mer *als*-specific probe was end-labeled to specific activity at 1.2 X 10^5 cpm/μg. Hybridization was carried out under high stringency at 60°C overnight. The washing condition was at 60°C, with the blot soaking in a rolling tube set in rotating motion in the prewarmed incubator. The next two changes of 6X SSPE for 5 min and 10 min was performed in the same way.

The inserts of the positive λ phage DNA confirmed by Southern blotting were excised by digestion with SfiI, isolated by agarose gel electrophoresis, and purified using a GeneClean Kit (BIO101). The λ DNA fragments obtained from the purification were subcloned into the same digestion site of the phagemid pBluescript KS (+) *E.coli* strains SRB (Stratagene) transformation.

The nucleotide sequence of each clone was determined by automated sequencing performed by the Sequencing & Separations Facility at Hunter College (Center for Study of Gene Structure & Function, Hunter College) using the *als*-specifc probe as sequencing primer. The resulting sequencing information from genomic clones was analyzed using nucleotide-nucleotide BLAST program (Altschul *et al.* 1997) from both NCBI GenBank database (*Drosophila* genomic DNA) and Flybase to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

compare with the *als* gene and its 5'-flanking sequence.

2. Result

Four positive clones designated "C", "M", "Y" and "Z" were isolated after three rounds of genomic library screening. They were initially analyzed by restriction enzyme digestion followed by gel electrophoresis and visualized on ethidium bromide-staining agarose gel. The four clones were shown to be different, with clones "M" and "Z" sharing very similar digestion patterns.

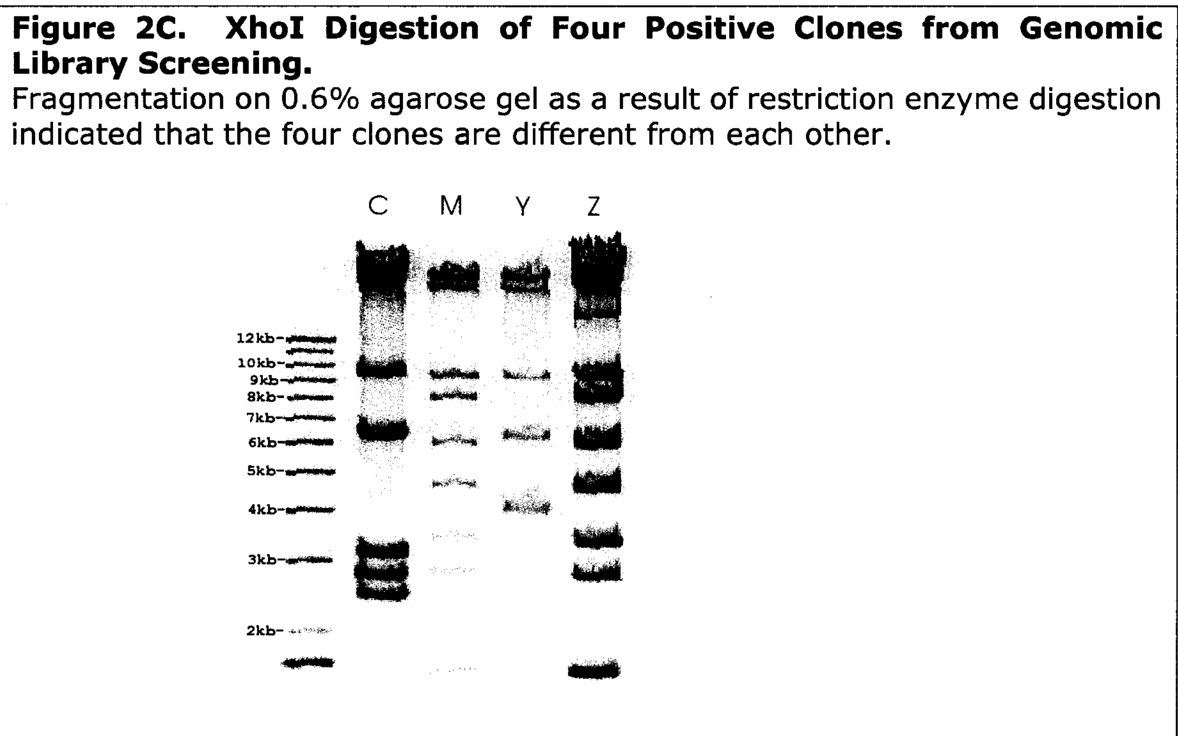
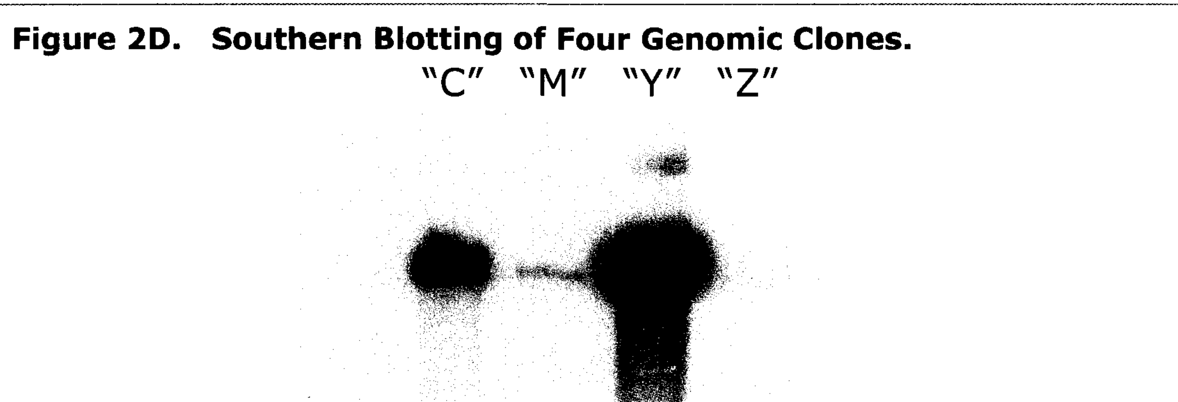


Figure 2C. XhoI Digestion of Four Positive Clones from Genomic Library Screening.
Fragmentation on 0.6% agarose gel as a result of restriction enzyme digestion indicated that the four clones are different from each other.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

To further confirm truly positive clones, the λ DNA of each clone was analyzed by southern blotting using the same 24 mer *als*-specific probe. As shown in **Figure 2D**, clone "C" and "Y" displayed strong positive signals, while clone "M" displayed weak signal. Clone "C" and "Y" were chosen for further analysis.

Figure 2D. Southern Blotting of Four Genomic Clones.



The inserts of the recombinant λ phage DNA "C" and "Y" were subcloned into pBluescript KS (+) for amplification and DNA sequencing analysis. Technique obstacles arose during the DNA sequencing process at Sequencing and Synthesis Facility of Hunter College in 1997. It was a frustrating period of repeated troubleshooting cycles scrutinizing quality and quantity of DNA purified from the subcloned λ DNA clones. Finally, the DNA sequences were obtained from the DNA Sequencing Resource Center at the Rockefeller University. However, based on computer analysis (BLASTN 2.0.8) of the resulting sequencing

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

information from genomic clone "C" and "Y", none of them matched the *als* gene and its 5'-flanking sequence.

II. Experimental Approach II

The Post-Genome Era and the Bacterial Artificial Chromosomes (BAC) – An Alternative strategy to identify and manipulate the *als* locus

1. Material and Method

Large-scale sequencing of human and model organism genomes, cDNAs and expressed sequence tags (ESTs) is identifying tens of thousands of genes about which little is known. Several clone-based physical maps have been described previously. Low-resolution yeast artificial chromosome maps of the genome have been produced by polytene chromosome *in situ* hybridization (Garza *et al.* 1989), and cosmid maps of regions of the X chromosome have been made by STS content and fingerprint mapping (Siden-Kiamos *et al.* 1990). The most complete previous map is the P1-based map constructed by polymerase chain reaction-based 348 sets of contiguously overlapping clones (contigs), each with at least two STS markers (Kimmerly *et al.* 1996). However, the contiguity of the P1 map was limited by the shallow genome coverage of the library (about sixfold) and the relative small

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

insert size of the clones (80 kb).

Because of its historical importance, a large research community, and powerful research tools, as well as its modest genome size, *Drosophila* was chosen in 1990 as a test system to explore the applicability of the whole-genome shotgun (WGS) sequencing for large and complex eukaryotic genomes (Venter *et al.* 1998). Under the auspices of the federally funded Human Genome Project, genome projects in the United States, Europe, and Canada have produced a battery of genome-wide resources. In 1999-2000, the Berkeley and European *Drosophila* Genome Projects (BDGP and EDGP) initiated genomic sequencing and produced a BAC-based physical map and defined a tiling path of overlapping BAC clones that span the euchromatic portion of the genome. The BAC map and other genomic resources available for *Drosophila* serve both as an independent confirmation of the assembly of data from the shotgun strategy and as a set of resources for further biological analysis of the genome (Hoskins *et al.* 2000). Celera Genomics collaborated with BDGP in this vast project with its WGS approach. The annotated genome sequencing of *Drosophila melanogaster* was published on the March 24th 2000 issue of *Science*. The nucleotide sequence of nearly all of the ~120 Mb euchromatic portion of the *Drosophila* genome had been

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

determined using WGS sequencing strategy supported by extensive clone-based sequence and a high-quality bacterial artificial chromosome physical map. The sequence of all the scaffolds (including heterochromatin), predicted transcripts, and predicted proteins are also available on a download page of BDGP (http://www.fruitfly.org/sequence/release4genomic.shtml).

BAC technology allowed researchers to overcome the limitations of the cloning vectors. The backbone of a BAC clone is a ~7 kb vector that can stably hold several hundred kilobases of exogenous DNA (Frengen *et al*. 1999; Hoskins et al. 2000). The BAC is propagated in *E. coli* as a supercoiled plasmid at a copy-number of one per cell. As the BACs are propagated in recombination-deficient E. coli, they are not prone to chimerism (Shizuya 1992).

The RPCI-98 *Drosophila melanogaster* BAC library has been constructed in Children's Hospital Oakland Research Institute (CHORI) laboratory by Aaron Mammoser and Kazutoyo Osoegawa in collaboration with BDGP headed by Gerald M. Rubin and co-directed by Susan E. Celniker at Lawrence Berkeley National Laboratory. This library has two different names as a result of different nomenclature guidelines used by the BGDP and CHORI (BACPAC Resources).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

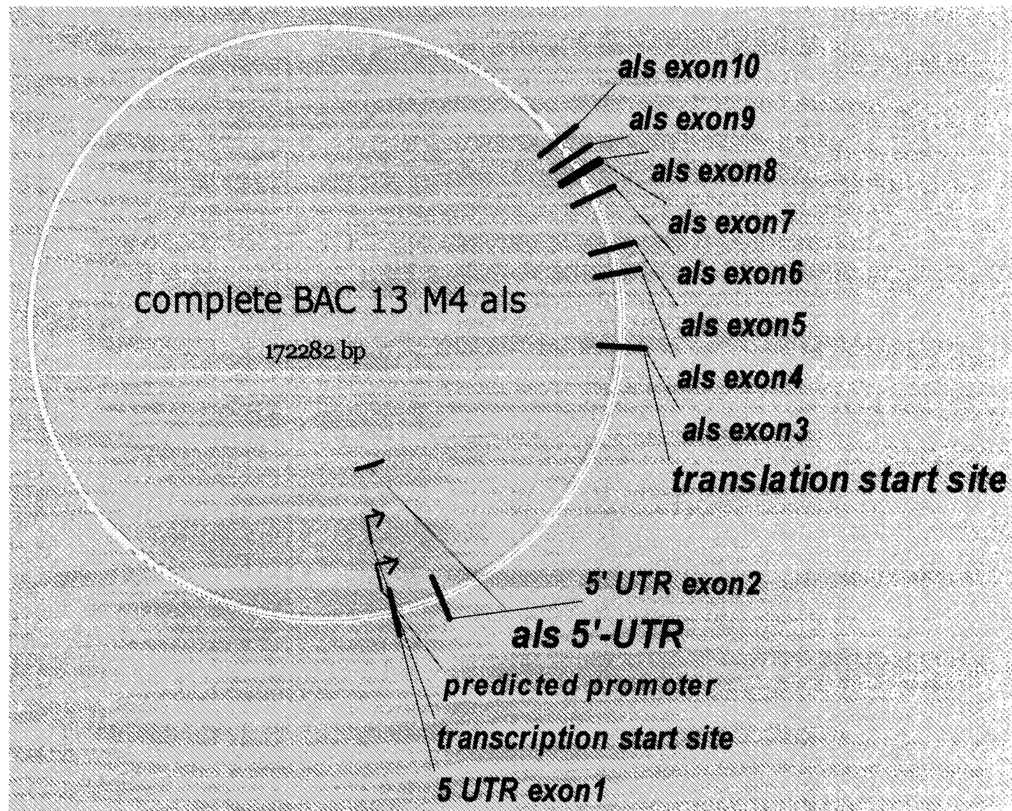
2. Result

The availability of nucleotide sequences of BAC clones allowed us to identify the genomic source that carries the *als* gene and it's 5'-flanking sequence. Blast search of the *als* published cDNA sequence (Bossy *et al.* 1988) against the *Drosophila* database, resulted in the BAC clone "RPCI 98-13.M.4" or "BACR13M04" that carries the gene of interest (Figure 3).

The BDGP followed the established nomenclature for various *Drosophila* BAC libraries and BACPAC Resources followed the NCBI recommendations. Hence, our clone called "BACR13M04" by the BDGP is referred to as "RP98-13M4". The BAC clone was ordered from CHORI (BACPAC Resources), as bacterial LB agar stab culture. Single colonies were streaked onto LB agar plate supplemented with 20 μg/ml of chloramphenicol. BAC clone DNA was purified using QIAGEN Large-construct DNA Preparation Kit. The identity of the BAC clone was checked by PCR with *als*-specific primer set (see Table 3).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 3. The *als* Gene Versus BACR 13M04



The vector NTI molecular map of 172 kb BACR13M4 genomic DNA clone which carries the complete *als* gene and its 5'-flanking sequences. The identified non-coding exon 1 and exon 2, transcription start site, coding exon 3 through exon 10; translation start site, as well as computer-predicted promoter are all indicated on the map as we know now.

Molecular cloning of the upstream *als* sequence using the above described BAC clone for *in vivo* analysis will be further described in Chapter 5.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Discussion

The initial attempt to screen genomic library with radioactive-labeled *als*-specific probe was designed with various cautious steps, including the choice of location at a highly variable cytoplasmic region of *als* gene; its high GC content (75%) for high melting temperature that allows relatively high stringency hybridization conditions; while using several confirming approaches available at the time to make sure it is sequence –specific to the *als* gene. Both nucleotide sequence alignment using GCG pile-up and Blast search (BLASTN 1.4.9MP, 3-26-1996) against the *Drosophila* database were carried out. The original record of blast search with the probe on 10/30/1996 showed a satisfying single match to the *als* published mRNA. However, no expected results were obtained after DNA sequencing verification.

Given the fact that the main repository nucleotide database has been updated in the pubic domain along with rapid progression of on-going genome sequencing projects in the past few years. In 1999, the *als* probe was checked again against the Genbank (BLASTN 2.0.8, 1-5-99). Our record showed four matches: in addition to *als*, there were three mouse genomic matches. As for blast search with the DNA sequencing results of clone C and clone Y, they were found to match an unrelated

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

segment on BACR48E02. We suspect that the short 24 mer *als* probe actually found its way during hybridization not to the *als*-containing clones, but rather any possible matching genomic regions in the vast genome of *Drosophila*. To prove that, an updated blast search with the same 24 mer probe was conducted. As compared to the combined database of GenBank, EMBL, DDBJ, PDB from that of October 1996 with only 274,177 sequences, 386,744,762 total letter, now there is a myriad of sequence database containing 35,532,003 sequences, 40,325,321,348 total letter. Not surprisingly, the current blast search results in more than a dozen matches other than the expected clone (**Table 2**). As people always say, "Hindsight 20/20". This is definetly a good lesson for anyone who designs any primers for various molecular approaches.

After all, the year 2000 completion of the euchromatic DNA sequences of *Drosophila* is a breakthrough that highlight important unanswered questions. For example, only one-fifth of the 124 Mb of Drosophila euchromatic DNA codes for protein. The function of the remaining 100 Mb of mostly unique DNA is largely unknown. Some proportion of this non-reading frame DNA must encode the functional recognition sites targeted by the approximately 700 sequence-specific DNA binding proteins that regulate transcription in *Drosophila*.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ion channels are the physical substrates that underlie neuronal signaling and behavior of higher organisms. With the completion of the *C. elegans* and *Drosophila melanogaster* genome sequences predicting 18,424 (*C. elegans* Sequencing Consortium, 1998) and 13,601 (Adams *et al.* 2000) protein-coding genes, while the human genome sequence nearing completion, its genome is likely to surpass 80,000 protein-coding genes. An initial evaluation and characterization of the molecular machinery that drives three strikingly different nervous systems, with distinct morphologies and levels of complexity: the simple nervous system of worms with 302 neurons, the more complex *Drosophila* nervous system with ~250,000 neurons, and the substantially more elaborate human brain with an estimated 10^{12} neurons. Despite the huge differences in complexity, comparison of genes encoding ion channels and synapse assembly proteins reveals considerable conservation at the molecular level. The basic framework for neuronal signaling and synaptic organization was completed early in metazoan evolution, and evolution of much more complicated nervous systems did not entail extensive alteration of this framework. Thus, genomic analysis fully confirms the belief that one can obtain fundamental neurobiological insights from studies of simple model organism.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2. List of Matching Clones with 24 mer *als*-specific oligonucleotide probe.

List of Matching Clones with 24 mer *als*-specific oligonucleotide probe against All *Drosophila* Database Sequences. Besides the expected *als* gene and the specific genomic region containing *als* gene, other clones with 100% nucleotide similarity were shown in this table.

High Score Matching Clones	Clone Length (matching length)	Genebank Name of the Matching Clone	Chromosome Location of Matching Clone
gb|X07194	3498 (24)	X07194	3R: 96A-96A
AC008366	172282 (24)	BACR13M04	3R: 96A-96A
AC008204	199016 (24)	BACR04E17	3R: 95E-95F
BACR18B10-TET3 (STS)	1109 (24)	BACR18B10	n/a
BACR15G24-TET3 (STS)	977 (24)	BACR15G24	n/a
BACR10A19-TET3 (STS)	933 (24)	BACR10A19	n/a
SEG:AE003747	206741 (24)	gadfly-seqname:AE003747	3R: 95F5-96A2
AC010706	159513 (16)	BACR36D15	X, 13C-13E
AC007581	156508 (16)	BACR03I24	2R, 60E-60E
AC007574	171151 (16)	BACR11C07	2R, 60D-60D
CG9198-RA	6380 (16)	FBgn0030639	X: 13C3-13C4
AC016792	164970 (15)	BACR26A02	3R: 88B-88B
AC016368	80017 (15)	BACR36A03	2R: 42A-42B
SD09926.complete	4797 (15)	EST SD09926	2R: 42A12-42A13

CHAPTER 2.

IDENTIFYING THE TRANSCRIPTION START SITE OF THE *als* GENE

The major transcriptional machinery of many eukaryotic genes are located upstream of their transcription start site(s). We assumed this to be the case for the *als* gene. Therefore, extensive effort has been made to define the transcription start site. A combination of RT-PCR, primer extension assay, RACE-PCR and computer analysis were employed to address this major issue.

When Spierer's lab first reported (Bossy *et al.* 1988) isolation of the *als* gene with 96-195 of the chick neuronal nAChR $\alpha 2$ gene, the structure of the ALS-encoding gene differed from the structure of its vertebrate counterparts and other known *Drosophila* neuronal nAChRs. First, the *als* gene spans more than 50 kb of DNA and its major transcript is 10.5 kb long, the abundance of this transcript reached its maximum at late embryogenesis and remains stable up to early pupal stage, with 2 fold decrease at late pupal stage, followed by a further 3-fold decrease at adulthood (7-days after eclosion). Second, it possesses one of the longest (1282 nt) 5′ leader ever mentioned in literature. The first two

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

exons contain 5′ leader sequences which contain only 15% U and has many GC-rich stretches frequently separated by groups of 2-6 A. This leader sequence is separated from the eight coding exons by 28 kb of intervening sequence (Bossy *et al.* 1988).

Besides these noteworthy features of the *als* gene and its transcript, I had reasons to believe that the 5′-end of the published *als* mRNA was erroneous. Blast search of the existing *als* mRNA (with GenBank Accession Number X07194) against *Drosophila* nucleotide database revealed cDNA versus genomic DNA mismatching. The gaps were between 36-51nt, 745-776nt, 979-1049nt. At this point, it became critical to sort out the sequence puzzle of the 5′-end of the *als* gene before any serious steps can be taken to design experiments to map the transcription start site. To address this important issue, *als*-specific primer sets were designed to "walk" through exon 5 to exon 1 towards the 5′ -end (**Figure 4A**) to ensure that (1) the two non-coding exons do exist, despite of its encoding long leader sequence being separated 29 kb from the coding exons; (2) to verify the sequence contiguity from the non-coding exons to coding exons; and (3) to get some ideas about the transcription start site by assembling the sequencing information.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

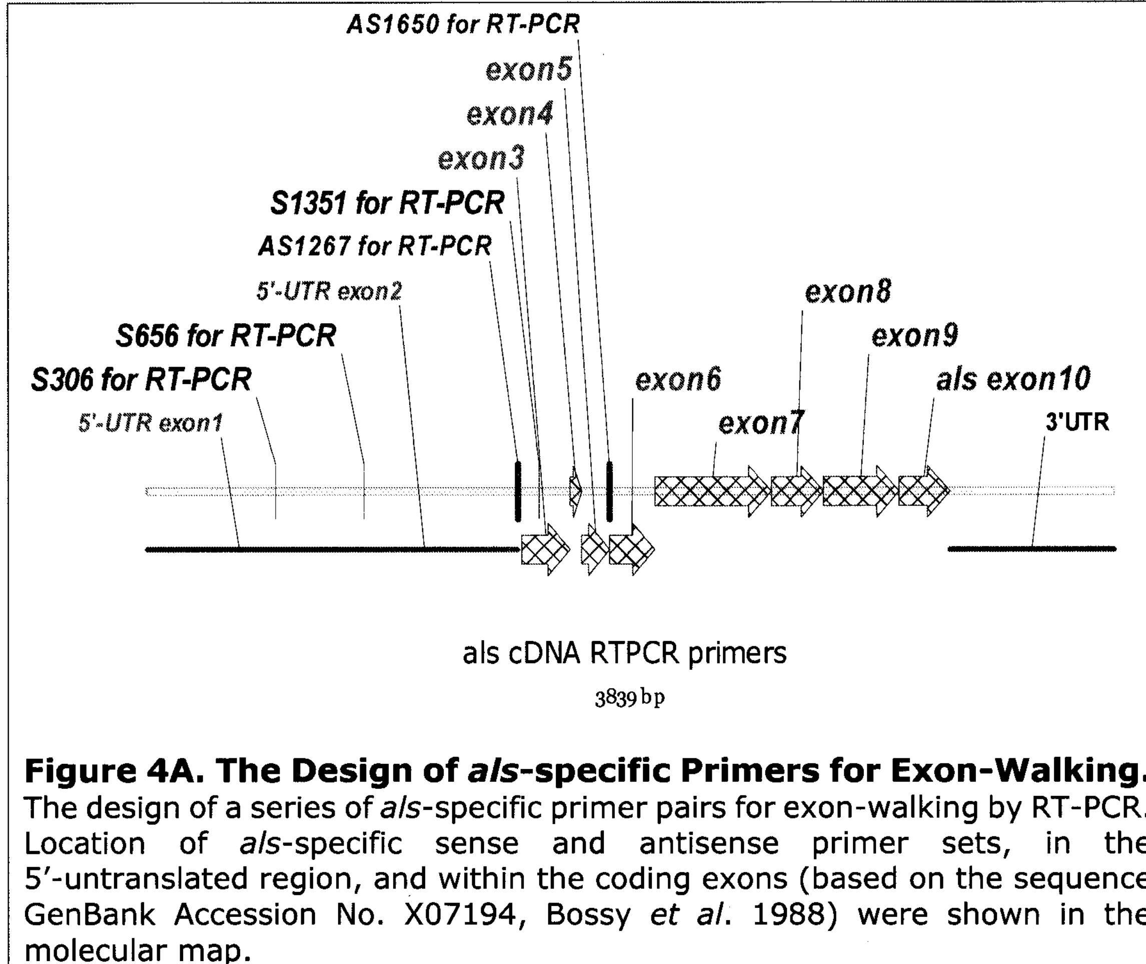


Figure 4A. The Design of *als*-specific Primers for Exon-Walking. The design of a series of *als*-specific primer pairs for exon-walking by RT-PCR. Location of *als*-specific sense and antisense primer sets, in the 5′-untranslated region, and within the coding exons (based on the sequence GenBank Accession No. X07194, Bossy *et al.* 1988) were shown in the molecular map.

In this chapter, I will describe three major experimental approaches that are substantial to studying molecular structure of the *als* gene: (1) RT-PCR for *als* exon-walking; (2) Primer extension analysis; (3) RACE-PCR. The latter two techniques allow mapping of the transcription start site of the *als* gene. Two sections: "Material and Method", "Results" are described below. A combined discussion following bioinformatic analysis will be found in Part I of the next chapter.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MATERIAL AND METHOD

1. Fly Culture

Wild-type Oregon R flies were maintained at 25°C on cornmeal agar media supplemented with yeast. Standard *Drosophila* husbandry protocols were followed in the care, handling, and collection of embryos (Ashburner, 1989). For embryo collection, wild-type flies were propagated in large population cage. Prior to a timed collection, flies were provided with a pre-warmed, fresh and heavily-yeasted food tray to minimize retention of fertilized embryos. Embryos were collected on agar trays in large population cage and aged at 25°C before harvesting. Larvae and pupae were from 6-hour egg collections that were aged and collected in large plastic boxes containing fly food.

2. Preparation of Total and Poly(A)$^+$ RNA

Total RNA was prepared using RNeasy Midi Kits (QIAGEN). 250 mg of frozen tissue from *Drosophila* 12-24 hour embryos was used as starting material. *Drosophila* polyadenylated RNA from 12-24 hour embryos was isolated according to the manufacturer's (Pharmacia) instructions. 1.25mg of total RNA was precipitated from the ethanol suspension and dissolved in 1ml of Elution buffer (10mM TrisCl, pH 7.4

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and 1mM EDTA). The RNA solution was heated to 65°C for 5 minutes and then chilled on ice. After adding 0.2ml of sample solution (Elution buffer containing 3M NaCl), the RNA was applied to an oligo-dT cellulose column which had been pre-equilibrated with high-salt buffer (elution buffer containing 0.5M NaCl). Once the entire RNA solution passed through the column under gravity, the column was centrifuged at 350xg for 2 minutes. Using the same centrifugation routine, the column was washed twice with 0.25ml high-salt buffer and three times with 0.25ml low-salt buffer (elution buffer containing 0.1M NaCl). The bound poly(A)$^+$ mRNA was eluted with four successive 0.25ml aliquots of prewarmed (65°C) elution buffer in the same manner. The poly(A)$^+$ mRNA was stored frozen at -75°C as an ethanol precipitate.

3. RNA gel electrophoresis

The integrity and size distribution of total RNA purified with QIAGEN RNeasy Kits was checked by formaldehyde-agarose (FA) gel electrophoresis and ethidium bromide staining. 1.2% of FA gel was made with 1.2g agarose, 10 ml of 10X FA Gel buffer (200mM 3-[N-morpholino] propanesulfonic acid (MOPS), 50mM sodium acetate, 10mM EDTA, pH to 7.0 with NaOH), and RNase-free water to bring the volume to 100ml. The above mixture was heated to melt agarose,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cooled to 65°C in a waterbath, added 1.8 ml of 37% (12.3M) formaldehyde and 1 μl of a 10 mg/ml ethidium bromide stock solution, mixed thoroughly and poured onto gel support. Prior to running the gel, it was equilibrated in 1X FA gel running buffer for at least 30 min.

RNA samples were added in a 4:1 ratio with 5X loading buffer (16 μl of saturated aqueous bromophenol blue solution, 80 μl of 500mM EDTA, pH 8.0, 720 μl of 37% formaldehyde, 2ml of 100% glycerol, 3084 μl formamide, 4 ml 10X FA gel buffer, bring volume up to 10 ml with RNase-free water). The sample mixture was incubated for 3-5 min at 65°C, chilled on ice, and loaded onto the equilibrated FA gel. The gel was run at 80 V/cm in 1X FA running buffer (100ml 10X FA gel buffer, 20ml 37% formaldehyde, 880 ml RNase-free water).

4. RT-PCR

Successful RT-PCR requires proper design of optimal gene-specific primer pairs, the use of appropriate primer concentrations, and the correct storage of primer solutions. For the purpose of verifying two *als* 5′-non-coding exons (exon 1 and 2) and the coding exons (exon 3, 4 and 5) that are 28 kb far apart, primer pairs were designed (see **Figure 4A**). Following are the Sense(S) and Antisense(AS) *als*-specific primers

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

designed: *als* S656 (located within exon2) 5′-AACATCAGGGACAGGCAGGT-3′; *als* S306 (located within exon1) 5′-CGCATCCAAGCAGCAAACTA-3′; *als* AS1267 (located within exon3) 5′-CCATCTTAGGCCAGGTCCAA-3′; *als* S1351 (located within exon3)5′-AGATGCGAAGCGACTCTACG-3′; *als* antisense primer AS1650 (located within exon5) 5′- ACTTCATAGTTGCCATCGGC-3′ (also see Table 3, Page 61-62).

0.96 μg of poly(A) mRNA prepared from late embryo was used as starting material for RT-PCR analysis, which was carried out using the QIAGEN OneStep RT-PCR kit. RT-PCR conditions (primer concentrations, input RNA, choice of RT primer, cycling conditions) were initially optimized and were identical for all samples. Appropriate precautions e.g., dedicated areas for sample preparation and analysis, use of RNase-free pipette tips and DEPC-treated distilled water were taken to avoid contamination and RNA degradation. Both control (not shown) and sample were prepared using common master mixes containing the same RT and PCR reagents and were run in parallel.

With QIAGEN OneStep RT-PCR kit, reverse transcription and RT-PCR were carried out sequentially in the same tube. All components required for both reactions were added during setup. The reaction was performed in a total volume of 50 μl containing: 5X RT-PCR buffer 1X,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

400 μM of each dNTP, 0.6 μM of forward and reverse *als*-specific primer, 5 units of RNase Inhibitor, RT-PCR enzyme mix and 0.96 μg of poly(A) mRNA. Thermal cycler conditions were as follows: reverse transcription at 50°C for 30 min, initial PCR action step at 95°C for 15 min, 30 rounds of cycling with: denaturation at 94°C for 1 min, annealing at 62°C for 1 min, and extension at 72°C for 1min; final extension took place at 72°C for 10 min. All RT-PCR products were run on 1.4% gel electrophoresis, visualized and purified with GeneClean Spin Kit (BIO101) to remove salts, proteins, enzymes, oligomers, primers and trace of detergents or organic solvents, followed by DNA Sequencing.

5. Primer extension analysis

To map the 5′ termini of the *als* transcript, poly(A)mRNA was hybridized with an excess of a single-stranded *als*-specific oligodeoxynucleotide primer, which is complementary to the target RNA, radiolabeled at its 5′ terminus prior to hybridization. Reverse transcriptase was then used to extend the primer. The resulting cDNA was complementary to the RNA template and was equal in length to the distance between the 5′-end of the priming oligonucleotide and the 5′ terminus of the RNA. The protocol from Current Protocols in Molecular Biology (Ausubel *et al.*, 1988) and part of the manufacture instruction

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

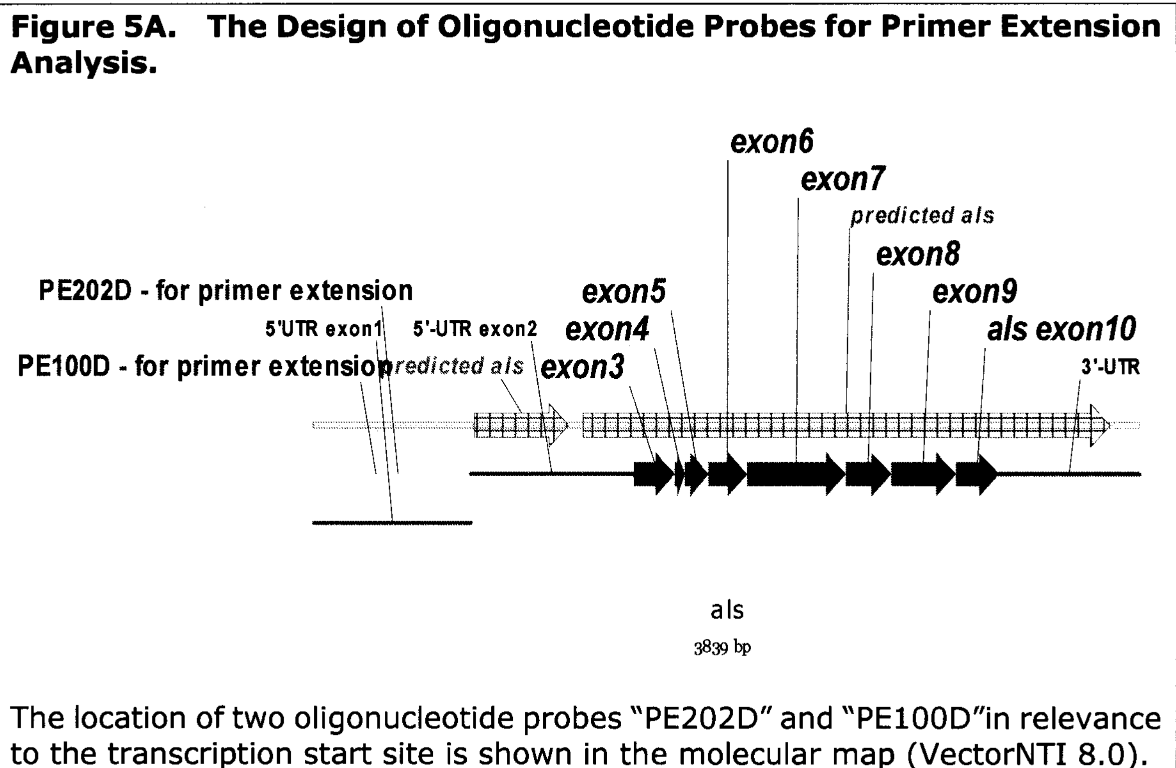
(5'-AmpliFINDER Kit, Clontech lab, Inc.) was modified and adapted.

Two oligonucleotide probes were designed based on the published *als* sequence that are complementary to the *als* coding sequence close to the 5′-end (**Figure 5A**). They are:

(1) "PE100" 5′-GCTAGCGGGCTCATTTGCCTGCGATTCGTGGCA-3′

(2) "PE202" 5′-AATGCGGATACGTAGACGGATGCGTATGCGGAT-3′

Figure 5A. The Design of Oligonucleotide Probes for Primer Extension Analysis.



The location of two oligonucleotide probes "PE202D" and "PE100D"in relevance to the transcription start site is shown in the molecular map (VectorNTI 8.0).

The two primers were synthesized by the Sequencing & Separations Facility at Hunter College (Center for Study of Gene Structure &

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Function). The ATP-labelled oligonucleotide (3μCi/mmol) was mixed with 9μg of poly(A)$^+$ mRNA in 30μl of RNA hybridization buffer (4 part of formamide and 1 part of 5x stock solution; 5x stock solution contains 200mM PIPES, pH 6.4, 2M NaCl and 5mM EDTA). After heating at 85°C for 10 minutes, the mixture was quickly moved to a 37°C water bath and incubated overnight. Following hybridization, the annealed RNA was precipitated in ethanol and redissolved in 10 μl of dH_2O.

The reverse transcription reaction was carried out by adding 20 μl of master mix (1.5xReverse transcriptase buffer, 50units of RNase inhibitor, 1.5mM dNTP mix and 12.5 units of AMV reverse transcriptase) and incubated at 37°C for 2 hours. The reaction was stopped by adding 1μl of 0.5 M EDTA and the RNA template was removed by digesting with RNase (0.33μg/μl) at 30°C for 30 minutes. After phenol/chloroform extraction, the cDNA was precipitated in ethanol and redissolved in 10μl of TE buffer. The cDNA product was analyzed in a 6% polyacrylamide gel/7M urea followed by autoradiography to reveal the length of extended primer.

6. RACE-PCR

To ascertain and define precisely the *als* transcription start site,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

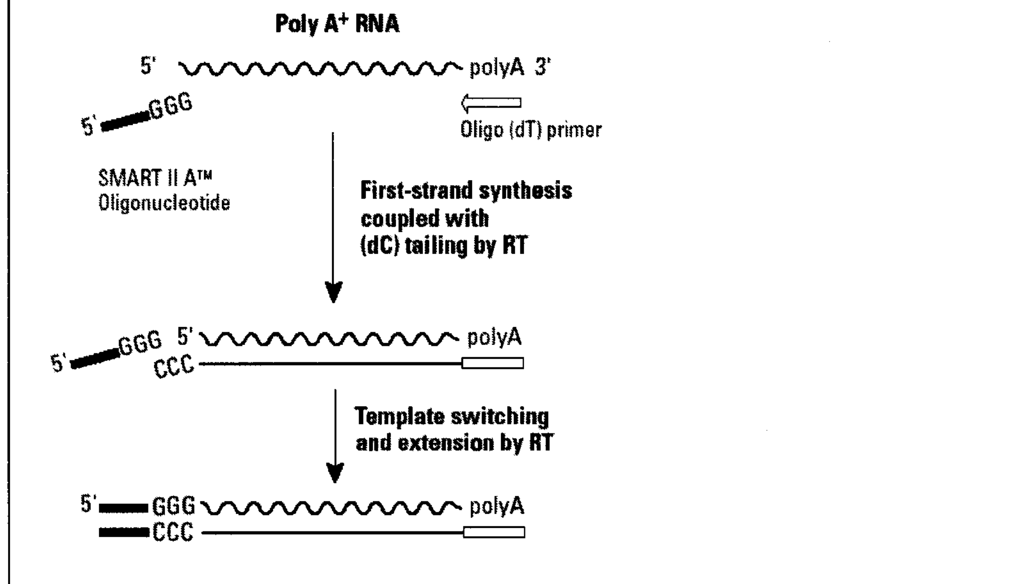
RACE-PCR experiments were designed and conducted.

Nothing in molecular cloning is more frustrating than to isolate a cDNA clone that lacks sequence representative of the 5′ end of the corresponding mRNA. RACE-PCR involves three sequential enzymatic steps: reverse transcription, addition of homopolymeric tails, and PCR, none of which is foolproof. The specificity of amplification depends entirely on the anchoring primer and is generally low. The population of amplified DNAs is therefore enriched for 5′ sequences of the target mRNA but is by no means pure.

The SMART cDNA systhesis technology by CLONTECH was adopted (**Figure 6A**). First-strand cDNA synthesis was primed using a modified oligo(dT) primer. After reverse transcriptase reaches the end of the mRNA template, it adds several dC residues. The SMART II Oligonucleotide anneals to the tail of the cDNA and serves as an extended template for Superscript II RT.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 6A. Mechanism of SMART cDNA synthesis.
The Schematic drawing is adopted from Clontech, BD Biosciences.



(1)Primer design:

The SMART II Oligonucleotide and other anchor primers were synthesized as follows:

SMART II Oligonucleotide (30mer):

5'-AAGCAGTGGTAACAACGCAGAGTACGCGGG-3'

Tm = 70.1°C

5'-RACE cDNA Synthesis Primer (5CDS, 27mer)

$5'-(T)_{25}$ V N - 3'

(V = A, G, or C; N = A, C, G, or T)

Tm = 51.1 °C

3'-RACE cDNA Synthesis Primer (3CDS, 57mer)

5'-AAGCAGTGGTAACAACGCAGAGTAC$T_{(30)}$V N-3'

(V = A,G or C; N = A,C,G or T)

Tm = 72.2 °C

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Long Universal Primer (LUP)

5'- CTAAT ACGAC TCACT ATAGG GCAAG CAGTG GTAAC AACGC AGAGT -3'   Tm = 72.9°C

Short Universal Primer (SUP, 22 mer)

5'-CTAATACGACTCACTATAGGGC-3'

Tm = 45.1 °C

***als* antisense primer A448** for 5'-RACE cDNA synthesis

(GSP1, 29 mer) T_m = 73.8°C

5'-CGGTGCGGAACGGAACGTCTCTAAGCGGA-3'

***als* antisense primer A323** for 5'-RACE cDNA synthesis

(nested GSP1, 29 mer) Tm = 74.8 °C

5'- GTTTGCTGCTTGGATGCGAATGGCGGACG -3'

***als* antisense primer A280** for 5'-RACE cDNA synthesis

(nested GSP1, 28 mer) T_m = 74.0°C

5'-TGCTGGCGGTCCGCTTTGGCTGAACACT-3'

***als* sense primer S2992** for 3'-RACE cDNA synthesis

(GSP2, 26 mer) Tm = 73.5 °C

5'-ACTTGGGTTCGCGGCTGGCTGGTGGA-3'

***als* sense primer S2031** for 3'-RACE cDNA synthesis

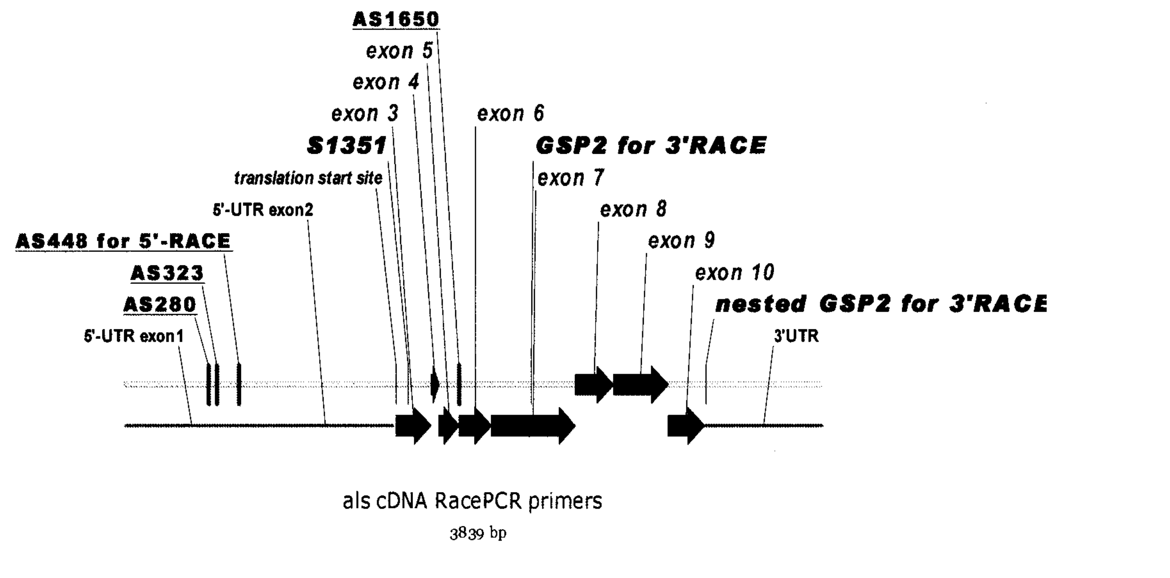
(nested GSP2, 30 mer) Tm = 74.1 °C

5'-CCTGCGTGGGCATCTCGTTCCTGTCCGTTC-3'

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 6B. The Design of primers for RACE-PCR.

Schematic drawing indicating primer locations of 5′- and 3′- RACE-PCR. **AS280, AS323, AS448** and **AS1650** are a series of nested 5′- antisense primers for 5′-RACE in **light bold underlined letters; *S2992, S2031*** and ***S1351*** are nested sense primers for 3′-RACE, in ***bold black italicized letters***.



(2)First-strand cDNA synthesis:

A total of 1.2 μg of polyadenylated RNA from *Drosophila* late embryo was mixed with SMART oligo and 5′ RACE cDNA synthesis primer (10 μM of each) in 5 μl, incubate at 72°C for 2 minutes, cool on ice for 2 minutes; then added 5X first-strand buffer (250 mM Tris-HCl pH 8.3, 375 mM KCl, 30 mM $MgCl_2$), 20mM of DTT and 50mM of dNTP (1μl of each), the above mixture was converted to 5 prime-ready cDNA employing MMLV RNase H^- point mutant reverse transcriptase (Superscript II, GIBCO BRL) in a final volume of 10 μl at 42°C for 1 hour in an air incubator. The resulting

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

full-length, single-stranded cDNA contains the complete 5'end of the mRNA, as well as sequences that are complementary to the SMART oligonucleotide. The smart anchor sequence and the poly A sequence serve as universal priming sites for end-to-end cDNA amplification. The reaction was terminated by placing the tube on ice, then diluted with 250 μl of Tricine-EDTA buffer (10 mM Tricine-KOH pH 8.5, 1.0 mM EDTA), heat at 72°C for 7 minutes, cool down on ice briefly and stored in -75°C.

3'-ready cDNA was prepared at a similar fashion except without adding SMART oligo, plus using 3' RACE cDNA synthesis primer to carry out the reverse transcription reaction.

(3)5'- and 3'- RACE-PCR

5'- or 3'-RACE-PCR reactions were performed in 50 μl of mix containing: High-fidelity PCR Buffer 1X, 200 μM of dNTPs, 1 unit of Platinum Taq DNA Polymerase High Fidelity DNA Polymerase (GIBCO BRL), 200 μM of long and short universal primers as adaptor anchor primers, plus 200 μM of reverse (GSP1, for 5'RACE) or forward (GSP2, for 3'RACE) *als*-specific primer in two separate reactions with 2.5 μl of the 5'- or 3'-RACE-ready cDNA. Thermocycler conditions were as follows: 5 cycles at 94 °C, 5 sec, and 72 °C, 3 min; 5 cycles at 94 °C, 5

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sec, 70 °C, 10 sec and 72 °C, 3 min; 20 cycles at 94 °C, 5 sec, 68 °C, 10 sec, and 72 °C, 3 mins.

(4)Cloning and verification of RACE-PCR products

RACE-PCR products were checked by nested PCR and then cloned into pGEM-T Easy vector (Promega) employing the TA cloning technique, transformed into JM 109 competent cells, >10^8cfu/μg (Promega), amplified and sequenced by the Sequencing Facility at Hunter College.

RESULTS

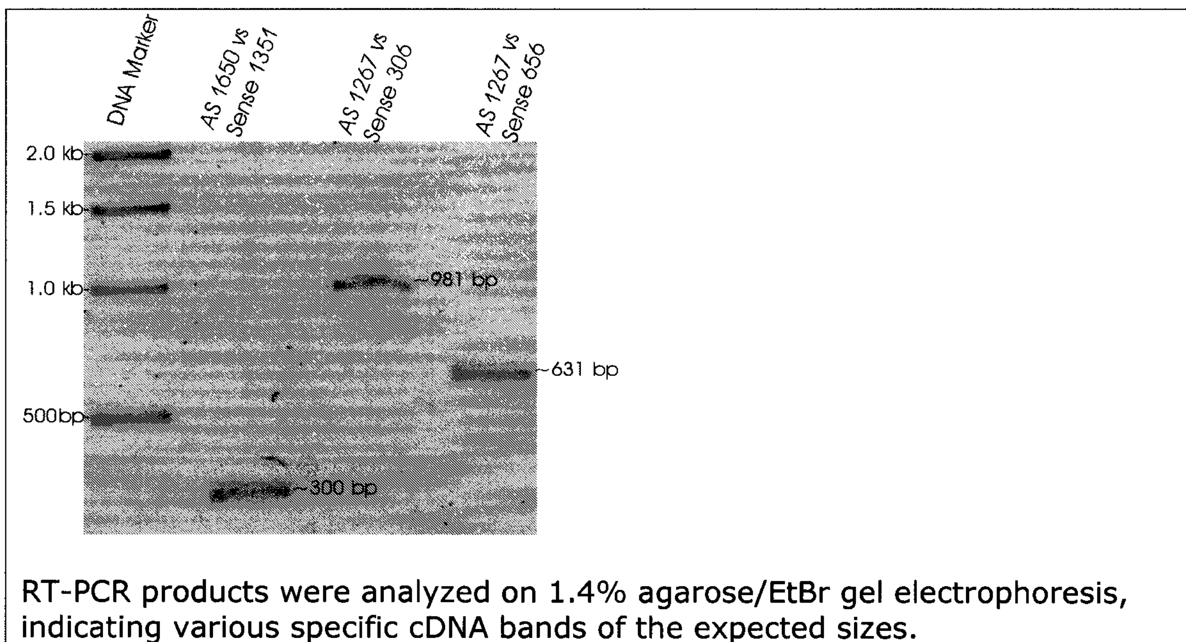
RT-PCR

In order to ascertain the presence of leader exons and the far-apart (29 kb) coding exons, we set up a series of RT-PCR studies using forward oligonucleotides chosen from within the leader exons (S306, S656) that matches the genomic sequence, from coding exon (S1351); and reverse oligonucleotides from exon 3 and exon 5. The RT-PCR products obtained in reproducible experiments with polyA$^+$mRNA extracted from late embryos are shown in **Figure 4B** to illustrate the results yielded by exon-walking. In the case of leader exon 1 and 2, RT-PCR products were at the expected sizes 980 bps and 600 bps. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

primer set in exon 3 and exon 5 generated a product of ~270 bps. In all experiments, RT-PCR products were never detected for negative controls. DNA sequencing data verified the identity of various length of *als* cDNAs.

Figure 4B. RT-PCR Products.



RT-PCR products were analyzed on 1.4% agarose/EtBr gel electrophoresis, indicating various specific cDNA bands of the expected sizes.

These experiments ensure that: (1) the existing exon 5 through exon 1 are present in the *als* transcript; (2) there seems to be no alternative splicing in the 29 kb intron between the first untranslated exon and the third exon where translation start site is located; (3) DNA sequencing data of various resulting RT-PCR products allowed us to assemble the *als* cDNA sequence and further confirmed by comparing with the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Drosophila genomic database for their presence and contiguity. However, the nucleotide sequence discrepancy at the very 5′ end still remains between the *als* mRNA and the genomic sequences in the *Drosophila*. The transcription start site of the *als* gene awaits to be identified.

Primer extension analysis

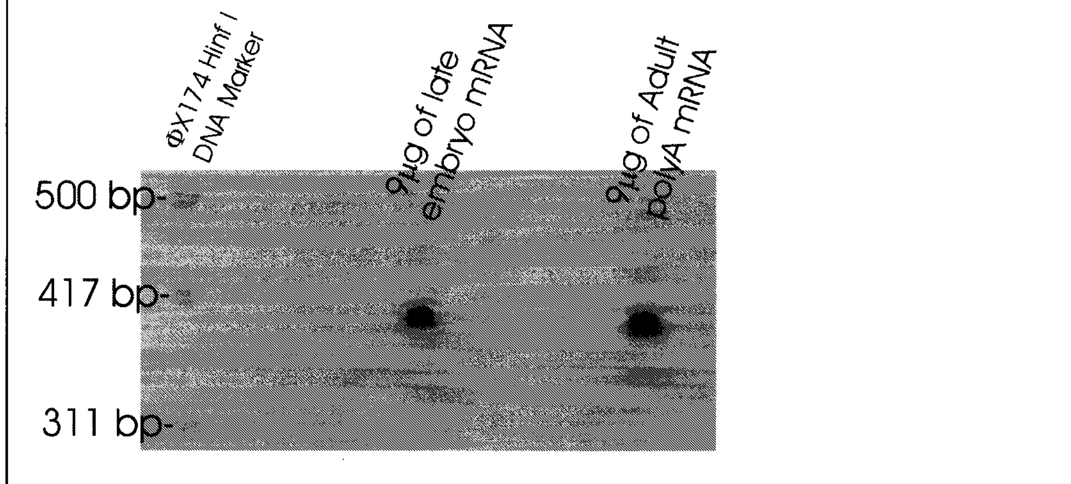
Primer extension was one independent approach utilized to map the transcription start site. In order to determine the size of the untranslated exon in the *als* gene, two different primers were designed based on the Spierer's published *als* cDNA sequence in question, one 100 bp, the other 202 bp near the questionable 5′-end. Both primer sequences corresponded to the genomic sequence in the *Drosophila* database based on blast search. Parallel primer extension experiments were carried out using mRNA from 12-24 hour embryos and from 1-day old adult fly head. The extension reaction yielded a product of around 400 base pairs, which for the first time indicated the distance of the 5′ end of the *als* transcript. Consistent expression was observed in both late embryo and adult fly head with the PE202 primer (**Figure 5B**). This result reasonably matches our precise mapping of 5′ end of *als* with the RACE-PCR approach. It also corresponds to the predicted promoter

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

identified by computer-assisted analysis (see Chapter 3).

Figure 5B. Autoradiography of Primer Extension Products.

Autoradiography showing a primer extension product around 400 bp obtained from *Drosophila* late embryo poly-A$^+$mRNA (two right lanes). Lane 1 shows a ^{32}P-labeled ØX174 HinfI dephosphorylated marker with band sizes as indicated.



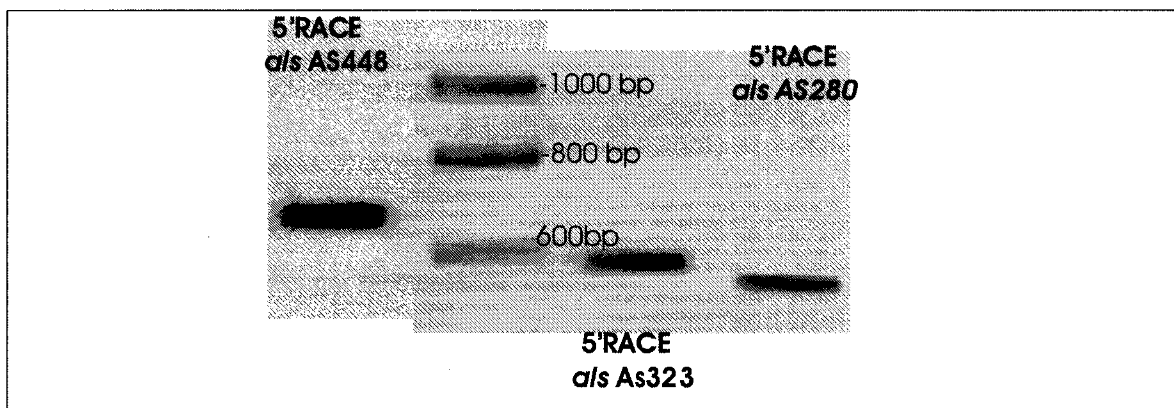
RACE-PCR

Based on sequence identity of the *als* gene exon-walking established by previous RT-PCR results, we have designed *als*-specific antisense primers aimed at extending *als* progressively to the transcription start site (primers involved as shown on page 52, 53 and in **Figure 6B**, page 54). 5'-RACE-PCR generated three fragments of 653-bp, 528-bp, and 485-bp (**Figure 6C**) with repeated efforts using *als* antisense primers A448, A323 and A280. Sequencing data from various 5'-RACE PCR

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

products yielded identical sequence that were 746-bp upstream of the published mRNA by Spierer's lab (GenBank accession No. X07194). This led to the identification of "true" *als* transcription start site, with 5'-noncoding sequence extended further upstream.

Figure 6C. Results of 5'RACE-PCR



By the same token, *als*-specific sense primers (see Table 3 and Figure 6A) were designed for 3'-RACE-PCR experiments, which generated a 1.8-kb fragment, and a 850-bp fragment with nested PCR primer (Figure 6D), the sequencing data indicated an additional 136 nucleotides (with reference to GenBank X07194) in 3'-noncoding region.

Therefore, full-length *als* cDNA was established based on the 5'- and 3'-RACE-PCR experiment. The 5'-end of *als* was further extended by 746 base pairs, which included a new non-coding exon of the *als* gene.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

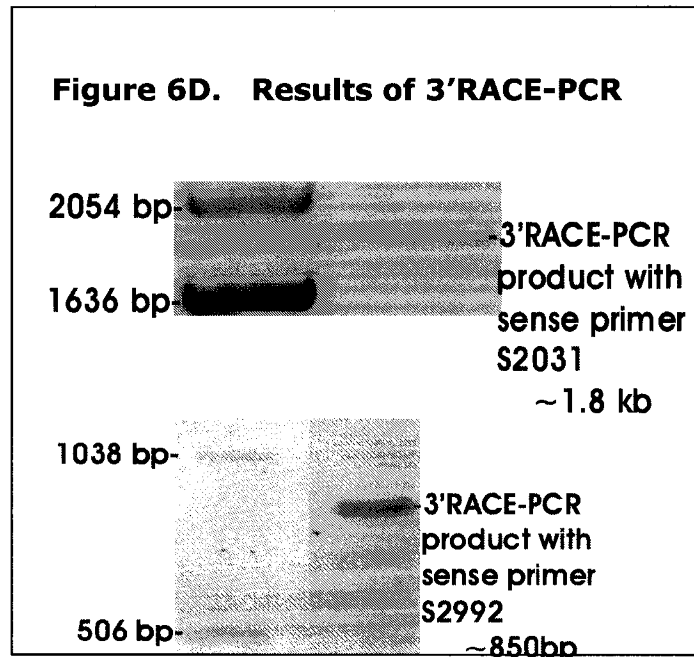


Figure 6D. Results of 3'RACE-PCR

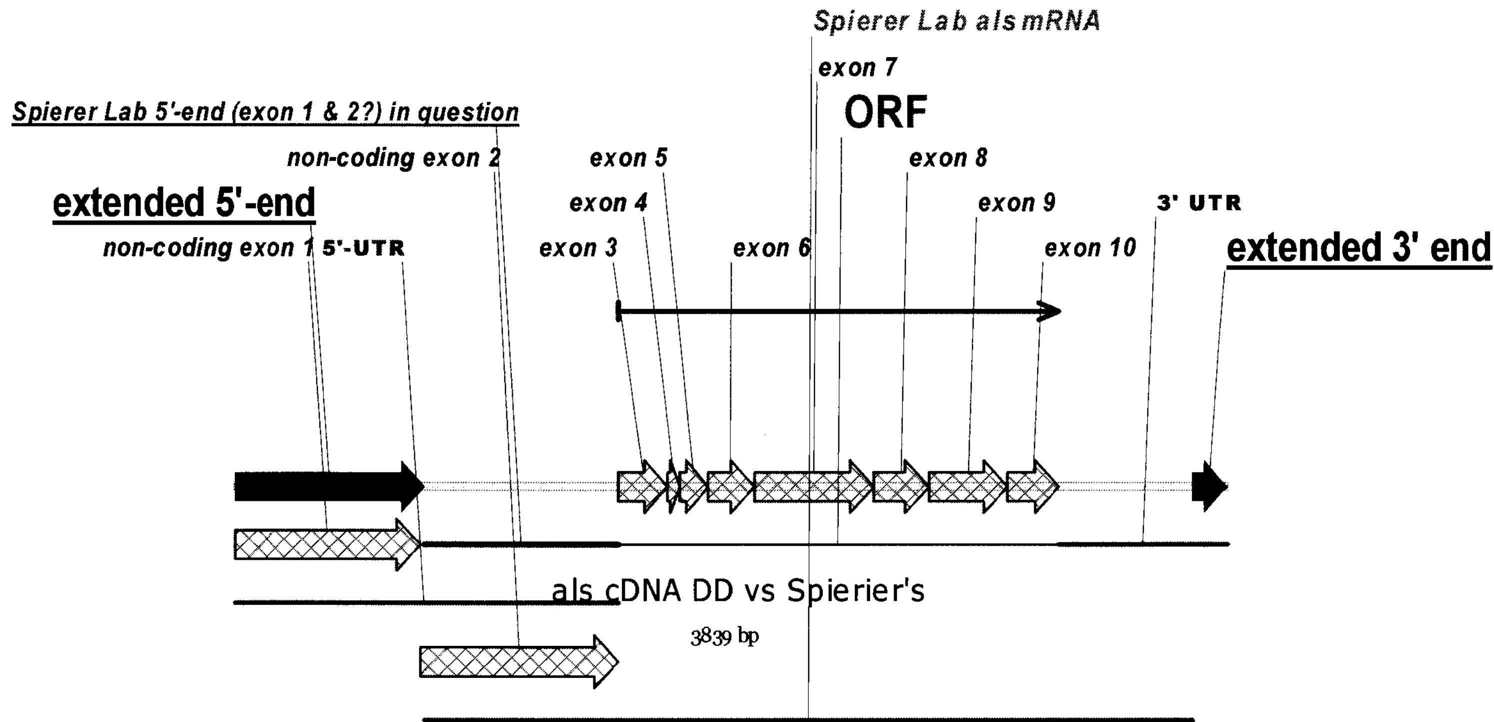
And 3'-end was extended by 136 base pairs, as depicted in **Figure 7.** The *als* primary transcript was shown to have a longer size than previously reported (Bossy *et al.* 1988) at 3,839 base pairs. The 5' leader has 1487 nucleotides, including two non-coding exons 730 bps and 757 bps in length respectively. The 3'-UTR has 651 nucleotides. Reassuringly, the *als* cDNA sequence in both 5'- and 3'-RACE region has a perfect match to the corresponding BDGP genomic region (AC008366 and BACR13R4M).

To complete the structural description of the *als* gene, computer-assisted analysis using an array of updated bioinformatics softwares will be further discussed in the next chapter.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 7. Full-length *als* cDNA. Full-length *als* cDNA was fully established based on the 5'- and 3'-RACE-PCR experiment. The 5 prime of *als* was further extended for 746 base pairs; and to the 3prime end, 136 base pairs (shown in black arrow bars). The first non-coding exon contains 730 nucleotides, and the second 757 nucleotides. All ten exons are represented by open shadowed arrow bars. The *als* primary transcript has a longer size at 3,839 base pairs, as compared to original Spierer's lab published mRNA indicated by the gray line at the bottom of the graph. The 5'-UTR was 1487 bps, and the 3'-UTR was 651 bps. The *als* cDNA sequence in both 5'- and 3'-RACE region matched perfectly to the BDGP genomic sequence (AC008366). The original 5' end in question is in light underlined legend.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3. Primers Designed for Various Applications as Described in the Text

Experiment	Primer name	Primer sequence
Genomic screening	*als*2545-2568	5'- CGCAGAACCGCCGCTGCCCTCCTC -3'
Primer Extension	PE100	5'-GCTAGCGGGCTCATTTGCCTGCGATTCGTGGCA -3'
Primer Extension	PE202	5-AATGCGGATACGTAGACGGATGCGTATGCGGAT-3'
RT-PCR	S1351	5'- AGATGCGAAGCGACTCTACG -3'
RT-PCR	S656	5'- AACATCAGGGACAGGCAGGT -3'
RT-PCR	S306	5'- CGCATCCAAGCAGCAAACTA -3'
RT-PCR	AS1267	5'- CCATCTTAGGCCAGGTCCAA -3'
RT-PCR	AS1650	5'- ACTTCATAGTTGCCATCGGC -3'
1st strand cDNA synthesis for RACE	SMART II Oligo	5'- AAGCAGTGGTAACAACGCAGAGTACGCGGG -3'
als 5' RACE	5'RACE cDNA (5CDS) Synthesis Primer	A A 5'- $(T)_{25}$ G C - 3' C G T
als 3' RACE	3'RACE cDNA (3CDS) Synthesis Primer	5'- AAGCAGTGGTAACAACGCAGAGTACT$_{(30)}$VN -3' (N=A,C,G or T and V = A,G or C)
RACE adaptor anchor primer	Long Universal Primer(LUP)	5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTAACAACGCAGAGT -3'

(Table 3. Continued)

RACE nested adaptor anchor primer	Short Universal Primer(SUP)	5′- CTAATACGACTCACTATAGGGC -3′
als 5' RACE	A448-GSP1	5′-CGGTGCGGAACGGAACGTCTCTAAGCGGA -3′
als 5' RACE	A323	5′-GTTTGCTGCTTGGATGCGAATGGCGGACG -3′
als 5' RACE	A280	5′-TGCTGGCGGTCCGCTTTGGCTGAACACT-3′
als 3' RACE	S2031-GSP2	5′-CCTGCGTGGGCATCTCGTTCCTGTCCGTTC-3′
als 3' RACE	S2992	5′-ACTTGGGTTCGCGGCTGGCTGGTGGA-3′
als upstream seq amplification	Forward "7646"	5′-CAGATAAGTTCAATGATGTCCAGTGC-3′
als upstream seq amplification	Reverse "7873"	5′-TGTCCAGACCAATGCCTCCC-3′
Sequencing Primer of pCaSpeAUGβgal	Forward	5′-CAGATAAGTTCAATGATGTCCAGTGC-3′
Sequencing Primer of pCaSpeAUGβgal	Reverse	5′-TGTCCAGACCAATGCCTCCC-3′

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 3.

BIOINFORMATIC ANALYSIS OF THE *als* GENE AND ITS ENCODING PROTEIN

In this chapter, further exploration of the *als* gene and its encoding protein using various bioinformatics tools is depicted and analyzed. These include genomic organization of the *als* gene, amino acid deduction from the *als* cDNA, peptide alignment, promoter and regulatory sequence analyses, and finally, molecular phylogenetic study of the *als* gene and its large family of siblings and cousins. "Material and Method", "Result" and "Discussion" are presented under two sections:

Part 1. The *als* gene and the ALS protein;

Part 2, Phylogenetic study of nAChR subunit genes.

PART 1. The *als* Gene and the ALS Protein

I. MATERIAL AND METHOD

1. Analysis of Genomic Organization

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The 3,839 bp cDNA sequence of the *als* gene was compared with the genomic sequence of *Drosophila* in GenBank. Both the DotPlot and the Gap methods were used to deduce the intron-exon boundaries. Genebank/Flybase accession numbers of the nAChR subunit sequences used in this study were as follows: *Drosophila als* gene, nAcRalpha-96Aa, DNA: X07194; protein: CAA30172; BACR13M04, AC008366; CG5610-PA, NM_079757. The experimentally defined 3,839 bp of the *als* cDNA has not been deposited into the GenBank.

2. Amino Acid Sequence Analysis and Multiple Peptide Sequence Alignment

The *als*-encoding protein was deduced by the translation software provided by VectorNTI Suite 8.0 (Informax), also in parallel by the "DNA to protein" program available at the public domain of ExPASy Proteomics Server (www.expasy.ch).

The deduced ALS protein was aligned using AlignX program, an integrated component of VectorNTI Suite 8.0 (Informax).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. Computer-assisted analysis of promoter regions and motif-scanning of transcription factor binding sites

A 32 kb genomic region spanning the *als* 5'flanking sequence, the two leader exons, and the third exon that carries the translation start site was submitted to software McPromoter MM:II – The Markov Chain Promoter Predication Server (http://genes.mit.edu/McPromoter.html/) and the Neural Network Promoter Prediction for promoter prediction (www.fruitfly.org/seq_tools/promoter.html/).

The 2.9 kb (**Figure 14**, page 110) and 1.25 kb (**Figure 11A**, page 79) fragments containing 482 bp of 5'-untranslated exon were scanned for potential consensus transcription factor binding sites using three different programs:

(1) TESS: Transcription Element Search System (www.cbil.upenn.edu/tess/);

(2) TRANSFAC 6.0 (www.gene-regulation.com/pub/database.html#transfac/);

(3) MatInspector version 2.1. (http://www.gsf.de/biodv/matinspector.html/).

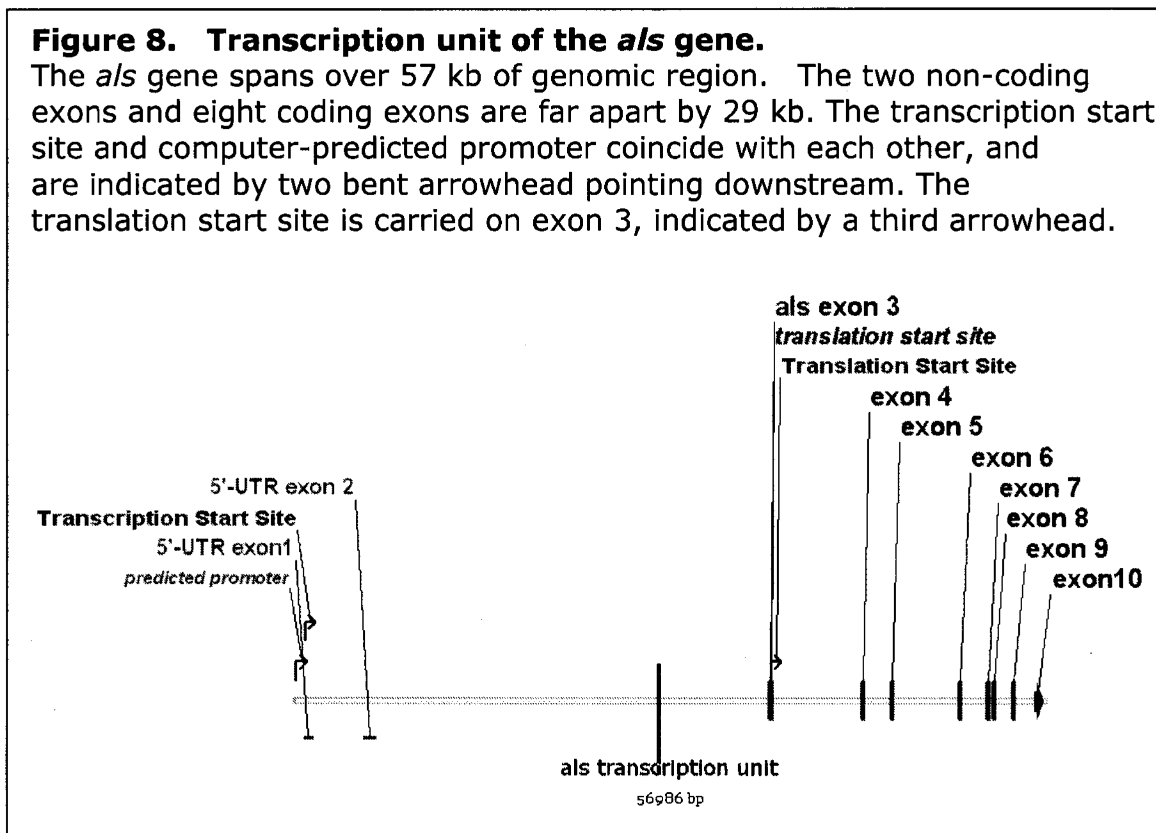
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

RESULT

1. Genomic Organization of the *als* gene

The transcription unit of the *als* gene spread over ~57 kb of genomic region, which contains ten interrupted exons. The first two 5′-non-coding exons comprise the unusual long leader sequence of 1,487 base pairs and are separated by an intron of 29 kb from the eight exons that encode the ALS subunit. These eight coding exons map within a 21 kb chromosomal interval (**Figure 8**).

Figure 8. Transcription unit of the *als* gene.
The *als* gene spans over 57 kb of genomic region. The two non-coding exons and eight coding exons are far apart by 29 kb. The transcription start site and computer-predicted promoter coincide with each other, and are indicated by two bent arrowhead pointing downstream. The translation start site is carried on exon 3, indicated by a third arrowhead.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The exon-intron boundaries of the *als* gene is shown in **Table 4**. Four out of seven introns in the ALS coding sequence have splice site localizations that are perfectly conserved throughout evolution since they occur at precisely the same nucleotide in vertebrates (Bossy *et al.* 1988). The three remaining introns are located in two poorly conserved coding regions which vary in length depending upon the subunit type. One intron which splits the signal peptide coding sequence into two exons in all known vertebrate nAChR genes is absent from the *als* gene (Bossy *et al.* 1988). No introns separated the first three ALS transmembrane coding sequences, a situation that also occurs in vertebrate neuronal nAChR genes (Nef *et al.* 1988), whereas in endplate nAChR genes introns do separate the coding sequences for transmembrane regions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 4. Exon/Intron Boundaries of *Drosophila als* Gene.

Capital bold letters and lower cases indicate exon and intron sequences, respectively. The length of intron is indicated above.

Upstream exon	Intron length / Boundary sequence	Downstream exon
	3,801 bp	
(exon 1)	**GCCCTGTAAT**gtaagtcaga... ...ttttgcag**CGACAACAAC**	(exon 2)
	29,900 bp	
(exon 2)	**TGCAGCAACG**gtaagtgtag......ctttacag**GGTTGGACCT**	(exon 3)
	6,803 bp	
(exon 3)	**GATCGATGTG**gtaagtgcgg......acttgcag**AATCTAAAGA**	(exon 4)
	2,225 bp	
(exon 4)	**GTGGAACAG**gtaagttatt... ...tgcaattt**CAGGAATGGAA**	(exon 5)
	5,009 bp	
(exon 5)	**TCTATAACAA**gtaagtacc... ...atttcagc**GCCGATGGCAA**	(exon 6)
	1,888 bp	
(exon 6)	**TGGTTACATG**gtgagttgg... ...cttttccag**GTGGACTTGA**	(exon 7)
	75 bp	
(exon 7)	**TGTGAATTTT**aggtaagat......cttcctcttt**AGATCCCCTGT**	(exon 8)
	1,170 bp	
(exon 8)	**ACGGCATTCC**aggtacga......ttttcccca**GCACTACCCGC**	(exon 9)
	1,266 bp	
(exon 9)	**TTTGAGAGT**gtgagtaat......acctttcag**GTGGAAGAGGA**	(exon 10)

2. Amino Acid Sequence Analysis

The full-length 3,839-bp *als* cDNA defines a single long open reading frame (ORF) of 1,701 nucleotides. The main ORF encodes 567 amino acids presenting structural domains homologous to those of the vertebrate neuronal nAChR subunits. The entire nucleotide sequence and the deduced amino acid sequence is shown in **Figure 9**. There are two conserved Cys at positions 128 and 142, which identify all the α subunits.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 9. Nucleotide sequences and putative amino acid sequences of the *als* gene.

Nucleotides are numbered in the 5'- to 3'-direction. The putative amino acid sequences are shown above the nucleotide sequences in three letter code. A single major open reading frame of 1701 bp is present encoding a protein of 567amino acids. Four transmembrane (TM) domains are shaded in gray. In-frame stop codons upstream to the start codon are underlined. Several small ORFs in the 5' UTRs are underlined by wave lines.

```
          5'-UTR-exon 1
   1      AGTTGTCTCA GAGCCCGCGG GTTGAAGAAA TCGTACGCAT CGAGCGAGTG ACAAGCGCCG CGTCTCACCA
  71      AACAAAGAAA AGCAAAAGCA GAGCCCCAAA AAAAGGAAAA AGAAAAAACA AAAATCCAAA TAAAAAACAA
 141      CAAAAATCCA AGCTAAAACA ACAGAAAAAC GTAGAAAACC CGCTACACTG ACACAAAAAC ATAGCCAGTC
 211      TAACGAATAT TTTGTAGAAA CCTCCCACAA CAGAAAAAAA AAAAACATTC GCAACAAAAT GCTGCCACGA
 281      ATCGCAGGCA AATGAGCCCG CTAGCCCCTA ACGGCTCCCA TCTGCGGCTG CTGCACAGCA GCGAGCCCGA
 351      GCCCGGTCCC GCATCCGCAC TCGCATCCGC ATACGCATCC GTCTACGTAT CCGCATTCTC ATCCACATCC
 421      GCATCAGCCA CTCGCCGCGT CTCCCAGTGT TTGTGCCAGT GTTCAGCCAA AGCGGACCGC CAGCAATTGC
 491      AGTGTAAGGC GTCCGCCATT CGCATCCAAG CAGCAAACTA ATCTAAACGG CGGCGAGTGT CGCATCTGAA
 561      ATTGAGTTTC GCTTCCGGTT CCGCCGCGCG TCGCCCAATC CGTTTGGCCA AGTCGGTCGG TTGGTCCGCT
 631      TAGAGACGTT CCGTTCCGCA CCGGTCGGTC GGTCGTTCGT TCGTCCGTTC AGACAGTTTT TCCTGTTCAA
                                                                                5-UTR exon2
 701      AAATTGATGA CAACAGCGCA GCCCTGTAAT CGACAACAAC ATTCAAGACT GCTAACTGGA ATTGCTTCGA
 771      ACCATCACCA TTAAAAAACG AAGAAACAAC CACAACAAAG TGCTGCCAGC AACAAAAACG TAATCAGGAA
 841      TTTAAGCAAA ACCAAAAAGA AACATCAGGG ACAGGCAGGT AGACCCAGGG AGCATGTTTA TGGAATTGTC
 911      ATCGTTAGCG AGCAACTGAC GCAACAGCTC AAGGCATGTG GCAACAGCAG CAGCAACAGC AGCAGCAGCA
 981      ACACTGAGCG GCAGGAGCAG AAACACCTGG CAAGAGAAGC AACTAGGAGC AGGACAACCT CCGCGGACAG
1051      GTGGTGTGCT GCTATAAAGC AGCACCATAG CTGCCATTAA AATAGACGCC GCGGCGCAAC AGCAACGCTG
1121      CCGTCGTCAT AATCAACGCC AGCGACAACA CCACCCGCAT CATCGCCATC GCAGACGACG ACGGCGCAGC
1191      AGCAACATCA GCAGCAACAT CAGCAGCAAC AGCAACTGCA ACATCAGCAA CTGCAGCAGC AGCGGGCTGT
1261      TGGGCGCAGT GGGCGTGGCC GGCAGCTGCC CATTGCTCGC CACACAGCAA CTAATCGAGG AGCCTGGCAG
1331      CAGCAACAAC ACAGGCGACG ACGACGGCGG CGCAACAACG CAGCAGCAAC ACTGCAGCCT GCAACATCAG
1401      CAATCCGGTC GACGGCGACA CTACGACGGC CTGTCATCTG TAGGCAATGC AACGCGCTGT GCAGCAACGG
                          SP → Met GlySerValLeu PheAlaAla ValPheIle AlaLeuHisPhe AlaThrGly·
          ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1471      GTTGGACCTG GCCTAAGATG GGTAGCGTGC TATTCGCAGC TGTATTCATA GCATTACACT TTGCCACCGG
          Exon 3
```


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

```
      ·GlyLeuAla AsnProAspAla LysArgLeu TyrAspAsp LeuLeuSerAsn TyrAsnArg LeuIleArg
1541  CGGCCTGGCC AACCCAGATG CGAAGCGACT CTACGACGAC CTGCTGAGCA ACTACAATCG CCTCATCCGA

      ProValGlyAsn AsnSerAsp ArgLeuThr ValLysMetGly LeuArgLeu SerGlnLeu IleAspValAsn·
1611  CCCGTGGGCA ACAACTCGGA CCGTCTCACC GTCAAGATGG GTCTGCGCCT CTCCCAGCTG ATCGATGTGA

      ·ALeuLysAsn GlnIleMet ThrThrAsnVal TrpValGlu GlnGluTrp AsnAspTyrLys LeuLysTrp·
1681  ATCTAAAGAA TCAAATTATG ACAACCAATG TCTGGGTGGA ACAGGAATGG AACGACTATA AATTGAAATG
      Exon 4                                                                  Exon 5

      ·AsnProAsp AspTyrGlyGly ValAspThr LeuHisVal ProSerGluHis IleTrpHis ProAspIle
1751  GAATCCGGAT GACTATGGCG GCGTGGACAC TTTGCACGTT CCCTCCGAGC ATATATGGCA TCCGGATATT

      ValLeuTyrAsn AsnAlaAsp GlyAsnTyr GluValThrIle MetThrLys AlaIleLeu HisHisThrGly·
1821  GTGCTCTATA ACAACGCCGA TGGCAACTAT GAAGTGACAA TAATGACAAA AGCAATTCTT CATCACACGG
                                                                              Exon 6

      ·GLysValVal TrpLysPro ProAlaIleTyr LysSerPhe CysGluIle AspValGluTyr PheProPhe·
1891  GCAAAGTGGT GTGGAAACCG CCCGCCATTT ACAAATCCTT CTGCGAAATT GATGTCGAGT ACTTTCCCTT

      ·AspGluGln ThrCysPheMet LysPheGly SerTrpThr TyrAspGlyTyr MetValAsp LeuArgHis
1961  CGACGAGCAG ACATGTTTCA TGAAGTTCGG ATCCTGGACC TACGATGGTT ATATGGTGGA CTTGAGGCAC
                                                                              Exon 7

      LeuLysGlnThr AlaAspSer AspAsnIle GluValGlyIle AspLeuGln AspTyrTyr IleSerValGlu·
2031  TTGAAGCAGA CCGCCGATTC GGACAACATC GAGGTGGGCA TCGACCTGCA GGACTACTAC ATCTCCGTGG

      ·GTrpAspIle MetArgVal ProAlaValArg AsnGluLys PheTyrSer CysCysGluGlu ProTyrLeu·
2101  AGTGGGACAT CATGCGTGTG CCGGCGGTGC GGAACGAGAA GTTCTACAGC TGCTGCGAGG AGCCGTATCT

                                                                               ← TM1
      ·AspIleVal PheAsnLeuThr LeuArgArg LysThrLeu PheTyrThrVal AsnLeuIle IleProCys
2171  GGACATAGTG TTCAACCTGA CGCTCCGCCG GAAGACCCTC TTCTACACGG TCAACCTGAT CATACCCTGC

                                                                               ← TM2·
      ValGlyIleSer PheLeuSer ValLeuVal PheTyrLeuPro SerAspSer GlyGluLys IleSerLeuCys
2241  GTGGGCATCT CGTTCCTGTC CGTTCTGGTC TTCTACCTGC CCAGCGACTC TGGCGAGAAG ATCTCGCTCT

      ·CIleSerIle LeuLeuSer LeuThrValPhe PheLeuLeu LeuAlaGlu IleIleProPro ThrSerLeu·
2311  GCATCAGCAT CCTGCTCTCG CTTACCGTGT TCTTCCTCCT GCTGGCCGAG ATCATTCCGC CCACGTCGCT

                                                                               ← TM3
      ·ThrValPro LeuLeuGlyLys TyrLeuLeu PheThrMet MetLeuValThr LeuSerVal ValValThr
2381  GACGGTGCCG CTGCTGGGAA AGTATCTGCT CTTCACCATG ATGCTGGTCA CGCTCTCTGT CGTGGTCACC

      IleAlaValLeu AsnValAsn PheArgSer ProValThrHis ArgMetAla ProTrpValGlnArgLeuPhe·
2451  ATTGCCGTGC TCAATGTGAA TTTTAGATCC CCTGTCACGC ATCGCATGGC ACCGTGGGTG CAGCGCCTCT
                                                                              Exon 8
```


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

```
        ·PIleGlnIle LeuProLys LeuLeuCysIle GluArgPro LysLysGlu GluProGluGlu AspGlnPro·
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2521    TCATCCAGAT CCTGCCCAAG CTGCTCTGCA TCGAGCGGCC CAAGAAGGAG GAGCCCGAGG AGGACCAGCC

        ·ProGluVal LeuThrAspVal TyrHisLeu ProProAsp ValAspLysPhe ValAsnTyr AspSerLys
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2591    GCCCGAAGTG CTCACCGATG TCTATCACCT GCCGCCGGAT GTGGACAAGT TTGTCAACTA CGATTCGAAG

        ArgPheSerGly AspTyrGly IleProAla LeuProAlaSer HisArgPhe AspLeuAla AlaAlaGlyGly·
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2661    CGTTTCAGCG GCGACTACGG CATTCCAGCA CTACCCGCTT CGCATCGCTT CGATTTGGCT GCGGCGGGCG
                                                                              Exon 9

        ·GIleSerAla HisCysPhe AlaGluProPro LeuProSer SerLeuPro LeuProGlyAla AspAspAsp·
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2731    GGATTAGCGC CCACTGTTTC GCAGAACCGC CGCTGCCCTC CTCACTGCCA CTCCCGGGCG CCGACGACGA

        ·LeuPheSer ProSerGlyLeu AsnGlyAsp IleSerPro GlyCysCysPro AlaAlaAla AlaAlaAla
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2801    TCTGTTCAGC CCGTCGGGCC TCAACGGAGA CATCAGTCCC GGCTGCTGCC CGGCCGCTGC AGCCGCCGCT

                                                                               ← AM
       AlaAlaAspLeu SerProThr PheGluLys ProTyrAlaArg GluMetGluLysThrIle GluGlySerArg
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2871    GCCGCCGATC TCAGTCCTAC GTTCGAGAAG CCCTACGCCC GCGAAATGGA GAAGACCATC GAAGGATCCC

        ·PheIleAla GlnHisVal LysAsnLysAsp LysPheGlu SerValGlu GluAspTrpLys TyrValAla·
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2941    GCTTCATCGC GCAGCACGTG AAGAACAAGG ATAAGTTTGA GAGTGTGGAA GAGGACTGGA AGTACGTTGC
                                                                              Exon10

                                                                               ← TM4
        ·MetValLeu AspArgMetPhe LeuTrpIle PheAlaIle AlaCysValVal GlyThrAla LeuIleIle
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
3011    CATGGTATTG GATCGTATGT TTCTGTGGAT CTTCGCGATC GCTTGCGTGG TCGGCACAGC GCTGATTATC

        LeuGlnAlaPro SerLeuHis AspGlnSer GlnProIleAsp IleLeuTyr SerLysIle AlaLysLysLys·
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
3081    CTGCAGGCTC CCAGCTTGCA CGACCAATCG CAGCCGATCG ATATACTCTA CTCGAAGATT GCCAAGAAGA

        ·LPheGluLeu LeuLysMet GlySerGluAsn ThrLeu
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
3151    AGTTCGAGCT GCTCAAGATG GGCAGCGAGA ACACCTTATA GCGACCACTT GGGTTCGCGG CTGGCTGGTG

3221    GAATCTAATC ATCAGCTCGC TGCGGACCCT GTTCGCCGCT CGCGACGATG ATCGTGGATA ATATACGATG

3291    AGCCAAACCA TGAGTATTTC CTCGGCGTTG TACAACACTA ACAGCAGCAG CAGCAGCATC AGTATGAGCA

3361    ACGCCAGTGA CAGTGGCAGT GTCATCGGCA GTGCTGTAAA GGTTATTGTG AGTAGTGTTG GCAGCACCAA

3431    CGAAACGGAT GACTCGTGGA GCCCATCGTA AATGTGCGCG AGGCGAGTGA AGTGAACATT GCGAAATGAA

3501    TTGAGATACA ATCGAATGGA GGCAAGTTTT TGGAGCAGTT GATCAAAATA TGTATGTAAT GTAATTTGAT

3571    GGGGTTGTAG TGTTGATTAA CCGTGACAGT GCTGGACAGC GACGCAAGCC ACAAGATCGC TGCGGAGCGG

3641    AGATTAGTTT GTAAGTTTTT TGATGTACTT GATGTATGAT TGTGTGGTCG AAGATGTTTA GGGTGGTTCC

3711    CGAACCAGTG TTGTAAAAAG CGTATGGAGT TTTGTAAGTT ATATTGTATT TTTGAGCGCC AAAAAGTTCC

3781    TTGTTACATC TTTTAGTTAT AAAGTTCCAC ACAACACAAA GCAACAAAGC TAACAAAAA
```


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As observed previously (Bossy *et al*. 1988), *als* was found to share 42% overall sequence identity to ARD subunit. An updated alignment of the deduced amino acid sequences of all *Drosophila* nAChR subunits was performed using AlignX provided by VectorNTI Suite 8.0 (Informax, Inc.). The result is shown in **Figure 10**, along with pairwise subunit identity score indicated in **Table 5**. Several highly conserved regions in the transmembrane segments M1-M4 were revealed by the alignment. Other regions such as the N-terminus and cytoplasmic domains are generally more variable. Three of the *Drosophila* nAChRs are most closely related to mammalian $\alpha7$ subunits, the predominantly expressed bungarotoxin-sensitive α subunits in mammalian brain. In contrast, the seven remaining fly nAChRs are more closely related to each other than to any known mammalian α or β subunits. The encoded protein of 567 amino acids displays high sequence identity with Dα2 (76%), SBD (75%), Dα4 (74%) and *ard* (68%). Identity scores between the neuronal subunits demonstrate that ALS, DALPHA2, DALPHA4 and SBD are significantly more homologous to one another than to DALPHA3, DALPHA5, DALPHA6 AND DBETA3, this agrees with the result from our phylogenetic studies.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 10. Protein Sequence Pile-Up of all known *Drosophila* nAChR subunits. Comparison of the deduced amino acid sequences of all currently known *Drosophila* nAChRs.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

```
(1050)          1050      1060      1070      1080      1090      1100      1110      1120      1130      1140      1150      1160      1170      1180           1195
ALS       (413) SSLPLPGRDDDLFSP------SGLNGDISPGCCPAAAAAAAAADLSPTFEKPYAREMEKIIEGSRFIAQHVRNKDKFESVEEDWKYVAMVLDRMFLWIFAIACVVGTALTILQAPSLYDQSQPIDILYSKIAKKKFELLKMGSENTL
DALPHA2   (460) LSGYNGLPSVLSGLD--------------------------DSLSDVAARKKYPFELEKAIHNVMFIQHHMQRQDEFNAEDQDWGFVAMVMDRLFLWLFMIASLVGTFVILGEAPSLYDDTKAIDVQLSDVAKCIYNLTEKKN----
ARD       (419) SGGELGLSDGCR-------------------------R----ESESSDSILLSPEASKATEAVEFIAEHLRNEDLYIQTREDWKYVAMVIDRLQLYIFFIVTTAGTVGILMDAPHIFEYVDQDRIIEIYRGK-------------
DALPHA5   (712) QGDDGSIGPIGSTR------------------------MPDAVTHHTCIKSSTEYELGLIIKEIRFITTQLRKDDECNDIANDWKFAAMVVDRLCLIIFTMFAIIATIAVLLSAPHIIVS--------------------------
DALPHA7   (441) QGDDGSVGPVGPAGP----------------V--VDGRLHEAIGHTCLTSSAEYELALIIKELRWITDQLKKEDETSDITRDWKFAAMVVDRLCLIIFTLFTIIATLAVLFSAPHFIFP--------------------------
DALPHA6   (429) -GSQTAIGSSASFG------------------------RPTTVEEHHTAIGCNHKDLHLIIKELQFITAKMRKADEAEEIGDWKFAAMVVDRFCLIVFTLFTILATVTVLLSAPHIIVQ--------------------------
DBETA3    (371) ---ES--------------------------------------QAEKTKEMDEHPYEEADEQESSPLSINHTEVPGAKANQFDWALLATAVDRISEVSFSLAFLILAIRCSV-------------------------------------
DALPHA3   (629) AAAAAAAAAAAAGTPPPPPPPAGVCDILPACQREGGPHCCSGRGNVRQRWHTCPELHKANDGVTYIADQIRKEEESTRVKEDWKYVAMVLDRLFLWIFTLAVVVGTAGIILQAPTLYDTRMPIDIKLSEIASTTAKPNVARPVL--
DALPHA4   (428) -------------------------------------KCLP---------EIHKSCFCVRFIAEHTKMQEDSTKVKEDWKYVAMVLDRLFLWIFTLAVVVGTAGIILQAPTLYDDRIPIIEQKDLDFSGTSAGRCRPPC---
SBD       (406) NGMQN-------------------------------SVDSDNVIPRNLTPEVLQALRAVRFIACHIKKDADKDNEIVEDWKFVSMVLDRFFLWLFTLSCVFGTLAIICQSPSLYDTRSPIDRQLSEIPLRKNNFMLPPDIVRQ
Consensus(1050)  G     IG                               S         EL KAL  VRFIAEHLKK DE   I EDWKFVAMVLDRLFLWIFTLA VVGTIAILL APSLYD    I
```

Table 5. Amino Acid Sequence Identity Score of Pileuped nAChR subunits.

Level of amino acid sequence identity (%) among ten *Drosophila* nAChR subunits. Alignments were calculated using Clustal W as an integrated component of the VectorNTI Suite 8 program.

	ALS	DALPHA2	ARD	DALPHA5	DALPHA7	DALPHA6	DBETA3	DALPHA3	DALPHA4	SBD
ALS	100	**76**	**68**	40	63	57	56	58	**74**	**75**
DALPHA2		100	**68**	41	63	57	56	53	70	**73**
ARD			100	44	66	61	60	50	68	**71**
DALPHA5				100	65	50	37	23	40	42
DALPHA7					100	**72**	57	46	63	65
DALPHA6						100	56	39	57	60
DBETA3							100	36	57	61
DALPHA3								100	62	55
DALPHA4									100	**74**
SBD										100

3. Computer-assisted analysis of promoter regions

The *als* gene 1.2 kb region including the first leader exon was submitted to software McPromoter MM:II – The Markov Chain Promoter Predication Server (http://genes.mit.edu/McPromoter.html) and the Neural Network Promoter Prediction (www.fruitfly.org/seq_tools/promoter.html). Both programs predict the same promoter (from -43

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to +7) that coincides with our transcription start site (+1).

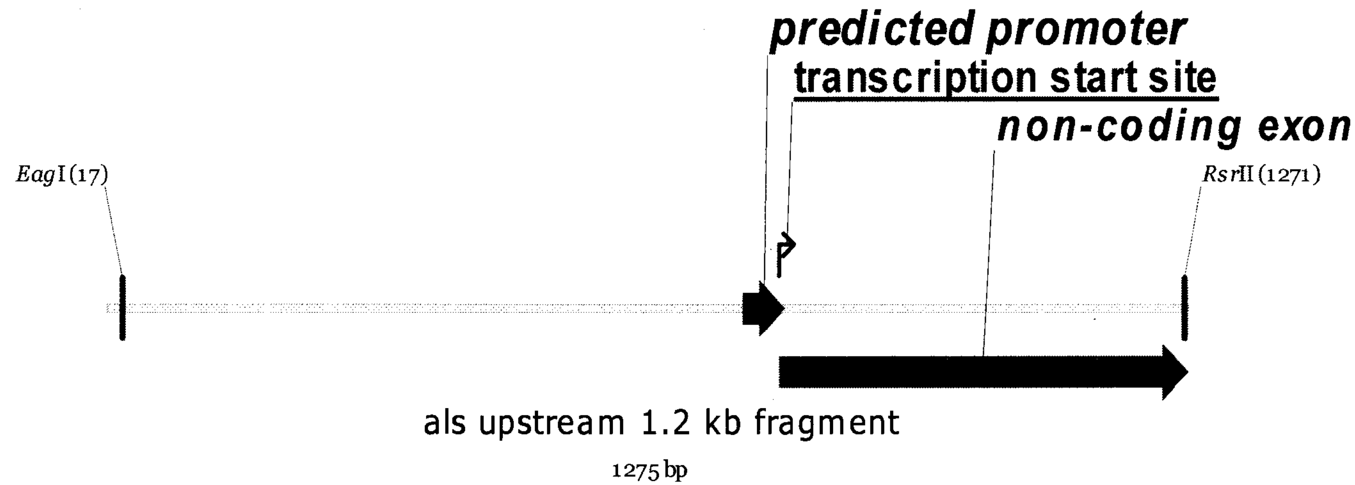
Motifs of transcription factor binding sites were searched for in the 2.9 kb and 1.25 kb fragments containing a common region of 482 bp from the 5′-untranslated exon. Analysis of the upstream sequences revealed the absence of TATA and CCAAT boxes, although one TATA-box was found at -357 to -353, three CCAAT-boxes were observed from -460 to -457; -259 to -255; -372 to -369 (on the reverse strand), they were too far away from the vicinity of the transcription start site to be associated with transcriptional control of the *als* gene.

Potential binding motifs for several known regulatory sequences were observed as illustrated in **Figure 11B (1.25 kb upstream fragment)**. There were GC-like boxes (GGGCGG) at -53 to -44 and -13 to -7 located immediately upstream from the first exon. Two E-boxes were found from -40 to -35, and from +322 to +327. Five Zeste binding sites were revealed at -586 to -581, -224 to -219, -68 to -63, +45 to +50, +371 to +376 and two potential GAGA binding sites at -216 to -210, -124 to -118 were also discovered. All of the above sites form around the cluster within 793 bps upstream of the *als* transcription start site.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 11A. Molecular Map of the 1.25 kb Fragment.
The 1.25 kb fragment contains 482 bp of the first non-coding exon and 793 bp nucleotide upstream of the *als* transcription start site. The partial first exon is indicated by the solid arrow bar; the predicted promoter is shown in a short arrowhead immediately upstream of the transcription start site.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 11B. Analysis of 1.25 kb Upstream Fragment of the *als* Gene.
DNA sequence of the 1.25 kb *als* upstream fragment surrounding the *als* transcription start site. Exon sequences are in ***bold italic letters*** with the transcription start site designated as "+1", the initial nucleotide "A" is shown taller. The predicted *als* promoter is underlined and shaded gray. Proximal transcription factor binding elements are indicated. Four GAGA motifs are shown in filled box, with double arrow line on top; Zeste binding sites are in open boxes; Two-E boxes are enclosed by oval frame; two GC-like boxes are underlined by one-way arrowhead line.

-793 CGCTGGAGAG TTTTTCGGCC GGAGAAAAAT CCACGCCCTA TATGTAAATT ATTAGCCTGA CTGCTTTTTG TTCGCCCAAC

-713 TCGCGGCGAC ACAAGTCCCC TGGAGTTTCT CGCTAACAAT TTGCGGGCGA CTGAGAAGTG GTGGCCGAAA AAAAAACCCA

-633 TGAGAGACAC CAAATAAAAA AGAAAGAGAC ACTTTCAAGC GGAAGAGTGA GTGAGTTGGG CGAAGGATGC TCTCGGTTTG

-553 TAGACTTTTG CCATGTGGCT CAAAGTAAAA CGCCCTTTTG GAGTTGCAGA TAAGAGAGTG AAAGGCTAAG AGAACCATTT

-473 ATTTTTATTT TGTCAATAGA ACAAAATTAC AAAACTAATC TAATAAAATA AGGGGAATAC AATTAACTAA AATCTTGACA

-393 CATCTTTTCG TACGAAAATC TATTGCCTAC CTACAATATA ATCTATCAAC GCACCGCTGA AATTTCTTAA CTAATTTCAA

-313 GCTAGCGCAA GACATTACCA AACACGAAGG CTGTGACCGC TTGAGAACCA AAGCCCAATG CTTACTTACC ATCATGCCAA

-233 AGGACCTGCC ACTCGAAGAG AGCGAGAAAG GTCCTGTATC TCAGAGATTC AAAGCAAAAA GGGTGGCTGC CCGCTCTTCA

-153 AATTTTGTCG CTTTGTGGCC CAGCATGCGC AGAGCGGCGG AGAGAACGTT GGCGAGCGCC ACACGGCCAC ATCCCGAGCA

-73 GCATCCACTC CAAAGGACCA GCGGCGGCGG CCCCAGCTGC GAAAAAGTCT GCCGTGCAGC GGCGGCCATC AGT***AGTTGTC***

8 ***TCAGAGCCCG CGGGTTGAAG AAATCGTACG CATCGAGCGA GTGACAAGCG CCGCGTCTCA CCAAACAAAG AAAAGCAAAA***

88 ***GCAGAGCCCC AAAAAAAAGG AAAAAGAAAA AACAAAAATC CAAATAAAAA ACAACAAAAA TCCAAGCTAA AACAACAGAA***

168 ***AAACGTAGAA AACCCGCTAC ACTGACACAA AAACATAGCC AGTCTAACGA ATATTTTGTA GAAACCTCCC ACAACAGAAA***

248 ***AAAAAAAAAA ACATTCGCAA CAAAATGCTG CCACGAATCG CAGGCAAATG AGCCCGCTAG CCCCTAACGG CTCCCATCTG***

328 ***CGGCTGCTGC ACAGCAGCGA GCCCGAGCCC GGTCCCGCAT CCGCACTCGC ATCCGCATAC GCATCCGTCT ACGTATCCGC***

408 ***ATTCTCATCC ACATCCGCAT CAGCCACTCG CCGCGTCTCC CAGTGTTTGT GCCAGTGTTC AGCCAAAGCG GACCG***

DISCUSSION

One of the most surprising and important discoveries in the field of metazoan transcription is the demonstration that large protein complexes are required for linking sequence-specific upstream regulatory factors to the Pol II transcription machinery. Metazoan development requires precise temporal and spatial expression of numerous genes. Studies on gene activation in evolutionary diverse organisms is consistent with the idea that activator proteins recruit basal transcriptional machinery such as TATA-binding protein (TBP) and RNA polymerase II holoenzyme to the promoter (Ptashne and Gann 1997).

The primer extension analysis and RACE-PCR experiment suggested one transcription start of the *als* gene, which is located 29 kb upstream from the translation start site. Common start sites were identified in both late embryonic stage and one-day-old adult, indicating a conservation of transcription initiation during development in embryonic as well as one-day-old adult stages.

Computer-assisted analysis of the nucleotide sequence upstream of the cluster of the transcription start sites reveals the absence of TATA and CCAAT boxes. Although one TATA and three CCAAT boxes are found further upstream, they are located too far away to be associated

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

with control of *als* expresson. The absence of a TATA-box is a notable feature of many housekeeping genes, oncogenes, growth factors, and transcription factors. There are two classes of promoters that lack obvious TATA boxes, one is GC-rich and the other is not. The GC-rich promoters are found primarily in housekeeping genes that are constitutively active, contain several transcriptional initiation sites spread over a fairly large region, and usually contain several binding sites (i.e. 5'-CCCCTCCC-3') for the transcription factor Sp1 (Dynan and Tjian , 1983). The other class which does not contain GC-rich sequence are regulated during development and initiate transcription at a few tightly clustered sites. In addition, this type of promoters contains multiple copies of the binding sites for the transcription factors, ZESTE and GAGA (Biggin and Tjian 1988; Mulholland and Kingston *et al*. 2003). The promoter structures of several developmentally important *Drosophila* genes, such as *engrailed* (*en*), *Ultrabithorax* (*Ubx*), *Antennapedia* (*Antp*) and *E74* have been reported to fall in this class (Biggin and Tjian, 1988; Thummel, 1989).

Searching for more possible sequence motifs of known *cis*-regulatory elements has suggested that the proximal promoter sequence might harbor binding sites for several other sequence-specific transcription factors including two GC-like boxes (-53 to -44;-13 to -7), two E-boxes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(-40 to -35, +322 to +327), five *Zeste* binding sites (-586 to -581, -224 to -219, -68 to -63, +45 to +50, +371 to +376) and at least three GAGA binding sites (-216 to -210, -124 to -118, -114 to -107).

The GAGA transcription factor is encoded by the *Trithorax-like (Trl)* gene (Farkas et al. 1994) and was first discovered as an activator of the *ubx* and *en* genes by its association with heat shock and histone gene promoters (Biggin and Tjian 1998; Soeller *et al.* 1998; Gilmour 1989). Normal expression of several developmental genes, including homeotic genes, requires GAGA (Farkas *et al.* 1994; Bhat *et al.* 1996). Immunofluorescent staining of polytene chromosomes revealed that many euchromatic genes include binding sites for GAGA, suggesting a general role in transcriptional control (Benyajati et al 1997). GAGA is known to enhance transcription from promoters containing d(GA·TC)n sequences (Tchoubrieva and Gibson 2004). It has been shown to counteract chromatin repression at all levels (Okada and Hirose 1998), and by so doing, trigging the active transcription of genes subject to repression. Evidence has been presented that GAGA is essential for enhancer trapping. Inspection of ~250 *Drosophila* promoter sequences (Arkhipova 1995) reveals that ~15% contain at least one optimal GAGA element within 50 bp 5′ of the transcription start site. The function of GAGA is not restricted to that of a gene-specific transcriptional activator.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

GAGA has also been implicated in the functioning of the polycomb response elements (Strutt *et al.* 1997). Immunolocalization studies revealed a strong association of GAGA with the GA-rich centric heterochromatin throughout the cell cycle in early embryos (Raff *et al.* 1994). Thus, GAGA is a multipurpose protein that mediates gene-specific regulation but also plays a global role in chromosome function. Although there is considerable variability in GAGA sites, optimal binding sequences contain a minimal GAGAG pentanucleotide.

Zeste, a member of the Trithorax group of proteins, binds DNA in a manner that is reminiscent of that of GAGA. Zeste is an interchromosomal facilitator required for the ability of enhancers to activate a promoter on a paired chromosome, for example, the interchromosomal interactions *Ultrabithorax* (*Ubx*) and *yellow*. Zeste binds to a specific DNA sequence through an amino-terminal DNA-binding domain (Benson and Pirrotta 1987), and self-interacts through hydrophobic repeats at the carboxyl terminus (Bickel and Pirrotta 1990). The interchromosomal effects of Zeste require both DNA binding and self-interaction. Thus it has been suggested that Zeste supports long-range and interchromosomal enhancer–promoter interactions by crosslinking enhancers and promoters together (Qian and Pirrotta 1992).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The E boxes with consensus sequences CANNTG is the binding sites for transcription factors containing a helix-loop-helix (HLH) domain. E boxes have been reported to exist in all 5' flanking regions of characterized muscle nAChR subunit genes within their putative promoter regions, although they are different in their number and distance to the transcriptional start site. At least two functional E box elements, separated only by a short stretch of nucleotides, are present in the promoter regions of the chicken, mouse and rat γ subunit gene and in the α subunit enhancer of chicken and mouse. In contrast, the mouse β, the rat and mouse ε and rat δ minimal promoter regions contain only one CANNTG consensus sequence. In addition, a number of transcription factors bearing a basic helix-loop-helix (bHLH) motif have been described that are required for regulation of the neural programme. In *Drosophila*, the proneural genes containing the HLH motif have also been found to be involved during neurogenesis (for review see Campos-Ortega, 1991). They are important for development of the central and peripheral nervous system in the embryo. Such proteins are thought to form two amphipathic helices separated by a loop that mediates dimer formation. This would enable the formation of homodimers or heterodimers with specific DNA-binding and transcription regulatory properties. Expression of these genes is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

thought to provide the cells with neural potential. Homologues of these genes have been found in a number of vertebrate species in which they are expressed in developing nerve cells. Analyzing the upstream sequence from the transcription start site, two E boxes are found within 350 bp region. One of them is within the critical sequence from -40 to -35. Whether this single E box is involved in the expression regulation or additional downstream E box is required needs further investigation.

So far, ten different subunits of nAChR genes in *Drosophila melanogaster* have been identified, i.e. ALS, Dα2, Dα3, Dα4, Dα5, Dα6, Dα7, ARD, SBD and Dβ3. Two of the subunits (i.e. ALS and ARD) have been suggested by using immunoprecipitation experiments to be components of the same receptor complex (Schloss *et al.*, 1991; Schloss *et al.,* 1992). In addition, ALS, Dα2, and ARD have been shown to share a widespread similar distribution in the central nervous system of *Drosophila* and their expression may be overlapping (Schuster *et al.,* 1993; Jonas *et al.,* 1994). These subunit genes must be regulated in a highly coordinate way. Therefore, it is interesting to compare the promoter regions between these subunits genes. To date, only the 5'-flanking regions required for transcriptional regulation of *ard* and *Dα2* have been reported (Hess *et al.,* 1994; Jonas *et al.,* 1994). Just like the *ard* gene, the *Dα2* gene does not have TATA box-like sequence. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

only common sequence between *ard* and *Dα2* in the 5' upstream regions is located between -998 and -984 relative to the translation start site of the *ard* gene and between -676 and -662 relative to the transcription start site of the *Dα2* gene. These two region share 14 identical sequences and include an E-box. It is noted that, in the *Dα2* gene, six GAGAG sequence motifs are found within the first 350 bp upstream of the transcription start site.

The variable transcription of the genes encoding nAChR subunits across the nervous system generates a diversity of distribution and thus of colocaliation of subunits. The combinatorial assembly of these subunits produces a wide structural diversity of receptor oligomers, targeted to different subcellular compartments, which exhibit variable electrical properties (conductance, ion selectivity, rectification), pharmacologic characteristics (affinities for agonists, competitive antagonists and allosteric effectors, potency orders) and kinetics of activation and desensitization. Different subunits of the same subtype may mutually replace each other within hetero-oligomeric receptors, yet with the obvious condition that they are expressed at the same cellular and subcellular locations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

PART 2. Phylogenetic Study of the nAChR Subunit Genes

MATERIAL AND METHOD

1. Sequence Collection

Nucleotide sequences (coding portion) were obtained from the public databases (NCBI Entrez: http://www.ncbi.nlm.nih.gov.entrez/) for a representative subset of 111 neuronal nAChR subunit nucleotide sequences from 24 different organisms. A combination of keyword search and BLAST probing with the *als* cDNA sequence against the nucleotide databases was performed with the BLASTN (Altschul *et al.* 1997) using default parameters. Blast probing returns a taxonomic view of BLAST hits, where each individual sequence was directly retrieved from the GenBank into VectorNTI Suite 8.0 (Informax, Inc.), where ORF were readily analyzed. Redundant data of subunit types for each species was excluded. The resulting 111 coding sequences were then collected into the initial dataset (**Table 6**).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 6. nAChR Subunit Nucleotide Sequences Used in Phylogenetic Study

Species	Common Name	Subunit	Abbreviation	Accession No.	References
Homo sapiens	Human	$\alpha 1$	Hu_alpha1	S77094	Gattenlohner *et al*. (1994)
		$\alpha 2$	Hu_alpha2	U62431	Elliott *et al*. (1996)
		$\alpha 3$	Hu_alpha3	M37981	Mihovilovic and Roses (1991)
		$\alpha 4$	Hu_alpha4	Y08421	Groot-Kormelink *et al*. (1997)
		$\alpha 5$	Hu_alpha5	Y08419	Groot-Kormelink *et al*. (1997)
		$\alpha 6$	Hu_alpha6	NM_004198	Ebihara *et al*. (2002)
		$\alpha 7$	Hu_alpha7	NM000746	Burghaus *et al*. (2003)
		$\beta 2$	Hu_beta2	X53179	Anand and Lindstrom (1990)
		$\beta 3$	Hu_beta3	Y08417	Groot Kormelink *et al*. (1997)
		$\beta 4$	Hu_beta4	NM_000750	Graham *et al*. (2003)
Bos Taurus	Cow	α	Cow_alpha	X57032	Craido *et al*. (1992)
		$\alpha 3$	Cow_alpha3	NM_174719	Maneu *et al*. (2002)
		$\alpha 5$	Cow_alpha5	AF487464	Campos-Caro *et al*. (1997)
		$\alpha 7$	Cow_alpha7	X93604	Garcia-Guzman *et al*. (1995)
		$\beta 1$	Cow_beta1	X00962	Tanabe *et al*. (1984)
		$\beta 4$	Cow_beta4	AF487465	Campos-Caro *et al*. (1997)
Rattus norvegicus	Rat	$\alpha 2$	Rat_alpha2	AY574253	Groot-Kormelink (2004)
		$\alpha 3$	Rat_alpha3	AY574254	Groot-Kormelink (2004)
		$\alpha 4$	Rat_alpha4	NM_024354	Cucchiaro and Commons

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

		α4	Rat_alpha4	NM_024354	Cucchiaro and Commons (2003)
		α5	Rat_alpha5	AY574255	Groot-Kormelink (2004)
		α6	Rat_alpha6	NM_057184	Boulter (unpublished data 1993)
		α7	Rat_alpha7	AY574256	Groot-Kormelink (2004)
		β2	Rat_beta2	AY574258	Groot-Kormelink (2004)
		β3	Rat_beta3	AY574259	Groot-Kormelink (2004)
		β4	Rat_beta4	AY574260	Groot-Kormelink (2004)
Mus musculus	House mouse	α2	Mo_alpha2	AY574261	Groot-Kormelink (2004)
		α3	Mo_alpha3	AY574262	Groot-Kormelink (2004)
		α4	Mo_alpha4	AY574263	Groot-Kormelink (2004)
		α5	Mo_alpha5	AY574264	Groot-Kormelink (2004)
		α6	Mo_alpha6	AY574265	Groot-Kormelink (2004)
		β2	Mo_beta2	AY574267	Groot-Kormelink (2004)
		β3	Mo_beta3	AY574268	Groot-Kormelink (2004)
		β4	Mo_beta4	AY574269	Groot-Kormelink (2004)
Gallus gallus	Chicken	α1	Ch_alpha1	NM_204816	Nef and Ballivet (1988)
		α2	Ch_alpha2	NM_204815	Nef and Ballivet (1988)
		α3	Ch_alpha3	M37336	Couturier and Bertrand (1990)
		α4	Ch_alpha4	AJ250361	Nef and Ballivet (1988)
		α5	Ch_alpha5	J05642	Couturier and Bertrand (1990)
		α6	Ch_alpha6	NM_205364	Fucile and Eusebi *et al.* (1998)
		α7	Ch_alpha7	X68246	Matter-Sadzininski and Matter (1992)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

		α8	Ch_alpha8	X52296	Schoepfer and Lindstrom *et al*. (1990)
		α9	Ch_alpha9	AF082192	Hiel and Fuchs *et al*. (2000)
		β2	Ch_beta2	AJ250362	Nef and Ballivet (1988)
		β3	Ch_beta3	X83739	Hernandez and Matter *et al*. (1995)
		β4	Ch_beta4	NM_204819	Couturier and Bertrand (1990)
Danio rerio	Zebrafish	α1	Zf_alpha1	U70438	Sepich *et al*. (unpublished data)
		α2	Zf_alpha2	AY256908	Zirger, J.M., *et al*. (2003)
		α7	Zf_alpha7	AY247962	Zirger, J.M., *et al*. (2003)
		β3	Zf_beta3	AY256909	Zirger, J.M., *et al*. (2003)
Carassius auratus	Goldfish	α2	Gf_alpha2	X14786	Cauley, Agranoff, and Goldman (1989)
		α3	Gf_alpha3	X54051	Hieber *et al*. (1990)
		β2	Gf_beta2	X54052	Hieber *et al*. (1990)
Fugu rubripes	Pufferfish	α2	Pf_alpha2	AY299460	Jones and Sattelle *et al*. (2003)
		α3	Pf_alpha3	AY299461	Jones and Sattelle *et al*. (2003)
		α4	Pf_alpha4	AY299462	Jones and Sattelle *et al*. (2003)
Hypnos monopterygium	Electric ray	α	Hyp_alpha	AY472106	Tierney and Howitt *et al*. (2003)
		β	Hyp_beta	AY472107	Tierney and Howitt *et al*. (2003)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Caenorhabditis	C elegans	α7-like	acr-7	NM_063246	Mongan and Sattelle *et al.* (2002)
elegans		α7-like	acr-9	NM_077884	Mongan and Sattelle *et al.* (2002)
			acr-10	NM_076291	Kamath and Ahringer *et al.* (2003)
		α7-like	acr-11	NM_059505	Mongan and Sattelle *et al.* (2002)
			acr-12	NM_077861	Mongan and Sattelle *et al.* (2002)
			acr-15	NM_072805	Mongan and Sattelle *et al.* (2002)
		α7-like	acr-16	NM_072806	Raymond and Sattelle *et al.* (2000)
			acr-19	NM_060028	Mongan and Sattelle *et al.* (2002)
Drosophila	Fruitfly	α-like (α1)	Dm_als	AAF56301	Bossy and Spierer *et al.* (1988)
melanogaster		α2	Dm_alpha2	SWP:P17644	Jonas *et al.* (1990)
		α3	Dm_alpha3	SPTREMBL:Q9W3G6	Schulz *et al.* (1998)
		α4	Dm_alpha4	SPTREMBL:Q9NFR5	Lansdell and Millar (2000)
		α5	Dm_alpha5	SPTREMBL:Q9VJT9	Grauso *et al.* (2002)
		α6	Dm_alpha6	SPTREMBL:Q8IPE2	Grauso *et al.* (2002)
		α7	Dm_alpha7	SPTREMBL:Q9VWI9	Grauso *et al.* (2002)
		β1	Dm_ard	SWP:P04755	Hermans-Borgmeyer *et al.* (1986)

		β2	Dm_sbd	SWP:P25162	Sawruk *et al*. (1990)
		β3	Dm_beta3	SPTREMBL:Q8 IPV6	Lansdell and Millar, (2002)
Myzus persicae	Green peach aphid	α1	Myz_alpha1	X81887	Sgard and Windass *et al*. (unpublished data 1997)
		α2	Myz_alpha2	X81888	Sgard and Windass *et al*. (unpublished data 1997)
		α3	Myz_alpha3	AJ236786	Huang and Millar (1999)
		α4	Myz_alpha4	AJ236787	Huang and Millar (1999)
Aphis gossypii	Cotton aphid	α1	Ag_alpha1	AF527784	Li, F. and Han, Z.J. (unpublished 2002)
		α2	Ag_alpha2	AF527783	Li, F. and Han, Z.J. (unpublished 2002)
		β1	Ag_beta1	AF527785	Li, F. and Han, Z.J. (unpublished 2002)
Nilaparavata lugens	Brown planthopper	α	BpA1	AY378698	Liu and Zhang *et al*. (unpublished 2003)
		α	BpA2	AY378699	Liu and Zhang *et al*. (unpublished 2003)
Locusta migratoria	Migratory	α1	Lm_alpha1	AJ000390	Hermsen and Maelicke *et al*. (1998)
	locust	α2	Lm_alpha2	AJ000391	Hermsen and Maelicke *et al*. (1998)
		α3	Lm_alpha3	AJ000392	Hermsen and Maelicke *et al*. (1998)
		β	Lm_beta	AJ000393	Hermsen and Maelicke *et al*. (1998)
Schistocerca gregaria	Desert locust	α	Sgre_alpha	X55439	Marshall and Barnard *et al*. (1990)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Haemonchus contortus	Barber pole worm	α	Hc_alpha	U72490	Hoekstra and Visser *et al.* (1997)
Manduca sexta	Tabacco hornworm	α-like (α1)	Ms_alpha	Y09795	Eastham and Wonnacott *et al.* (1998)
Chilo suppressalis		α	Chsup_alpha	AF418987	Han,Z. and Han, Z. Sr. (unpublished data 2001)
Heliothis virescens	Tobacco budworm	α1	Hv_alpha1	AJ000399	Jafarigorzini and Maelicke (unpublished data 1997)
		α2	Hv_alpha2	AF096878	Schulte and Adamczewski *et al.* (unpublished 1998)
		α3	Hv_alpha3	AF096879	Schulte and Adamczewski *et al.* (unpublished 1998)
		β1	Hv_beta1	AF096880	Schulte and Adamczewski *et al.* (unpublished 1998)
Schistosoma	Blood fluke	α	Sman_alpha	AY392154	Bentley and Agnew et al. (2004)
mansoni		non-α	Sman_beta	AY392155	Bentley and Agnew *et al.* (2004)
Schistosoma	Blood fluke	α	Shae_alpha	AY392150	Bentley and Agnew *et al.* (2004)
haematobium		non-α	Shae_beta	AY392151	Bentley and Agnew *et al.* (2004)
Schistosoma bouis		α	Sbou_alpha	AY392152	Bentley and Agnew *et al.* (2004)
		non-α	Sbou_beta	AY392152	Bentley and Agnew *et al.* (2004)

2. Phylogenetic reconstruction

The amino acid sequences translated from the 111 coding sequences were aligned by CLUSTAL W (Thompson *et al.* 1994). Divergent sequences (less than 30% sequence identity) were eliminated from the analysis, resulting a final dataset of 107 sequences.

In order to infer the ancestor-descendant relationships of the sequences, we used the bilateria *Schistosoma haematobium* (blood fluke) as the outgroup in this study.

The alignment were analyzed using a Bayesian approach implemented in MrBayes version 2.0 (Huelsenbeck and Ronquist 2001), based on a nucleotide evolution model with three separate substitution rates for the three codon positions. Four MCMC (Markov chain Monte Carlo) chains were run for 10 million generations to reach convergence. Trees were saved every 1,000 generations and a total of 10,000 trees were obtained. A majority-rule consensus tree with posterior branch supports were obtained from the last 2,000 sampled trees.

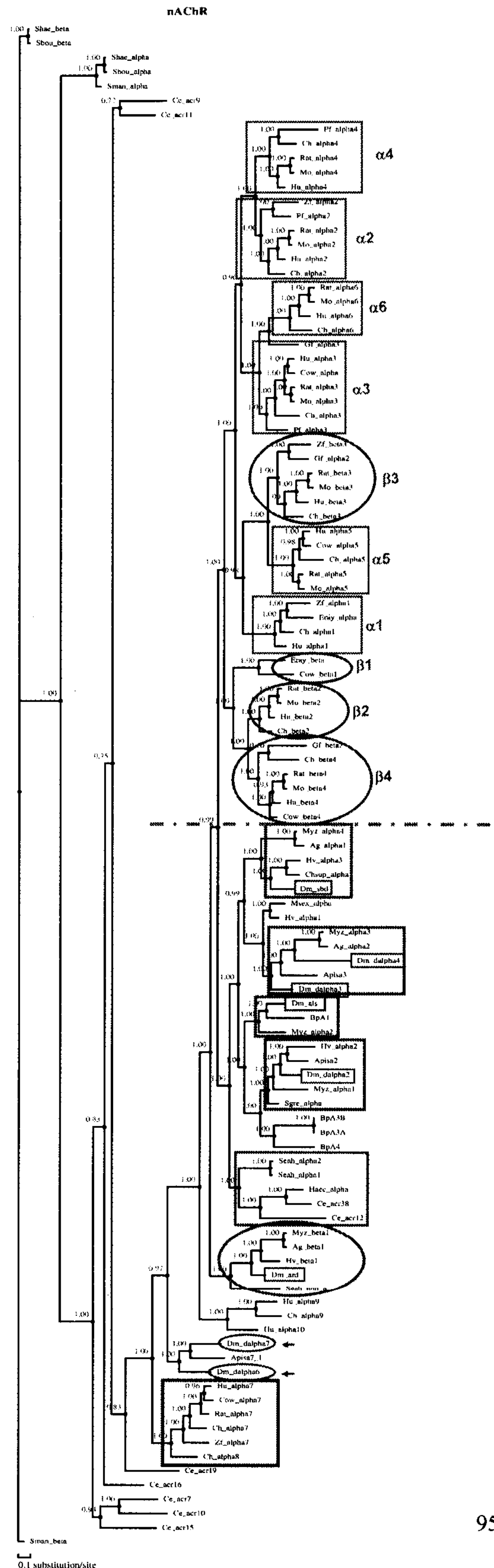
RESULT

Figure 12. Phylogenetic relationships among 107 nAChR subunit genes. Nucleotide coding sequences of 107 representative members of nAChR subunit genes from vertebrates and invertebrates were aligned using the clustalX program package. Numbers at the nodes represent the probability in support of each tree branch baesd on analysis of a Baysian approach. Abbreviations are shown in **Table 6**. (See next page)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Phylogenetic Study of Nicotinic Acetylcholine Receptors

Species	Common Name	Subunit	Abbreviation	Accession No.	References
Homo sapiens	Human	α1	Hu_alpha1	S77094	Gattenlohner,S., et al. (1994)
		α2	Hu_alpha2	U62431	Elliott, K.J., et al. (1996)
		α3	Hu_alpha3	M37981	Mihovilovic and Roses (1991)
		α4	Hu_alpha4	Y08421	Groot Kormelink, et al. (1997)
		α5	Hu_alpha5	Y08419	Groot Kormelink, et al. (1997)
		α6	Hu_alpha6	NM_004198	Ebihara, M., et al. (2002)
		α7	Hu_alpha7	NM000746	Borghese, L., et al. (2003)
		β2	Hu_beta2	X53179	Anand and Lindstrom (1990)
		β3	Hu_beta3	Y08417	Groot Kormelink, et al. (1997)
		β4	Hu_beta4	NM_000750	Graham, A.J. et al. (2003)
Bos Taurus	Cow	α	Cow_alpha	X57032	Criado and Navarro et al. (1992)
		α3	Cow_alpha3	NM_174719	Manou, V., et al. (2002)
		α5	Cow_alpha5	AF487464	Campos-Caro, A., et al. (1997)
		α7	Cow_alpha7	X93604	Garcia-Guzman et al. (1995)
		β1	Cow_beta1	X00962	Tanabe et al. (1984)
		β4	Cow_beta4	AF487463	Campos-Caro, A., et al. (1997)
Rattus norvegicus	Rat	α2	Rat_alpha2	AY574253	Groot-Kormelink, P.J. (2004)
		α3	Rat_alpha3	AY574254	Groot-Kormelink, P.J. (2004)
		α4	Rat_alpha4	NM_024354	Cucchiaro and Commons (2003)
		α5	Rat_alpha5	AY574255	Groot-Kormelink, P.J. (2004)
		α6	Rat_alpha6	NM_057184	Boulter, J. (unpublished data 1993)
		α7	Rat_alpha7	AY574256	Groot-Kormelink, P.J. (2004)
		β2	Rat_beta2	AY574258	Groot-Kormelink, P.J. (2004)
		β3	Rat_beta3	AY574259	Groot-Kormelink, P.J. (2004)
		β4	Rat_beta4	AY574260	Groot-Kormelink, P.J. (2004)
Mus musculus	House mouse	α2	Mo_alpha2	AY574261	Groot-Kormelink, P.J. (2004)
		α3	Mo_alpha3	AY574262	Groot-Kormelink, P.J. (2004)
		α4	Mo_alpha4	AY574263	Groot-Kormelink, P.J. (2004)
		α5	Mo_alpha5	AY574264	Groot-Kormelink, P.J. (2004)
		α6	Mo_alpha6	AY574265	Groot-Kormelink, P.J. (2004)
		β2	Mo_beta2	AY574267	Groot-Kormelink, P.J. (2004)
		β3	Mo_beta3	AY574268	Groot-Kormelink, P.J. (2004)
		β4	Mo_beta4	AY574269	Groot-Kormelink, P.J. (2004)
Gallus gallus	Chicken	α1	Ch_alpha1	NM_204816	Nef, P., and Ballivet, M. (1988)
		α2	Ch_alpha2	NM_204815	Nef, P., and Ballivet, M. (1988)
		α3	Ch_alpha3	M37336	Couturier, S., and Bertrand, D. (1990)
		α4	Ch_alpha4	AJ250361	Nef, P., and Ballivet, M. (1988)
		α5	Ch_alpha5	J05642	Couturier, S., and Bertrand, D. (1990)
		α6	Ch_alpha6	NM_205364	Fucile, S., and Eusebi, F., et al. (1998)
		α7	Ch_alpha7	X68246	Matter-Sadzinski and Matter (1992)
		α8	Ch_alpha8	X52296	Schoepfer and Lindstrom et al. (1990)
		α9	Ch_alpha9	AF082192	Hiel and Fuchs et al. (2000)
		β2	Ch_beta2	AJ250362	Nef and Ballivet (1988)
		β3	Ch_beta3	X83739	Hernandez and Matter et al. (1995)
		β4	Ch_beta4	NM_204819	Couturier and Bertrand (1990)
Danio rerio	Zebrafish	α1	Zf_alpha1	U70438	Sepich et al. (unpublished data)
		α2	Zf_alpha2	AY256908	Zirger, J.M., et al., (2003)
		α7	Zf_alpha7	AY247962	Zirger, J.M., et al., (2003)
		β3	Zf_beta3	AY256909	Zirger, J.M., et al., (2003)
Carassius auratus	Goldfish	α2	Gf_alpha2	X14786	Cauley, Agranoff, and Goldman (1989)
		α3	Gf_alpha3	X54051	Hieber et al. (1990)
		β2	Gf_beta2	X54052	Hieber et al. (1990)
Fugu rubripes	Pufferfish	α2	Pf_alpha2	AY299460	Jones and Sattelle et al. (2003)
		α3	Pf_alpha3	AY299461	Jones and Sattelle et al. (2003)
		α4	Pf_alpha4	AY299462	Jones and Sattelle et al. (2003)
Hypnos	Electric ray	α	Hyp_alpha	AY472108	Tierney and Howitt et al., (2003)
monopterygium		β	Hyp_beta	AY472107	Tierney and Howitt et al., (2003)
Caenorhabditis	C elegans	α7-like	acr-7	NM_065348	Mongan and Sattelle et al. (2002)
elegans		α7-like	acr-9	NM_077884	Mongan and Sattelle et al. (2002)
			acr-10	NM_076291	Kamath and Ahringer et al. (2003)
		α7-like	acr-11	NM_068808	Mongan and Sattelle et al. (2002)
			acr-12	NM_077881	Mongan and Sattelle et al. (2002)
			acr-15	NM_072806	Mongan and Sattelle et al. (2002)
		α7-like	acr-16	NM_072809	Raymond and Sattelle et al. (2000)
			acr-19	NM_060028	Mongan and Sattelle et al. (2002)
Drosophila	Fruitfly	α-like (α1)	Dm_als	AAF58301	Bossy B, Ballivet M, Spierer P. (1988)
melanogaster		α2	Dm_alpha2	SWP:P17644	Jones et al. (1990)
		α3	Dm_alpha3	SPTREMBL:Q9W3G6	Schulz et al. (1998)
		α4	Dm_alpha4	SPTREMBL:Q9NFR5	Lansdell and Millar (2000)
		α5	Dm_alpha5	SPTREMBL:Q9VJT9	Grauso et al. (2002)
		α6	Dm_alpha6	SPTREMBL:Q8IFE2	Grauso et al. (2002)
		α7	Dm_alpha7	SPTREMBL:Q9VW19	Grauso et al. (2002)
		β1	Dm_ard	SWP:P04755	Hermans-Borgmeyer et al. (1986)
		β2	Dm_sbd	SWP:P25162	Sawruk et al., (1990)
		β3	Dm_beta3	SPTREMBL:Q8IPV8	Lansdell and Millar, (2002)
Myzus persicae	Green peach aphid	α1	Myz_alpha1	X81887	Sgard and Windass et al.(unpublished data 1997)
		α2	Myz_alpha2	X81888	Sgard and Windass et al.(unpublished data 1997)
		α3	Myz_alpha3	AJ236786	Huang, Y., and Millar, N.S. (1999)
		α4	Myz_alpha4	AJ236787	Huang, Y., and Millar, N.S. (1999)
Aphis gossypii	Cotton aphid	α1	Ag_alpha1	AF527784	Li, F. and Han, Z.J. (unpublished 2002)
		α2	Ag_alpha2	AF527783	Li, F. and Han, Z.J. (unpublished 2002)
		β1	Ag_beta1	AF527785	Li, F. and Han, Z.J. (unpublished 2002)
Nilaparavata lugens	Brown planthopper	α	BpA1	AY378698	Liu and Zhang et al. (unpublished 2003)
		α	BpA2	AY378699	Liu and Zhang et al. (unpublished 2003)
Locusta migratoria	Migratory locust	α1	Lm_alpha1	AJ000390	Hermsen and Maelicke et al. (1998)
		α2	Lm_alpha2	AJ000391	Hermsen and Maelicke et al. (1998)
		α3	Lm_alpha3	AJ000392	Hermsen and Maelicke et al. (1998)
		β	Lm_beta	AJ000393	Hermsen and Maelicke et al. (1998)
Schistocerca gregaria	Desert locust	α	Sgre_alpha	X55439	Marshall, J. and Barnard, E.A. et al. (1990)
Haemonchus contortus	Barber pole worm	α	Hc_alpha	U72490	Hoekstra, R., Visser, A., et al. (1997)
Manduca sexta	Tobacco hornworm	α-like (α1)	Ms_alpha	Y09795	Eastham, H.M., and Wonnacott, S., et al. (1998)
Chilo suppressalis		α	Chsup_alpha	AF418987	Han,Z. and Han, Z. Sr. (unpublished data 2001)
Heliothis virescens	Tobacco budworm	α1	Hv_alpha1	AJ000399	Jafarigorzini and Maelicke (unpublished data 1997)
		α2	Hv_alpha2	AF096878	Schulte, Oellers and Adamczewski (unpublished 1998)
		α3	Hv_alpha3	AF096879	Schulte, Oellers and Adamczewski (unpublished 1998)
		β1	Hv_beta1	AF096880	Schulte, Oellers and Adamczewski (unpublished 1998)
Schistosoma	Blood fluke	α	Sman_alpha	AY392154	Bentley and Agnew et al. (2004)
mansoni		non-α	Sman_beta	AY392155	Bentley and Agnew et al. (2004)
Schistosoma	Blood fluke	α	Shae_alpha	AY392150	Bentley and Agnew et al. (2004)
haematobium		non-α	Shae_beta	AY392151	Bentley and Agnew et al. (2004)
Schistosoma bovis		α	Sbou_alpha	AY392152	Bentley and Agnew et al. (2004)
		non-α	Sbou_beta	AY392152	Bentley and Agnew et al. (2004)
Aplysia californica	California sea hare	α1	Ac_alpha1	AF467898	McCaman, Ond and Chang (unpublished data 2002)
		α2	Ac_alpha2	AF467899	Sadreyev et al. (unpublished 2000)
		non-α	Ac_beta	AF322876	Sadreyev et al. (unpublished 2000)



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 12 shows a phylogenetic tree constructed from 107 coding sequences of neuronal nAChR subunit genes from 24 different species in the CNS (see **Table 6**). The Bilateria *Schistosoma haematobium* (blood fluke) was selected as the outgroup in this study. This tree contains more subunits from invertebrates comparing with those in previous works (Gundelfinger 1995; Le Novere and Changeux 1995; Ortell and Lunt 1995), which reflects our immediate interest in the *Drosophila* nAChR subunits, their relationship with other insect nAChR subunits and their vertebrate counterparts.

The phylogenetic relationships among the nAChR subunit genes are well resovled by the Bayesian phylogenetic analysis. Almost all branches on the tree are highly supported by the data (most posterior probabilities are close to 1). Several conclusions can be drawn from this phylogenetic analysis:

(1) Ancestor of *Drosophila* nAChR genes

It has been hypothesized that the ancestral subunit which had a ligand-binding site appeared first in the nervous system (Le Novere and Changeux 1995; Ortells and Lunt 1995). The ancestral subunit may have functioned as a homooligomer in the primitive Bilateria (Tsunoyama and Gojobori 1998). The phylogenetic relationship demonstrated in this tree also indicated that the vertebrate $\alpha7$ and $\alpha9$ subunits, both of which are capable of forming homooligomeric receptors,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

diverged from the Nematode *C.elegans* subunit genes, who share the common ancestor Bilateria. *Drosophila* Dα6 and Dα7 (indicated by two arrowheads) are the most primitive *Drosophila* subunit genes that split early from the remainder, as shown in the purple square box. Both are derived from the vertebrate neuronal α7 and α9 subunits which can yield functional receptors as homooligomers, whereas the other α-type subunits need β-type subunits to form an integral single nAChR. All subunits of vertebrate α7, α9 and insects emerged after the divergence of the nematodes from the common ancestor, Bilateria. The subsequent duplication may have occurred during the evolution of Deuterostomia that have yielded two types of subunits, α and β.

(2) Ancestral gene duplication and functional divergence among paralogous genes

As highlighted on the phylogenetic tree, events of gene duplication and divergence were evident throughout evolutionary history of nAChR subunit genes in both invertebrates and vertebrates. At least five phylogenetic groups of nAChR subunit gene homologs: Dm_sbd, Dm_dalpha4, Dm_dalpha3, Dm_als, Dm_dalpha2 (marked in red rectangular boxes) can be pointed to on the tree; while ten functional groups of vertebrate nAChR subunit homologs are shown in rectangular (for α-type subunit genes) and oval (for β-type subunit genes).

The divergence of *Drosophila* nAChR subunits seems to be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lineage-dependent. In other words, its evolution took place independently from that of the vertebrates, although they shared a common ancestor. Noticeably shown in blue rectangular boxes along with their invertebrate counterparts from *Myzus persicae* (Myz, green peach aphid), *Aphis gossypii* (Ag, cotton aphid), *Apis mellifera* (Apisa, honey bee), *Heliothis virescens* (Hv, Tobacco budworm), *Drosophila* subunit genes *als*, Dα2, Dα3, Dα4 and *sbd* (each individually framed in a red/lighter box) appear to be paralogous genes resulting from gene duplication events. The only exception *ard* along with the other insect β1 subunit genes, seem to diverge earlier as an independent group directly from the mammalian α9 and α10, proceeding to all other vertebrate and invertebrate nAChR subunit genes. These earliest β subunit genes are probably the first ones serving as structural components in heteromeric receptor subtypes.

Meanwhile, phylogenetic groups of vertebrate nAChR genes appear to derive from the common ancestral homomeric receptor genes and undergo divergence through gene duplication in parallel with the invertebrates. This tree showed that an early separation took place between α1, α2, α3, α4, α5, α6, β3 and β1, β2, β4 vertebrate subunits, which is consistent with the view of Ortells and Lunt (1995), Tsunoyama and Gojobori (1998). It is believed that these two subunit groups separated after the emergence of the insect and nematode subunits,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

although there has been sharp disagreement among different research groups on the members constituting the two groups (Tsunoyama and Gojobori 1998).

(3) Diversification of nAChR during metazoan evolution are associated with the evolution of neural system.

The abundance of α- and β-type subunits genes resulting from gene duplication and divergence give rise to diversification of nAChR subunit genes. The possible varying combination of distinct subunit genes are most likely associated with the evolution of neural systems. This is evident on the phylogenetic tree which shows an increasing number of phylogenetic and functional groups in both invertebrates and vertebrate lineages.

DISCUSSION

In an attempt to understand the evolutionary relationship of *Drosophila* nAChR subunits in a larger frame of picture, we have performed phylogenetic analyses of their nucleotide and deduced amino acid sequences. Both the similarities between organisms as a consequence of their common descent and the differences between them that have occurred in their evolution are the objects of inquiry.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Over a billion years ago, the family of nicotinic receptor subunits began to emerge by a series of gene duplications from a single common subunit (LeNovere and Changeux 1995; Ortells and Lunt 1995; Cockcroft 1995). The first branch of the family evolved into the subunits found in neuronal bungarotoxin-sensitive receptors, which include $\alpha7$, $\alpha8$, and $\alpha9$ subunits. Of these three nicotinic subunits, only $\alpha7$ subunits are found in mammalian nervous tissue. $\alpha9$ subunits are only present in cochlea and vestibular organs (Elgoyhen *et al.* 1994; Vetter *et al.* 1999), whereas no mammalian homolog for chick a8 subunits has been observed (Elgoyhen *et al.* 1994). It has been suggested that among all the nicotinic receptor subtypes, neuronal bungarotoxin-sensitive receptors have the most features in common with the primordial nicotinic receptor. One important feature is that both have homomeric structure. All other mammalian nicotinic receptors, with perhaps the exception of $\alpha9$-containing receptors, appear to be heteromeric receptors (Lindstrom, 1997). It is hypothesized that the subunits of the primordial nicotinic homomer were folded and processed into different conformations just as we have observed for neuronal $\alpha7$ subunits. Such folding and processing may serve as a post-translational mechanism used to generate subunit diversity and may be critical for proper functioning of the receptors (Rakhilin et al., 1999). The different subunit conformations found in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

homomeric neuronal bungarotoxin-sensitive receptors could play a role similar to that of different subunit isoforms found within heteromeric nicotinic receptors. The additional proteins and/or factors needed to mediate the neuron-specific folding and processing of $\alpha7$ subunits may provide an important regulatory role by determining when and where functional homomeric receptors are produced. The evolutionary pressure for additional nicotinic subunit isoforms may have been to produce the different subunit conformations without the additional accessory proteins needed for proper folding (Drisdel and Green 2000).

It can be speculated that the power of generating variation by gene duplication and divergence may have played a role in the generation of complexity in the nervous system of animals with elaborate nervous systems. The conservation of sequence throughout evolution indicates selective pressures acting at the level of the protein, and therefore that if a protein is conserved it must be doing something important.

The importance of gene duplication in supplying raw genetic material to biological evolution has been recognized since the 1930s (Ohno 1970, Springer). Recent genomic sequence data provide substantial evidence for the abundance of duplicated genes in all organisms. Prevalence of gene duplication in all three domains of life has been observed. In Bacteria *Mycoplasma pneumoniae*, there are 298 duplicated genes, which account for 44% of total 677 genes; in Eukaryotes: 8,971

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

duplicate genes in *C.elegans*, 49% of total 18,424 genes; 5,536 duplicate genes in *Drosophila*, 41% of total 13,601 genes (Zhang 2003, see *table 1*).

The presence of duplicated genes can be beneficial simply because extra amounts of proteins or RNA products are produced. However, it is proposed that unless the presence of an extra amount of gene product is advantageous, two genes with identical functions are unlikely to be stably maintained when they differ in some aspects of their functions (Nowak 1997), which can occur by subfunctionalization, in which each daughter gene adopts part of the functions of their parental gene (Hughes 1994). One form of subfunctionalization that is potentially important in the evolution of development is division of gene expression after duplication (Force 1999). The other can occur at the protein function level and can lead to functional specialization when one of the duplicated genes becomes better at performing one of the original functions of the progenitor gene. The divergent event of *Drosophila* α- and β-type subunits seem to fit into the first form of subfunctionalization, in which following gene duplication, the β-type subunit genes lose one part of the *cis*-acting elements controlling expression of ligand-binding gene product; while the α-type retains the original function. Such hypothesis appears to explain the emergence of heteromeric receptor subtypes in higher animals and in a later time frame during the evolution

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of nAChRs in general.

The evolutionary analysis of the nAChR genes, although speculative, has proven to be a useful tool in the study from a non-evolutionary perspective, proving new insights into their possible structure and function. Also the knowledge of the phylogenetic relationships between various species could be an invaluable aid in the design of mutations, or chimeras, or in extrapolating known functional characteristics of a particular subunit to those that are evolutionarily related.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 4.
In Vivo Study of the Upstream Regulatory Sequence of the *als* Gene

Eukaryotic transcriptional elements and proteins have been studied extensively in *Drosophila* and vertebrates, and it is clear that many aspects of the transcription machinery are highly conserved, facilitating comparisons among systems. The similarity of complex elements in higher metazoans and the availability of advanced genetic techniques make *Drosophila* particularly well suited to studies of transcriptional regulatory systems. The most amazing advantage offered by fruit fly *Drosophila melanogaster* is the availability of vectors for stable germline transformation (Rubin and Spradling, 1982; Rubin 1988). Although large-scale genome sequencing projects have yielded comprehensive information about the protein coding capacity of entire *Drosophila* genome, however, transcriptional regulatory elements are considerably more difficult to identify from primary sequence alone. Identifying the function of putative regulatory regions requires direct experimental testing. The last episode of my study involves *in vivo* analysis of potential regulatory sequences coupled to reporter genes, permitting the analysis of native enhancer and promoter combinations in a physiologically relevant setting, while DNA sequences can be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

systematically manipulated and then transferred into the germline via P-element derived vectors.

Three general classes of experiments have been performed using P-element-mediated fly transformation. In the first class of experiments, a wild-type gene is inserted into a P-element vector and used to rescue a mutant phenotype by germline transformation. This provides evidence that the cloned DNA corresponds to the gene disrupted by the mutation (Haenlin *et al.* 1985). The second type of P-element experiment involves expression of an ORF under the control of an inducible promoter, usually derived from the *hsp70* heat-shock gene (Moran *et al.* 1979). In this case, a mild heat shock directs expression of the corresponding protein throughout the transformed fly (Lis et al., 1983). In the third class of P-element experiments, a *Drosophila* promoter fused to a reporter gene is used to trace the spatial and temporal activity of that promoter during development. The *lacZ* gene has frequently been used as a reporter due to the availability of chromogenic substrates for *β-galactosidase* (Lis *et al.* 1983; Hiromi and Gehring 1987).

My experimental approach employed the third class of P-element mediated transformation described above, that is, to construct fusion genes by linking putative regulatory elements, i.e. various DNA fragments from the 5′-upstream region of the *als* gene, to the *lacZ*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reporter gene into a P element transformation vector. These constructs can then be inserted in a P-element cloning vector and co-injected with a helper P-element encoding transposase into *Drosophila* eggs prior to the stage of pole cell formation, where the DNA transposes into the germline. The intention is to transform the germline cells so that a stable, transmissible transformant is created (Rubin and Spradling 1982).

Upon P-element mediated transformation, the fusion construct is expected to integrate onto the *Drosophila* chromosomes. The *lacZ*-encoding *E.coli β-galactosidase* is readily assayed in embryos, larvae, and adults. Expression of the fusion genes in transgenic flies can then be analyzed in the natural cellular environment both qualitively and quantitively. Based on experimental and computer-predicted data presented in the previous chapters, we expect to regenerate an appropriate pattern of *als* expression using defined segments of DNA containing potential transcriptional regulatory sequences.

We are aware that *Drosophila* has an endogenous *β-galactosidase* gene; however, this enzyme is expressed only in specific tissues and is not detectable at all developmental stages. Moreover, *Drosophila β-galactosidase* is localized in the lysozomes and has an acidic pH optimum. By performing enzyme assays at a slightly basic pH, we can

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

selectively detect bacterial *β-galactosidase* in transgenic flies (Lis *et al.* 1983).

In this chapter, I will describe "Making Fusion Construct", "Microinjection" and "Analyzing Transgenic Flies" in **Material and Method** section, followed by **Results** and **Discussion**.

MATERIAL AND METHOD

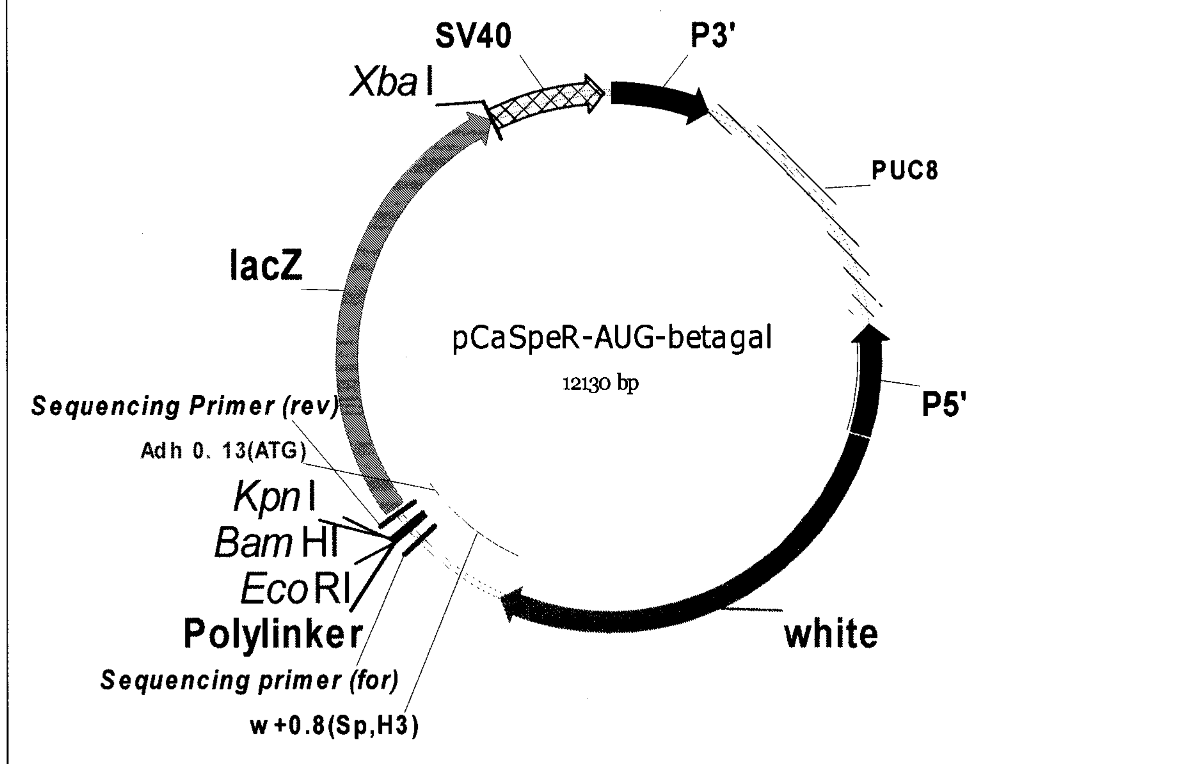
1. Making Fusion Construct:

Two upstream fragments: 2.9 kb and 1.25 kb of the *als* 5'-flanking sequence was fused in phase with the transcriptional reading frame, to the *E.coli* β-*galactosidase* gene.

The pCaSpeR-AUG-βgal vector contains a *Drosophila* $white^+$ gene, heat-shock promoter *Hsp 70* and pUC8 with ampicillin resistant gene (amp^R), flanked by P-element ends. The $white^+$gene is comprised of both genomic and cDNA sequences and contains 300 bp of 5'-flanking DNA. Transformants can be identified by screening for restoration of eye pigmentation in $white^-$ flies. There are three unique restriction sites for the insertion of foreign regulatory sequences: *Eco*RI, *Bam*HI and *Kpn*I (Tummel and Lipshitz 1988). The annotated map of this vector was constructed using the Vector NTI 8.0 program, shown in **Figure 13**.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 13. Molecular Map of the P-element vector pCaSpeR-AUG-βgal.



pCaSpeR-AUG-βgal vector was kindly provided by Professor Carl S. Thummel's lab from Howard Hughes Medical Institute. The restriction enzyme sites for EcoRI, BamHI and KpnI in front of *Adh-lacZ* reporter gene enable us to operate directional cloning of the *als* upstream fragments.

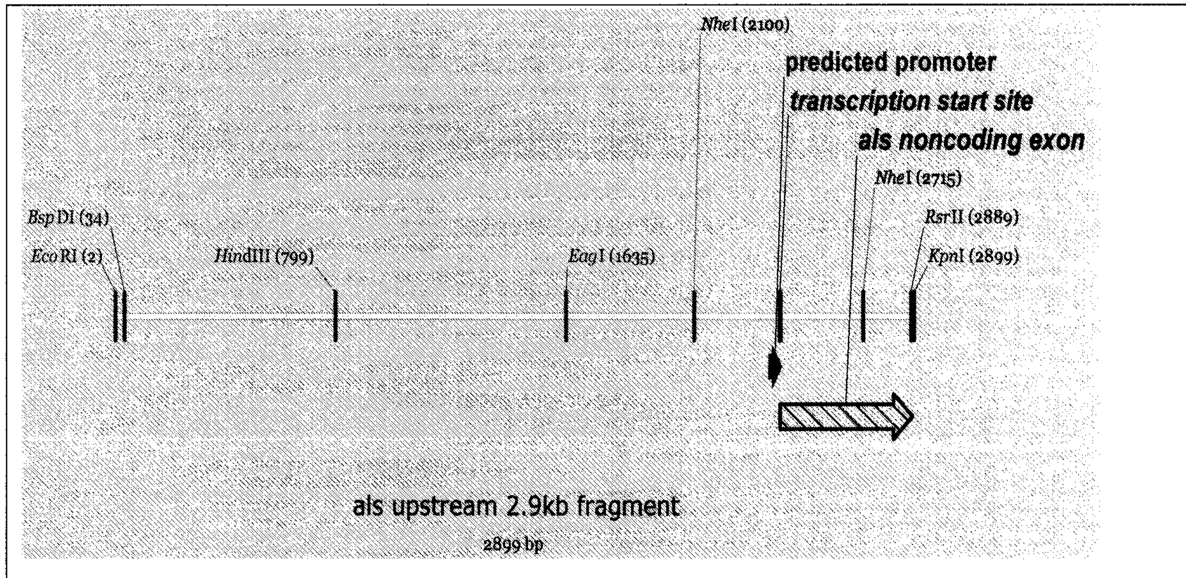
To facilitate construct making and sequence verification later on, forward and reverse sequencing primers flanking the polylinker region were designed as following: forward sequencing primer "7646": 5'-cagataagttcaatgatgtccagtgc-3' and reverse sequencing primer "7873": 5'-tgtccagaccaatgcctccc-3' (listed in **Table 3**).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As illustrated in the **Figure 11A and 14**, two upstream fragments were prepared by PCR amplification, and subsequently cloned into pCaSpeR-AUG-βgal vector. The primers designed to amplify genomic fragments are: M4 RsrII, 27 mer, 5'-ggtacccggtccgctttggctgaacac-3'; M4 Bsp, 30 mer, 5'-gaattccgatgtactcgtaatgaaatcccc-3'; M4 Eag, 25 mer, 5'-gaattcgctggagagtttttcggcc-3' (listed in **Table 3**). PCR reactions were carried out in 50μl total volume with 600ng of BAC DNA from "13 M4" as template, using 200μM of forward and reverse primers, in presence of High-fidelity DNA Polymerase buffer 1X, 200μM of dNTP mix and 1 unit of High-fidelity DNA Polymerase (Strategene). Thermal cycler parameters for touchdown PCR reaction was as follows: ten cycles at 94°C for 30 sec, 68°C for 1 min with 1.5 °C less each consecutive cycle, and 72°C for 2 min; 25 cycles at 94°C for 30 sec, 58°C for 1 min, and 72°C for 1 min and 40 sec; final extension at 72°C for 15 min. PCR product at expected size 1.25 kb and 2.9 kb were checked on 1.0% agarose gel electrophoresis, and then cloned in pCR4Blunt vector (Invtirogen). The two different inserts of 2.9 kb and 1.25 kb were verified by DNA sequencing performed by the Sequencing & Separations Facility at Hunter College (Center for Study of Gene Structure & Function).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 14. The 2.9 kb *als* upstream fragment.



2. Microinjection

Each fly lab has a different set-up for microinjection, especially where needle pullers, micromanipulators, and dissecting scopes are concerned. The following is an established protocol in our laboratory.

(1) Supplies:

- Agar plates with fruit juice (we use apple);
- Yeast paste to feed flies, which is also believed to encourage mating, store at 4°C;
- Embryo wash, "PBTx" (0.002% Triton X-100 in PBS) in a squeeze bottle;
- Small mesh basket for washing and dechorionating embryos;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Superfrost/Plus Microscope Slides (Fisher Scientific);
- Double-sided tape (Scotch, 3M);
- Halocarbon oil (Series 95; Halocarbon Products, SC);
- Moist chambers for incubating injected embryos - large Petri dishes with two supporting bars made from broken pipettes; moistened by putting wet Kimwipes;
- Needles ("home-made", method described later);
- Drying box with Drierite (Aldrich Chem. Co. Inc.);
- Inverted microscope stage (Model: Bio Star; Reichert, Inc.)
- 3-way timer
- Tissue Freezing Medium for frozen tissue specimens (Electron Microscopy Sciences, PA)

(2) Procedures:

i) Preparation of egg-laying plates

1.5 g of agar in 50 ml of apple juice was dissolved by heating in microwave oven, another 50 ml of apple juice was added, brought to boil by microwaving, cooled down and added 1.5g of sugar, mixed by swirling gently and poured onto small plates (about 5cm in diameter for up to 12 plates). Plates were stored at 4°C. Yeast paste was added

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

onto the center of the plate, warmed up before egg collection.

ii) Collection of embryos

yw flies that are deficient in the allele that is to mark the transformants, in this case, w^{1118}, must be in abundant supply, both for egg collections and later crosses. At least ~ 150 flies per egg laying plate are required to get a reasonable amount of embryos in 30 minutes.

Flies were transferred with a funnel into a disposable cup of which the diameter fits right onto the egg-laying plates. The cup was closed with the yeasted plate. The plates were turned over and let sit at least overnight. Usually flies were given two to three days to adapt to the chamber with frequent feeding, before egg collection for injection. Before injection began every day, eggs laid during the first 1~2 hours were discarded since they could have been kept in female flies too long. Ideal temperature range for injections range from 18 to 22 (including the microscope stage); low temperature improves the viscosity of the cytoplasm and slightly delays embryo development which gives higher success rate for injections.

yw fly embryos were collected on apple juice agar plates with yeast paste at 30 minutes intervals and brushed into a 25-ml beaker containing about 3 ml of tap water. Chorions were removed by filling up

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the beaker with an equal volume of 3% sodium hypochlorite and incubating for 1 min at room temperature. The embryo suspension was then poured through a small basket. The small basket was made by removing the bottom of a scintillation vial and drilling a hole into its cap, which was used to secure a nylon mesh (opening 140 um). Embryos were rinsed thoroughly in a gentle stream of tap water for at least 2 minutes, dried by touching with kimwipes, disassemble the nylon mesh basket, transfer the embryos to the non-taped part of a pre-taped slide by several gently touches of the nylon mesh.

iii) Lining up of embryos

A stretch of double sided sticky tape (1/4″ 3M, type 415 tape, Scotch) was mounted about 0.5 ~ 1mm to the edge of a microscopic slide lengthwise. Using a dull probe with tilted end at the tip, embryos were positioned to form a line without touching one another on the edge of a double sticky tape with each embryo oriented in the same direction (posterior poles of embryos facing the edge). The anterior end of an embryo can be recognized by the dorsal appendages, which are two string like parts that come off the egg with the chorion. Also located on the anterior end is the micropile, a small nipple-like extension from the egg that should be visable after removing the chorion. A space of about 1 to 1.5 embryo-widths were left between lined-up embryos, to avoid

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

oxygen depletion. 40-50 embryos were usually lined up on each slide, with one quarter to one fifth of their posterior ends over the edge of the double stick tape.

Usually the young embryos look homogeneous and dark provided they are well dechorionated; while older embryos look more transparent, sometimes cellularization in the peripheral area already begins.

To complete the preparation for injection, embryos were desiccated in a petri dish containing Drierite (Aldrich Chem. Co. Inc.,). The time for embryo lining-up and drying varied depending upon room temperature and humidity. Typically, 5 - 7 minutes for lining-up and 3 - 4 minutes for drying were applied for the first round. While the embryos of the first round were drying, a second round of lining-up was carried out (3 - 4 minutes) and then the second round embryos were dried for 1 -2 minutes. After desiccation, the embryos were covered with a thin layer of Halocarbon oil (Series 95; Halocarbon Products Corporation, North Augusta, SC, USA).

iv) Preparation of glass needle for injections

Needle-pulling technique depends on the configuration of the needle puller. In our lab, needles were pulled from capillaries (o.d. 1.0 mm, i.d. 0.5 mm; World Precision Instruments, Sarasota, FL, USA) using

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

a Vertical Pipette Puller (Model 700B; David Kopf Instruments, Pasadena, CA, USA). Examine pulled needles under dissection microscope, qualified needles were placed in a large Petri dish.

Before injection, sharp tips (proximal diameter of 2 μm) were generated by cutting the apex of the needle with a razor blade under dissection microscope.

v) Preparation of DNA

Three plasmids were used for microinjection: the RB-construct (2.9 kb insert); the RE-construct (1.2kb insert) and the pCaSpeR-AUG-βgal vector alone as control). The injection solution was prepared with a 3:1 molar ratio of insert to helper plasmid. 3 μg of purified transformation plasmid along with 1 μg of "helper" plasmid was ethanol precipitated, washed in 70% Ethanol, and resuspend in 10 μl injection buffer (0.1mM Sodium Phosphate pH7.0, 5mM KCl). Samples were usually spun at maximum speed for a few minutes to pellet junk before load onto the needle. 1 μl of DNA was pipetted into the needle. The tip of the needle should fill by capillary action in a few minutes. The needle was checked under dissection microscope for air bubbles near the tip.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

vi) Injection

The construct-filled glass needle was carefully inserted into a needle holder which was mounted on an inverted microscope stage (Model: Bio Star; Reichert, Inc.). The needle holder was joined by a polyethylene tubing (i.d. 1.14 mm; o.d. 1.57 mm) to a 60-ml syringe. Needle tip was centered under scope. Rear end of embryo was focused on, then needle tip was brought into focal plane. Embryos were slowly moved towards the needle until it punctures the membrane, the tip of the needle was positioned within the embryo at 5% - 10% egg length (posterior pole = 0%) by moving the stage toward it. A tiny amount of DNA solution was injected using air pressure generated by pressing the syringe and quickly withdrew. Enough DNA should be flowing that one can see stuff moving inside the embryo. Most of the DNA was left in the most posterior cytoplasm, where the pole cells will form. The membrane should stretch inward without crumpling, then instantly spring back when punctured. Wrinkled embryos and over aged embryos (indicated by clear cortical cytoplasm, a bumpy posterior surface) were avoided and destroyed.

vii) Post-injection

After microinjection, embryos were kept on the cover slip under oil in a moisture chamber and incubated in the fly room and scoop larvae the next day. Slides are usually kept and closely observed for two days.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Be sure to scoop the larvae before they crawl away. But do not collect larva until they have freed themselves from their egg shell.

Eclosed larvae were transfer to a labeled vial of fresh food containing cornmeal agar medium. The surface of the food were scored before adding scoopees. Place around 10 larva in each vial. Once they have pupated, male and virgin females of the wild-type flies should be collected.

viii) Collecting G_0s and making genetics crosses of transformants

After hatching, each newly eclosed adults of G_0 flies must be collected before mating. Approximately half of these flies will be sterile, due to damage of the developing sex organs done by injecting. G_0 survivals were crossed with the original marker stock, *yw*. Each G_0 was placed in a vial with four to five of *yw* of the opposite sex. Their progeny (G_1) were screened for transformants by identifying yw^+ with restored eye colors, ranging from light yellow, orange to read. Each individual transformant was mated in the same fashion as before.

The G_2 yw^+ flies were crossed to *yw* stocks containing appropriate Balancer chromosomes to establish homozygous stocks. Males/virgin females carring the P-transposon were mated with balancer stock of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

opposite sex. The following balancer stocked were kindly provided by the Bloomington Stock Center: "4728" of genotype *y[1] w[1] sn[3] f[36a] ado[4P1]/FM7a* detects transposon insertion on X chromosome; "4899" of genotype w[1118]; EcR[225]/CyO detects second chromosomal insertions; "3607" of genotype *w[1118]; TM3, Sb[1]/CxD* (genetic crosses were set up to generate heterozygous flies carring RE/RB insert on the 3rd chromosome).

3. Analysing Transgenic Flies

(1) β-galactosidase activity assay

β-galactosidase activity assay was determined according to Kitamoto et al. (1992) with slight modifications. Ten heads of 1-day old fly of each yw^+ transformant line were homogenized in 200 μl of homogenization buffer (50mM Tris-Cl pH 8.0, 0.1 M NaCl, and 0.5% Triton X-100). The homogenates were centrifuged at 14,000 rpm for 15 minutes at 4°C. 25 μl of the supernatant was mixed with 50 μl of 10 mM chlorophenol red-β-D-galactopyranoside (CPRG, Boehringer annheim) and 450 μl of 50 mM sodium phosphate (pH 7.5) and 1 mM $MgCl_2$. Reactions were incubated at 37°C for 1 hour, and the absorbance was read at 574 nm. Samples were assayed in duplicates.

The amount of protein was determined in parallel.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Microassay procedure from the Bio-Rad Protein Assay (Bio-Rad) was adopted. Serial dilutions of Rabbit gamma globulin (Sigma) were prepared which were representative of the protein solution to be tested. The linear range of the assay for IgG is 1.2 - 25 μg/ml. Pipeted 800 μl of each standard and sample solution into a clean, dry microtube. 200 μl of dye reagent concentrate was added to each tube and vortex. Incubated at room temperature for at least 5 minutes (samples should be incubated at room temperature for no more than 1 hour). Absorbance was measured at 595 nm. Samples were assayed in duplicates.

(2) β-galactosidase Histochemistry

i) Frozen Sections

Tissue was mounted in O.C.T. compound (Lab-Tek Division, Miles). One drop of O.C.T. was added onto the surface of specimen holder. The holder was immersed half-way into liquid nitrogen until the O.C.T. compound began to freeze at the bottom. The holder was removed from the liquid nitrogen and placed on a flat surface. The O.C.T. continued to freeze from bottom to top. A second smaller drop of O.C.T. was added when the unfrozen area had contracted to a ring of about 0.5 cm in diameter. To ensure proper contact between the two blocks, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

second drop of embedding medium was added before the first drop was completely frozen. The second drop, in which the tissue would be embedded had a diameter of approximately 0.5 cm. Before the second drop froze, the tissue was added – either one-day old fly whole body or only head. Late stage pupae or adult flies were transferred to the block using a fine-tip forceps. Usually tissues of interest were embedded at high density in the center of embedment but not in contact with each other. The block was immersed again half-way into liquid nitrogen and allowed it to freeze completely. The block was transferred into the cryostat chamber and equilibrated for approximately 30 min. After the block has been equilibrated to chamber temperature, it was mounted to the microtome and trimmed if necessary. 10 - 15 μm sections were cut and collected on subbed microscope slides. After the sections covered an area of about 18 x 36 mm of the slide, it was placed on a hot plate at about 52°C for 3 min to dry and flatten the sections onto the gelatin matrix of the slide. Now the slides were ready for fixation.

ii) X-gal Staining in Frozen Sections

The frozen sections on slides were fixed with 3% glutaraldehyde in PBS for 15 minutes at room temperature in a coplin jar. Slides were washed three times with PBS, 10 minutes each. After the wash, excess PBS was removed with a Kimwipe gently and the slides were placed on a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

level surface in a moist airtight chamber. A large Petri-dish with two sheets of water-soaked Kimwipes works well for this purpose. 50 μl of prewarmed X-gal staining solution (10 mM $NaPO_4$ (pH 7.2), 150 mM NaCl, 1 mM $MgCl_2$, 3 mM $K_4[FeII(CN)_6]$, 3 mM $K_3[FeIII(CN)_6]$, 0.3% Triton X-100, and 0.2% X-gal (in N,N-dimethyformamide, stored at -20°C); recipe obtained from Dr. Ballinger) was applied onto the dried slides as quickly as possible, put on all the coverslips, and immediately put the sealed chamber in a 37°C incubator. (50 μl suffices a 22 X 40 mm cover slip – adjust accordingly for different sizes). Allow the staining reaction proceed for as long as necessary (from 1 hour to overnight). The slides were examined periodically under a dissection microscope. When the staining appears to be complete, the cover slips were floated off in PBS and washed with 2 changes of PBT (PBS with 0.05% Tween 20), 1 change of 70% ethanol (the slides may be dehydrated through a graded ethanol series; the ethanol helps remove X-gal crystals deposited on the tissue sections during the staining reaction) and 2 changes of PBS sequentially. The slides were mounted in 90% glycerol in PBS. The glycerol mounting gives better morphology for photomicroscopy.

iii) X-gal Staining in Whole-mount Embryos

A protocol from Dr. Leslie Pick (personal communication) was adapted for X-gal staining of whole mount embryos. Aged embryos (10

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- 22 hours) were collected and dechorionated in 3% sodium hypochlorite for 90 seconds. Dechorionated embryos were then fixed in freshly prepared fixation solution (heptane : 50% glutaraldehyde : PBS = 2 : 1 : 0.5) for 10 to 15 minutes at room temperature. After settling, embryos in interface were transfered onto a glass slide in small drops and allowed heptane to evaporate. Embryos were picked up on the corner of a piece of double stick tape and placed on a new glass slide with embryos facing up. Covered with a drop of Ringer's solution, embryos were devitellinized by poking them with a needle. Devitellinized embryos were collected in a microfuge tube and rinsed once with warm staining solution (10 mM $NaPO_4$, 150 mM NaCl, 1 mM $MgCl_2$, 3.05 mM $K_4[Fe(III)(CN)_6]$, 3.5 mM $K_3[Fe(II)(CN)_6]$). The embryos were then stained at 37°C overnight in pre-warmed staining solution with 0.2% X-gal (in N,N-dimethyformamide). Stained embryos were rinsed with 70% ethanol and 100% ethanol once of each and stored in 90% glycerol/1X PBS overnight. The embryos were mounted on a slide between two small coverslips, covered with long coverslip and observed.

iv) β-Galactosidase Histochemistry in Larvae

Larvae from transgenic strains were also stained for β-galactosidase activity to assess *als*-reported expression in the larval CNS. The method is the same as the above for whole-mount embryo staining

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

except no need for the dechorionation step.

(3) Immunohistochemistry – Antibody stainings of frozen sections

Embryo, larvae, and one-day old adult flies from transgenic strains involving two different *–lacZ* fusion genes were immunostained with anti-β-galactosidase antibody to assess *als*-reported expression in the CNS. Fusion gene transformants were tested with polyclonal antiserum raised against ALS.

Frozen sections were prepared as described before, slides were fixed in 2.0% Glutaraldehyde in PBS at 4°C. For *Drosophila* heads, first the proboscis were removed and the head capsule fixed for 60-90 minutes. For other types of tissues, the fixation time can be adjusted according to their relative size and estimated efficiency of filtration. For Glutaraldehyde, depth of penetration is 2-3 mm/hour. The slides were washed in PBS and transfered to 12% sucrose in PBS at 4°C. Sucrose solution was allowed to infiltrate the tissue for 16 hrs. Then the tissue was removed from the sucrose solution and submerged in a drop of O.C.T. compound (Lab-Tek Division, Miles). Allow the tissue to be permeated by the O.C.T compound for 10-30 minutes at room temperature and then embeded in frozen O.C.T. compound using an ethanol-dry ice bath. 10-14 μm frozen sections were made in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cryostat and transfer section ribbons onto slides. Slides were heated briefly at ~40°C on a drying plate for no more than one minute to dry the sections onto the slides. Overheating may adversely affect the antigenicity of the protein of interest and the morphology of the tissue.

The sections were fixed immediately 0.5% glutaraldehyde in PBS for 20 minutes at room temperature and washed three times for 2-3 minutes in PBS. Slides were blocked for 30 minutes in PBSG (0.2% BSA, 1% goat serum, 0.01% saponin in PBS), then washed through several changes of PBS/0.01% saponin.

75-150 μl of the primary antibody was applied at 1:200 ~ 1:50 dilution in PBSG to each slide, cover slips were placed gently on the slides to evenly spread the antibody solution over the tissue sections. Slides in primary antibody were incubated for 30-60 minutes at room temperature or overnight at 4°C in a moist, airtight chamber, subsequently washed through several changes of PBS/0.01% saponin (float off the cover slips in the first wash).

The sections were incubated with secondary antibody at the appropriate dilution in PBSG exactly as described for the primary antibody. For Bio-Rad Laboratories HRP-conjugated goat-anti-rabbit IgG, a 1:200 dilution for 30-60 minutes at room temperature works well.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

For an HRP-conjugated secondary antibody, the slides were again washed through several changes of PBS/0.01% saponin and then incubated in 0.5 – 1.0 ml of staining solution per slide without cover slips or 100-150 μl with cover slips. (0.5 mg/ml diaminobenzidine (DAB), 0.003% H_2O_2, 1.5 mM $CoCl_2$, and 1.5 mM $NiCl_2$. The latter two components were optional. In their absence, a brown product formed, whereas in their presence, a more intense blue-black product was obtained (Since the extent and localization of background staining can vary considerably between intensified and non-intensified preparations, it is often worth trying both methods in parallel).

The staining reaction was monitored under a dissection microscope. It usually takes 5-30 minutes for the reaction to proceed to completion. The reaction was stopped by washing the slides through a few changes of PBS. Slides were mounted under cover slips in 70% glycerol/PBS after air drying without dehydration.

RESULTS

To identify transcriptional control elements regulating the *als* gene expression, transformant lines with two different upstream fragment at 2.9 kb (RB) and 1.25 kb (RE) were initially analyzed. As illustrated in **Figure 11A and Figure 13**, both constructs contain the common region

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

bp of upstream sequence flanking the *als* transcription start site, and the predicted promoter immediately proceed it. The 1.25 kb fragment was analyzed for transcription factor binding sites with computer-assisted programs. Several consensus *cis*-elements has been identified which may play important role in spatial and temporal control of the *als* gene expression. The resulting transgenic flies after microinjection of both fusion constructs were analyzed for expression patterns of reporter gene construct in late embryos, larvae, late pupae, and frozen section of one-day old whole flies and head using X-gal staining and anti-β-gal antibody.

Quantitative analysis of expression

Levels of expression from the transformed *lacZ* fusion genes were determined quantitatively with a simple liquid β-galactosidase assay system that uses CPRG as a substrate. CPRG produces a water-soluble product and is preferable to ONPG because it is reported to be ten times more sensitive (Simon and Lis 1987). For crude fly extracts, CPRG has the additional advantage over ONPG that the product absorbs at a wavelength not significantly interfered with by fly pigments. In CPRG assay of crude fly extracts, the rate of color development was constant over at least four hours of reaction time at 37°C and the relationship between extract concentration and CPRG cleavage was linear over a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

100-fold range of extract concentration (Simon and Lis 1987).

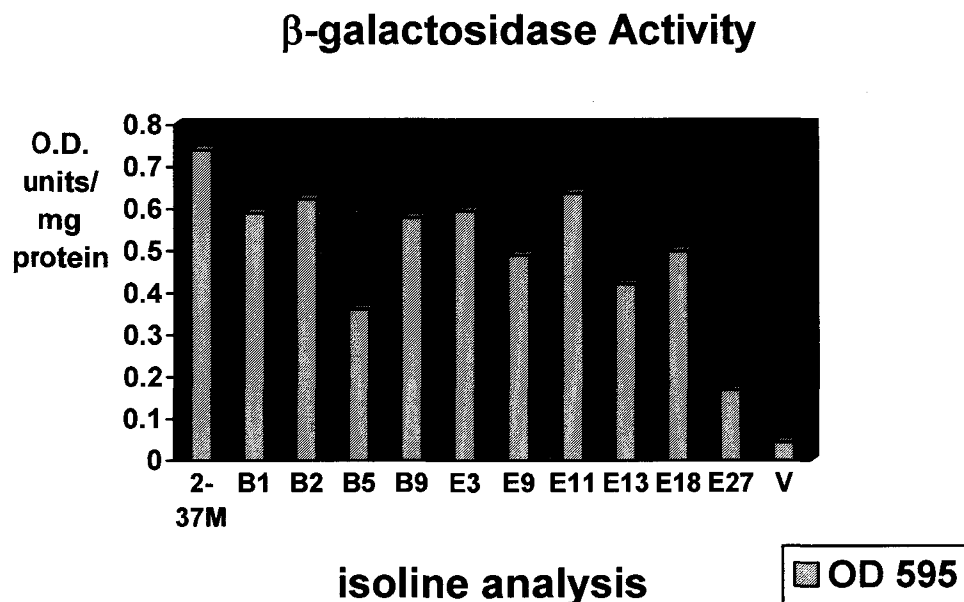
The results of β-galactosidase activity assay were shown in **Figure 15**. The transformant lines of RB and RE constructs were studied. The β-galactosidase activity was expressed in arbitrary units as O.D. units/mg protein. Background due to endogenous fly β-galactosidase activity and light scattering material present in the crude extracts was measured by using extracts prepared in parallel from the nontransformed strain (*yw*). This background was subtracted numerically from β-galactosidase levels in transformant extracts. In addition, the previously establish positive transformant line "2-37M" was selected as positive control. As expected, the transformant lines with the expression vector alone (i.e. V) has only background level of β-galactosidase activity. All of the transformant lines with various 5' flanking regions showed various levels of β-galactosidase activity. The β-galactosidase activities directed by E3, E9, E11, E13, E18, E27, and B1, B2, B5 and B9 were in a range between 0.3 and 0.8 O.D. units/mg protein. These results correspond to the results of X-gal staining.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 15. β-galactosidase Activity Assay.

Schematic representation of β-galactosidase levels in transformant extracts. The activity is expressed in arbitrary units as O.D. units/mg protein on Y-axis. "2-37M" is the positive control; "V" is the negative control. The bars indicate the average of all available isolines of each construct. The background was measured by using extracts prepared in parallel from the nontransformed strain (*yw*) subtracted numerically from β-galactosidase levels in transformant extracts.



Expression specificity in Transformed Lines

The expression patterns of one-day-old transgenic flies examined by X-gal staining were summarized in **Table 7**, along with the chromosomal location of P-element insertion of RB and RE constructs, respectively. The X-gal staining pattern observed for all positively expressing lines was qualitatively similar. No X-gal staining were detected in control transformant lines of *yw*. Most of RB and RE transformant lines (4 RB and 6 RE isolines of each) demonstrated similar expression patterns. Some transformed lines showed no significant X-gal staining (data not shown), possibly indicating insertion into a nonfunctional chromosome position.

Table 7. Summary of X-gal Staining in Isolines of Transgenic flies.

Constructs	Isoline Name	Chromosome Location	X-gal Staining (1 day-old adults)
2.9 kb	B1	II	+
	B2	II	+
	B5	III	+
	B9	II	+
1.2 kb	E3	II	+
	E9	II	+
	E11	II	+
	E13	II	+
	E18	III	+
	E27	II	-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Vector only	V	II	-

"+", positive staining in CNS

"-", no staining detectable

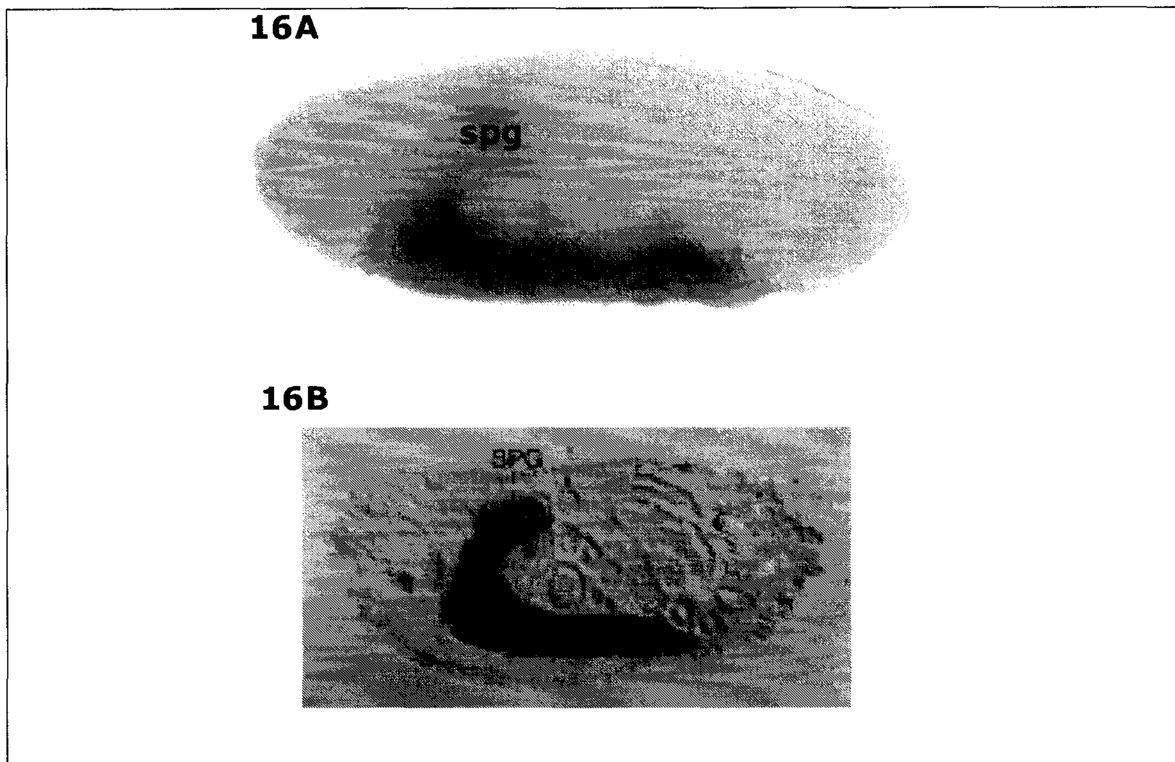
In one-day-old adult horizontal head section, X-gal staining is detected in the cortical regions of the midbrain, the cellular cortex of the optic lobe area and the cortical regions of thoracic ganglia (see **Figure 17A, 18A**). The intensity of staining varies in different CNS regions. The most intensely stained structures observed include the second and third order ganglia of the visual system, the optic lobes, large areas of the ventral-lateral and posterior protocerebrum, and the thoracic ganglion.

The patterns of β-galactosidase expression is also compared to that of the endogenous *als* gene expression. As shown in **Figure 16A**, the staining of whole-mount preparation of late embryo reveals intense X-gal staining pattern in subesophageal ganglion (sbg), supraesophageal ganglion (spg) and ventral cord (vc) which coincide with localization of *als* transcripts detected by whole mount *in situ* hybridization (**Figure 16B**, adopted from Sawruk and Schmitt 1990).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 16. β-galactosidase distribution pattern in embryos.

The X-gal staining patterns in whole mount preparation of late embryos (16A) in comparison with in-situ hybridization diction of *als* expression (16B) (Sawruk *et al.* 1995). vc, ventral cord; spg, supraesophageal ganglion; sbg, subesophageal ganglion.



As shown in **Figure 17A, 18B, 18C and 19B**, one-day old adult head section reveals a uniform X-gal staining in the cortical regions surrounding the lobula, lobular plate and medulla, while the cellular cortex of the lamina contains no X-gal staining. This agrees with the immunohistochemical data that the distal structure in lamina which is strongly immunoreactive with anti-ARD, but not with anti-ALS antisera (Schuster and Gundelfinger 1993). The optic lobe contains four

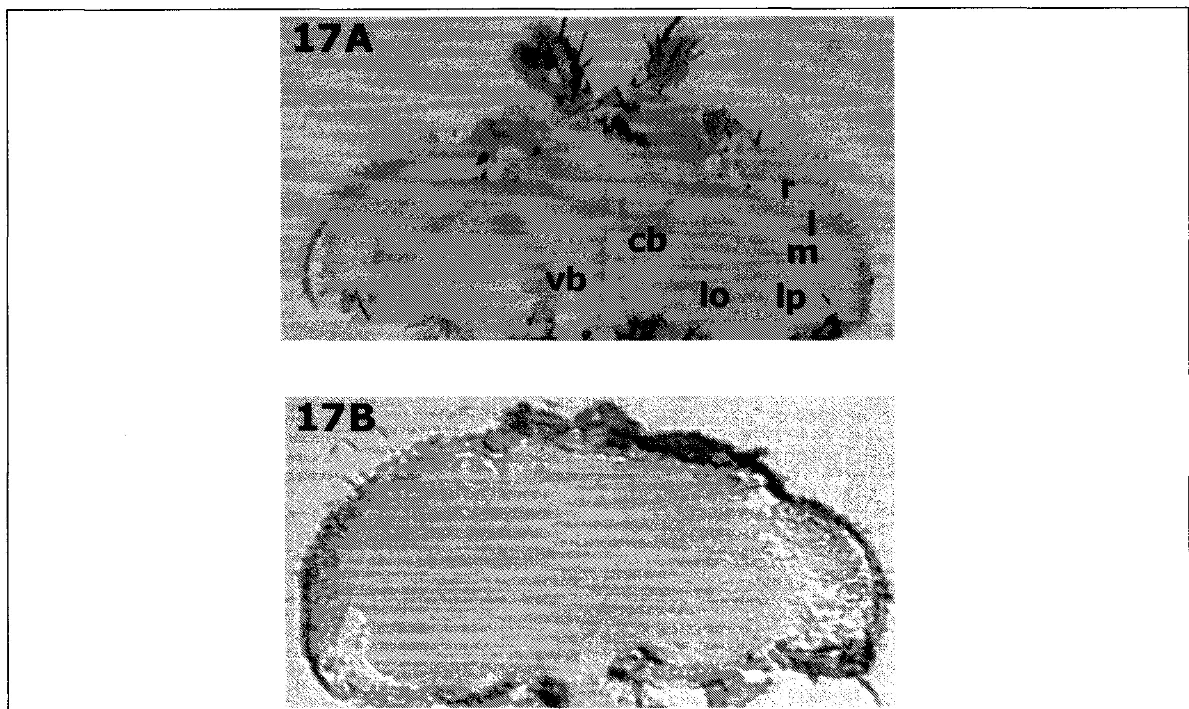
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

neuropil regions, i.e., from distal to proximal: the lamina, the medulla, the lobula, and the lobular plate (Flybrain, Interactive database (http://flybrain.neurobio. arizona.edu/). They are involved in the processing of visual information and its transmission to the central brain. All four neuropiles are organized in columns that are interconnected by perpendicularly running elements (Fischbach and Dittrich 1989). The lamina intrinsic (amacrine) cells and medulla neurons make their synapses in the lamina neuropile. The X-gal staining patterns observed in most of the transformant lines indicate the presence of moderate amounts of ALS-containing nAChRs in the entire synaptic region except for the lamina, where ARD might assemble with other a-like subunits to form a different nAChR subtypes. There are several evidence supporting this hypothesis. So far, the distribution of ALS, ARD, Dα2/SAD and Dα3 has been studied in the adult

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 17. Distribution of β-galactosidase activity in one-day old head section of RE-3 transformed isolines.

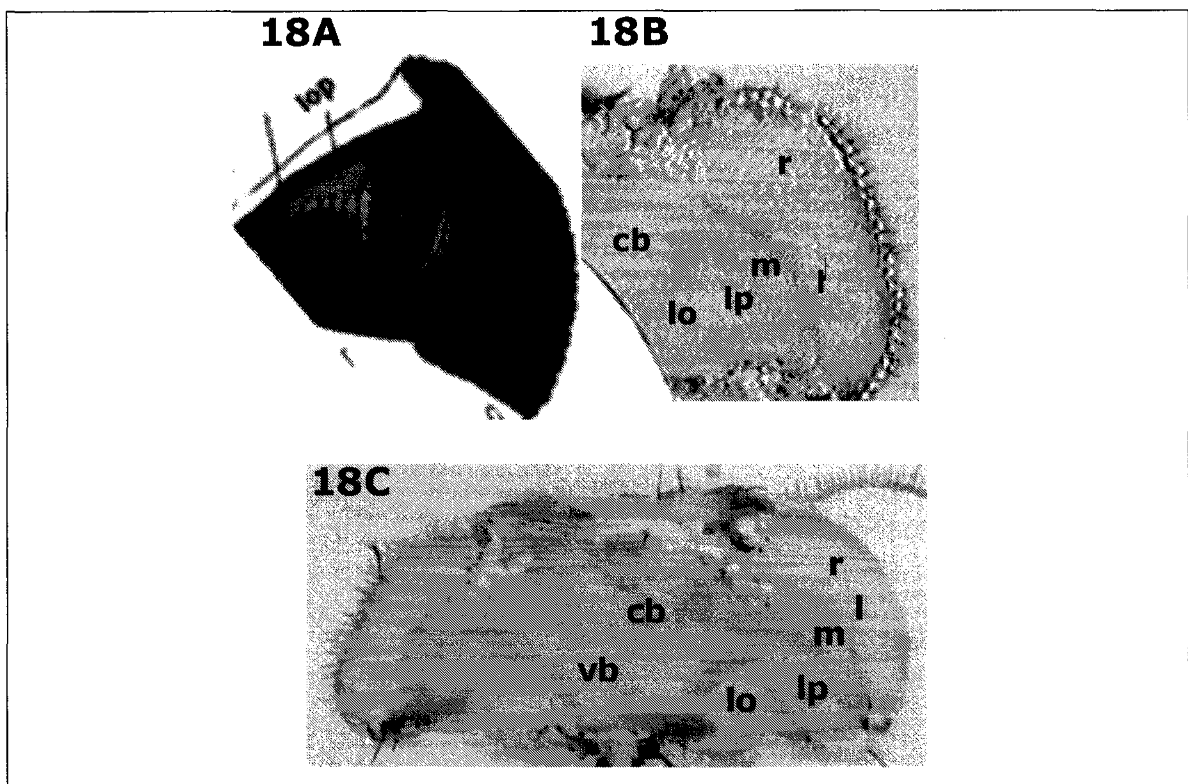
The X-gal staining pattern in frozen section of one-day old RE-3 adult head are shown in 16A, as compared to 16B, head section of *yw* strain as negative control. Abbreviations: cb, central brain; r, retina; m, medulla; lp, lobular plate; lo, lobular; v, ventral body; l, lamina



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 18. Distribution of β-galactosidase activity in one-day old head section of RB-2 transformed isolines.

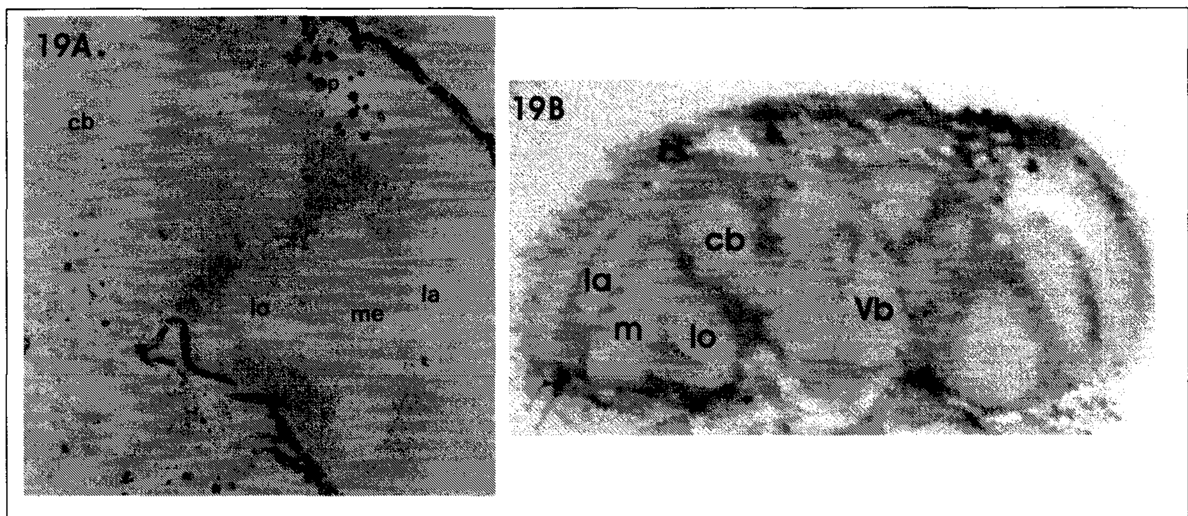
The X-gal staining pattern in frozen section of one-day old RB-2 adult head are shown in 17B (enlarged partial head), 17C (head section), as compared to 17A, a schematic drawing showing the optic lobe and the neuropile regions at its vicinity. Abbreviations: cb, central brain; r, retina; l, lamina; m, medulla; lp, lobular plate; lo, lobular; vb, ventral body;



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 19. Distribution of β-galactosidase activity in RE-11 transformed isolines

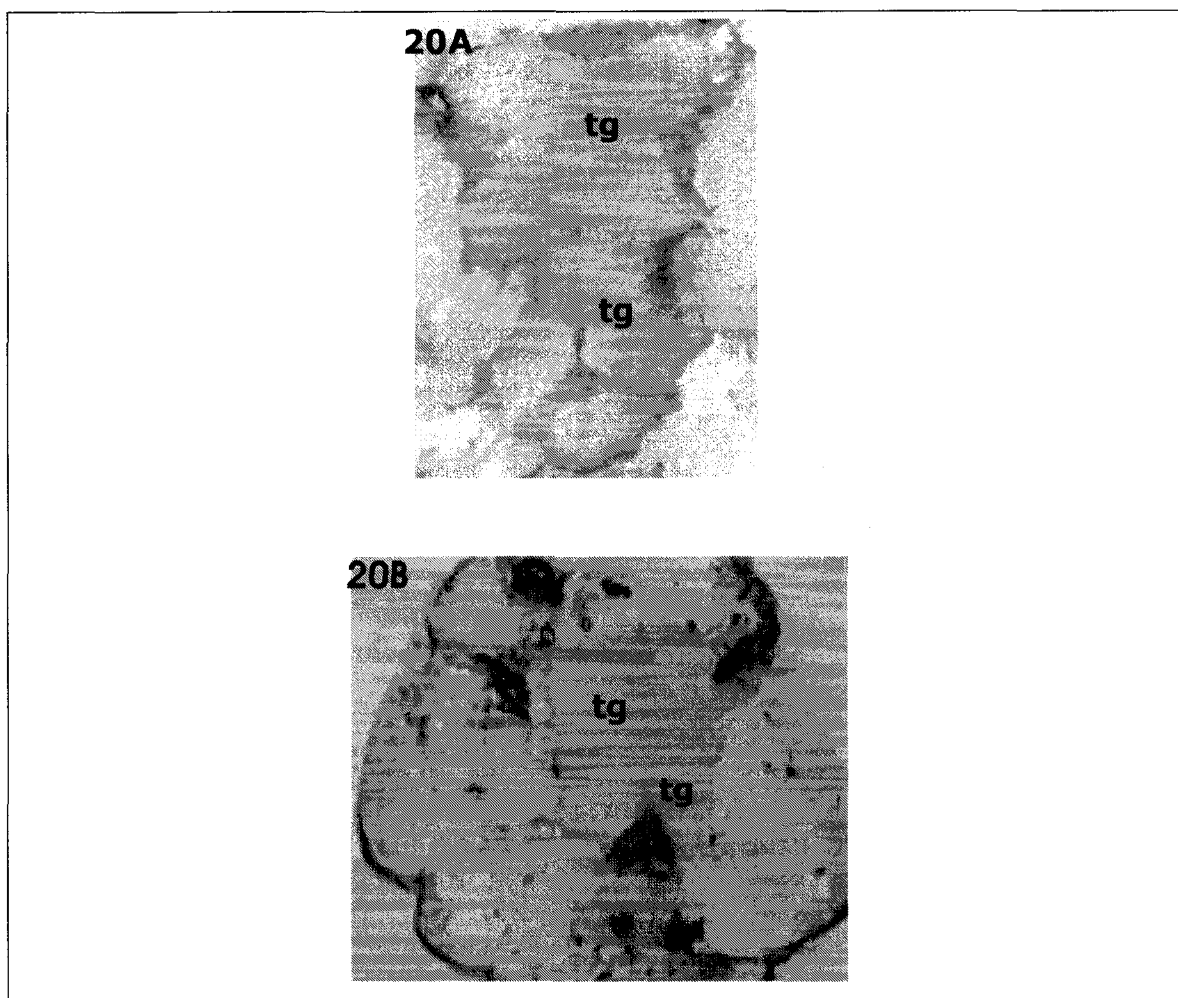
The X-gal staining patterns in frozen section of one-day old adult brain (19B) as compared to *ard* RNA transcript detected by *in situ* hybridization (Hermans-Borgmeyer and Gundelfinger *et al.* 1989) (19A). cb, central brain; la, lamina; m, medulla; lo, lobular; vb, ventral body.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 20. Distribution of β-galactosidase activity in thoracic region in RE-13 transformed isolines

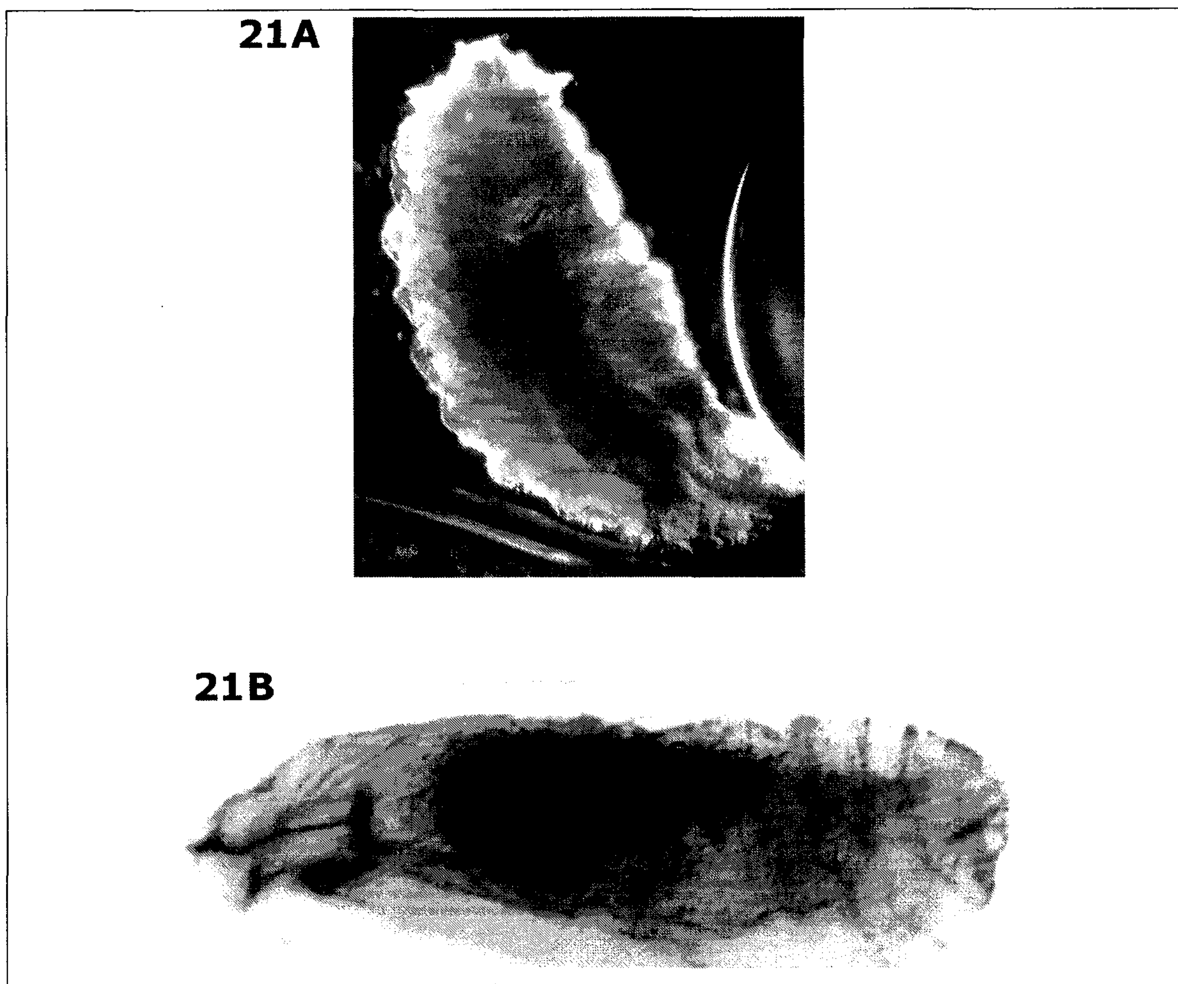
The X-gal staining patterns in frozen section of one-day old adult thoracic ganglionic center in an enlarged view (A) and in the thoracic part of a frozen section (B). tg, thoracic ganglion



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 21. Distribution of β-galactosidase activity in larvae of RE-18 transformed isolines.

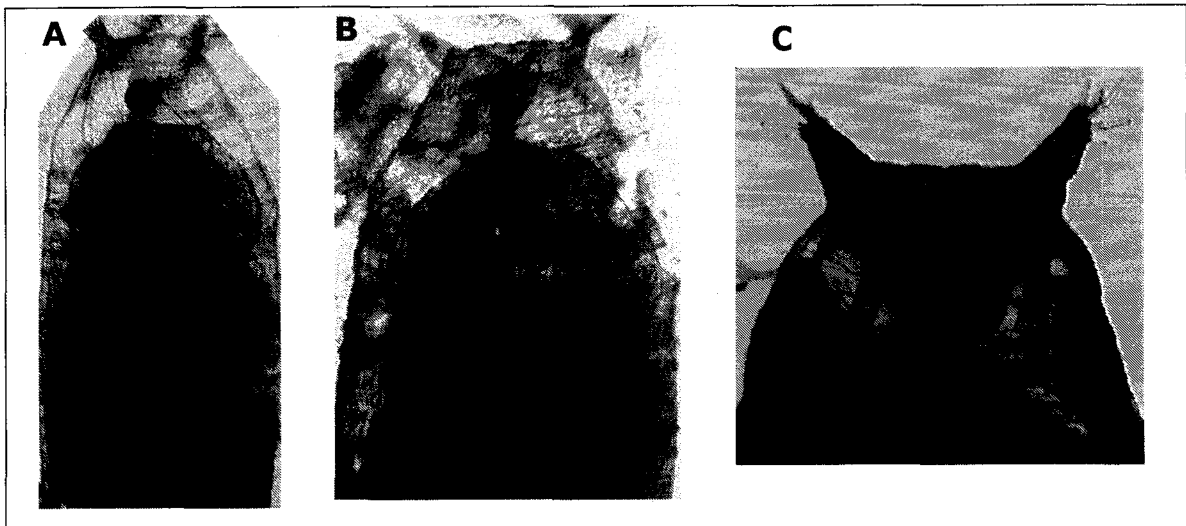
The X-gal staining patterns in frozen section of 1st instar larva (A) and in 2nd instar larva (picture was taken with its ventral side up) (B). vns, ventral nervous system.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 22. Distribution of β-galactosidase activity in prepupae of RB-9 transformed isolines.

The X-gal staining patterns in pre-pupae of control V-line (A), and transformant line RB-9 (B), an enlarged version of head region (C).



brain (Schuster and Gundelfinger 1993; Jonas and Gundelfinger 1994). While all four subunits are essentially co-distributed in most synaptic neurophil regions, only ARD and Dα3, but not ALS and Dα2, immunoreactivity is detected in globular cholinergic structures in the distal lamina. The identity of these structures is presently unclear (Chamaon *et al.* 2000). Second, ARD and Dα3 are proved to be integral components of the same receptor complex by co-immunoprecipitation of both protein with anti-ARD and anti-Da3 antibodies respectively, as well as by immunohistochemical studies showing co-distribution of ARD and Dα3 in synaptic neuropil regions of the optic lobe.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It would be interesting to elucidate the complete quaternary structure of insect nAChRs at the ultrastructural level to decide whether any of those nAChR subunits sharing similar distribution in the adult synaptic neuropil regions actually are present at the same synapse.

The patterns of β-galactosidase expression is determined by the extent of the specific putative regulatory DNA to drive *lacZ* expression. Expression specificity of the fusion constructs from transformant isolines of the 1.25 kb fusion gene containing 793 bp of upstream sequence was shown to be sufficient to drive *als* expression in CNS through different development stages, whereas the distinct staining pattern with X-gal identifying neurons in the CNS was qualitatively similar to that of the endogenous *als* gene expression by immunohistochemistry and in situ hybridization reported by other research groups. We have also exploited expression of fusion proteins in neuronal population by immunohischemistry using a monoclonal antibody against β-galactosidase (Promega Biotech). However no satisfactory result has been achieved so far due to some technique problems.

Based on our *in vivo* analysis of the transformed flies, we conclude that the 793 bp 5'-flanking sequence contains the essential *cis*-acting regulatory elements responsible for spatial and temporal *als* gene expression.

The coincidence of *als* gene expression with periods of formation of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the larval and adult nervous system in *Drosophila* raises the question whether the *als* gene product is required for proper differentiation or whether its role is restricted to signal transmission in the functional CNS. Genetic studies using mutant defective in vertebrate neuronal nAChRs suggested involvement of development and maintenance of the CNS. Production of mutants of the *als* gene and a subsequent genetic analysis will be necessary to understand the precise role of the ALS protein in *Drosophila* development.

DISCUSSION

The Isolation and characterization of the gene coding for the *Drosophila* nAChR *als* subunit gene has been described previously (Bossy *et al.* 1988). Distribution of the *als* transcript was detected in all developmental stages from late embryogenesis to adulthood, but not during the first 2 hours of development. The *als* transcript seems to reach its maximum of expression at late-embryogenesis and to stay relatively stable up to early pupal stage, followed by 2-fold decrease in late pupal stage and 3-fold decrease at adulthood (7 days after eclsion) (Bossy *et al.* 1988). Spatial expression pattern of the *als* gene was detectable in the developing central nervous system but was not in any other tissues. Our *in vivo* data is in accordance with the temporal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

distribution, the *als* transcript was detected at a very high level in the CNS during early embryogenesis, in larvae, and early pupae stage. The main goal of this study is to define the molecular structure which lay down the fundamental blueprint for temporal and spatial regulation of the *als* gene encoding the ligand-binding α subunit of nAChRs in central nervous system.

The primer extension analysis and RACE-PCR suggested one transcription start of the *als* gene, which is located 29 kb upstream from the translation start site. Common start sites were identified in both late embryonic stage and one-day-old adult, indicating a conservation of transcription initiation during development in embryonic as well as one-day-old adult stages. *In vivo* study was performed with fusion constructs that contain upstream fragments (RB -2.9 kb) and RE-1.25 kb) including transcription start site and part of the first non-coding exon. Analysis of transformed lines after P-element mediated transformation indicated the 1.25 kb construct with 793 upstream sequence contains sufficient promoter activity flanking the transcription start sites, i.e. 793 bp upstream sequence around the transcription start site is shown to drive tissue-specific expression during the development from early embryogenesis, larval, early pupal, late pupal to young adult stages. The temporal and spatial expression pattern of *lacZ* fusion gene is similar to that of the endogenous *als* gene expression. This confirms our

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mapping of the transcription start sites and our assumption that the regulatory sequence resides in vicinity of the transcription start site.

Given the fact that expression patterns of RE and RB constructs do not appear to differ, it may also be possible that the critical region is the promoter per se, independently of its type. Now with the essential 793 bp sequence required for correct expression of the *als* gene, the transcriptional regulation of the *als* gene is complex and needs further investigation.

The 793 bp *als* upstream region was further compared to the 5' flanking regions of other genes encoding proteins involved in neurotransmission in the *Drosophila* nervous system, such as acetylcholinesterase (*ace*) (Fournier *et al.,* 1989), choline acetyltransferase *(ChAT*) (Carbini *et al*. 1990). No significant sequence similarity has been found between the *als* regulatory region and the promoter regions of those genes.

In summary, by using P-element-mediated transformation, we were able to identify the upstream regulatory sequences which confer neuronal specificity on the *Drosophila* nAChR *als* subunit gene. During development (including late embryonic, late pupae and one-day-old adult stages), the 793 bp upstream sequence within a 1.25 kb fragment of the *als* gene directed reporter gene expression to specific subsets of neurons in CNS of transformed flies. The expression pattern is similar

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to that reported for the *als* transcript detected in embryo by *in situ* hybridization (Sawruk *et al.* 1990) and the *ard* transcript detected in adult CNS (Hermans-Borgmeyer *et al.* 1989). Our results suggest that the 1.25 kb DNA fragment around the transcription start site contains at least several *cis*-acting regulatory elements and is sufficient to confer spatial and temporal transcriptional specificity. However, it cannot be excluded that additional *cis*-acting elements further upstream of the 5' flanking DNA contribute to the expression pattern of the *als* subunit gene. In addition, some possible sequence motifs of known *cis*-regulatory elements have been identified within the 1.25 kb fragment while the functional significance of all potential *cis*-elements remains to be established.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

REFERENCES

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.

Anand, R., Conroy, W.G., Schoepfer, R., Whiting, P. and Lindstrom, J. (1991). "Neuronal nicotinic acetylcholine receptors expressed in Xenopus oocytes have a pentameric quaternary structure." J Biol Chem **266**: 11192-11198.

Antil-Delbeke, S., Gaillard, C., Tamiya, T., Corringer, P.J., Changeux, J.P., Servent, D., Menez, A. (2000). "Molecular determinants by which a long chain toxin from snake venom interacts with the neuronal alpha 7 nicotinic acetylcholine receptor." J Biol Chem **275**(38): 29594-601.

Arkhipova, I.R. (1995). "Promoter elements in *Drosophila melanogaster* revealed by sequence analysis." Genetics **139**(3): 1359-69.

Ashburner, M. (1989). "*Drosophila*: A Laboratory Manual." Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1988). "Current Protocols in Molecular Biology." Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., New York.

Ballivet, M., Alliod, C., Bertrand, S., Bertrand, D. (1996). "Nicotinic acetylcholine receptors in the nematode *Caenorhabditis elegans*." J Mol Biol **258**(2): 261-269.

Baumann, A., Jonas, P. and Gundelfinger, E.D. (1990). "Sequence of Dα2, a novel a-like subunit of *Drosophila* nicotinic acetylcholine receptors." Nucleic Acids Res **18**: 3640.

Bausenwein, B., Dittrich, A.P.M., and Fischbach, K.F. (1992). "The optic lobe of *Drosophila melanogaster* II. Sorting of retinotopic pathways in the medulla." Cell Tissue Res **267**: 17-28.

Belluardo, N., G. Mudo, M. Blum, G. Amato and K. Fuxe (2000). "Neurotrophic effects of central nicotinic receptor activation Central nicotinic receptors, neurotrophic factors and neuroprotection." J Neural Transm Suppl **113**(60): 227-45.

Belluardo, N., G. Mudo, M. Blum and K. Fuxe (2000). "Central nicotinic receptors,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

neurotrophic factors and neuroprotection." Behav Brain Res **113**(1-2): 21-34.

Benson, M. and V. Pirrotta (1987). "The product of the *Drosophila* zeste gene binds to specific DNA sequences in white and Ubx." Embo J **6**(5): 1387-92.

Benyajati, C. L.Mueller, N.Xu, M.Pappano, J.Gao, M.Mosammaparast, D.Conklin, H.Granok, C.Craig, S.Elgin (1997). "Multiple isoforms of GAGA factor, a critical component fo chromatin structure." Nucleic Acids Res **25**: 3345-3353.

Bertrand, D., Ballivet, M., Gomez, M., Bertrand, S., Phannavong, B., Gundelfinger, E.D. (1994). "Physiological properties of neuronal nicotinic receptors reconstituted from the vertebrate beta 2 subunit and *Drosophila* alpha subunits." Eur. J. Neurosci. **6**(5): 869-875.

Bhat, K., Farkas, G., Karch, F., Gyurkovics, H., Gausz, J. and Schedl, P. (1996). "The GAGA factor is required in the early *Drosophila* embryo not only for transcriptional regulation but also for nuclear division." Development **122**: 1113-1124.

Bickel, S. and V. Pirrotta (1990). "Self-association of the *Drosophila* zeste protein is responsible for transvection effects." Embo J **9**(9): 2959-67.

Biggin, M. D., S. Bickel, M. Benson, V. Pirrotta and R. Tjian (1988). "Zeste encodes a sequence-specific transcription factor that activates the Ultrabithorax promoter in vitro." Cell **53**(5): 713-22.

Bossy, B., Ballivet, M., Spierer, P. (1988). "Conservation of neural nicotinic acetylcholine receptors from *Drosophila* to vertebrate central nervous system." EMBO J. **7**(3): 611-618.

Brumwell, C.L., Johnson, J.L., Jacob, M.H. (2002). "Extrasynaptic α7-nicotinic acetylcholine receptor expression in developing neurons is regulated by inputs, targets, and activity." The Journal of Neuroscience **22**(18): 8101-8109.

Burns, C.M., H. Chu, S.M. Rueter, L.K. Hutchinson, H. Canton, E. Sanders-Bush and R.B. Emeson (1997). "Regulation of serotonin-2C receptor G-protein coupling by RNA editing." Nature **387**(6630): 303-8.

Caceres, J. F. and A. R. Kornblihtt (2002). "Alternative splicing: multiple control mechanisms and involvement in human disease." Trends Genet **18**(4): 186-93.

Camaon, K., Schulz, R., Smalla, K.H., Seidel, B., Gundelfinger, E.D. (2000). "Neuronal nicotinic acetylcholine receptors of *Drosophila melanogaster*: the alpha-subunit dalpha3 and the beta-type subunit ARD co-assemble within the same receptor complex." FEBS Lett. **482**(3): 189-192.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Camaon, K., Smalla, K.H., Thomas, U., Gundelfinger, E.D. (2002). "Nicotinic acetylcholine receptors of *Drosophila*: three subunits encoded by genomically linked genes can co-assemble into the same receptor complex." J Neurochem **80**(1): 149-157.

Campos-Caro, A., Smillie, F.I., Dominguez del Toro, E., Rovira, J.C., Vicente-Agullo, F., Chapuli, J., Juiz, J.M., Sala, S., Sala, F., Ballesta, J.J. and Criado, M. (1997). "Neuronal nicotinic acetylcholine receptors on bovine chromaffin cells: cloning, expression, and genomic organization of receptor subunits." J. Neurochem **68**(2): 488-497.

Campos-Ortega, J. A. and Y. N. Jan (1991). "Genetic and molecular bases of neurogenesis in *Drosophila melanogaster*." Annu Rev Neurosci **14**: 399-420.

Carbini, L. A., V. J. Munoz Maines and P. M. Salvaterra (1990). "Developmental expression of choline acetyltransferase mRNA in *Drosophila*." Neurochem Res **15**(11): 1089-96.

Cartier, G. E., Yoshikami, D., Gray, W.R., Luo, S., Olivera, B.M., McIntosh, J.M. (1996). "A new alpha-conotoxin which targets alpha3beta2 nicotinic acetylcholine receptors." J Biol Chem **271**(13): 7522-8.

Chang, K. T., Berg, D.K. (1999). "Nicotinic acetylcholine receptors containing alpha7 subunits are required for reliable synaptic transmission in situ." J Neurosci. **19**(10): 3701-10.

Changeux, J. P., Bertrand, D., Corringer, P.J., Dehaene, S., Edelstein, S., Lena, C., Le Novere, N., Marubio, L., Picciotto, M., Zoli, M. (1998). "Brain nicotinic receptros: structure and regulation, role in learning and reinforcement." Brain Res Brain Res Rev. **26**(2-3): 198-216.

Chen, D., Patrick, J.W. (1997). "The alpha-bungarotoxin-binding nicotinic acetylcholine receptor from rat brain contains only teh alpha7 subunit." J Biol Chem **272**: 24024-9.

Cockcroft, V., Ortells, M., Lunt, G. (1995). "Ligands, receptor models, and evolution." Ann N Y Acad Sci. **757**: 40-47.

Conroy, W. G., Wernallis, A.B., and Berg, D.K. (1992). "The α5 gene product assembles with multiple acetylcholine receptor subunits to form distinctive receptor subtypes in brain." Neuron **9**: 679-691.

Conroy, W. G. and D. K. Berg (1995). "Neurons can maintain multiple classes of nicotinic acetylcholine receptors distinguished by different subunit compositions." J Biol Chem **270**(9): 4424-31.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Cook, L., Ho, L.W., Taylor, A.E., Brayne,C. and Rubinsztein, D.C., *et al.* (2004). "Candidate gene association studies of the alpha 4 (CHRNA4) and beta 2 (CHRNB2) neuronal nicotinic acetylcholine receptor subunit genes in Alzheimer's disease." Neurosci Lett. **358**(2): 142-146.

Cooper, E., Couturier, S., and Ballivet, M. (1991). "Pentameric structure and subunit stoichiometry of a neuronal nicotinic acetylcholine receptor." Nature **350**: 235-238.

Corringer, P.-J., Le Novere, N. and Changeux, J-P. (2000). "Nicotine receptors at the amino acid level." Annu. Rev. Pharmacol. Toxicol. **40**: 431-458.

Corringer, P. J., Bertrand, S., Bohler, S., Edelstein, S.J., Changeux, J.P., Bertrand, D. (1998). "Critical elements determining diversity in agonist binding and desensitization of neuronal nicotinic acetylcholine receptors." J Neurosci. **18**(2): 648-657.

Corriveau, R. A. and D. K. Berg (1993). "Coexpression of multiple acetylcholine receptor genes in neurons: quantification of transcripts during development." J Neurosci **13**(6): 2662-71.

Couturier, S., L. Erkman, S. Valera, D. Rungger, S. Bertrand, J. Boulter, M. Ballivet and D. Bertrand (1990). "Alpha 5, alpha 3, and non-alpha 3. Three clustered avian genes encoding neuronal nicotinic acetylcholine receptor-related subunits." J Biol Chem **265**(29): 17560-7.

Cuevas, J., Roth, A.L., Berg, D.K. (2000). "Two distinct classes of functional α7 containing nicotinic receptor on rat superior cervical ganglion neurons." J Physiol. **525**(Pt 3): 735-746.

Dani, J. A. (2001). "Nicotinic receptor activity alters synaptic plasticity." Scientific World Journal **1**: 393-395.

De Fiebre, N. C., De Fiebre, C.M. (2003). "α7 nicotinic acetylcholine receptor-mediated protection against ethanol-induced neurotoxicity." Alcohol **31**: 149-153.

Deneris, E. S., Connolly, J., Rogers, S.W. and Duvoisin, R. (1991). "Pharmacological and functional diversity of neuronal nicotinic acetylcholine receptors." Trends in Pharmacological Sciences **12**: 34-40.

Devillers-Thiery, A., Giraudat, J., Bentaboulet, M., and Changeux, J.P. (1983). "Complete mRNA coding sequence of the acetylcholine binding α-subunit of *Torpedo marmorata* acetylcholine receptor: a model for the transmembrane organization of the polypeptide chain." Proc. Natl. Acad. Sci. **80**(7): 2067-2071.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Dougherty, J. J., J. Wu, R. A. Nichols, S. V. Nayak and J. M. McIntosh (2003). "Beta-amyloid regulation of presynaptic nicotinic receptors in rat hippocampus and neocortex Ca(2+) changes induced by different presynaptic nicotinic receptors in separate populations of individual striatal nerve terminals." J Neurosci **23**(17): 6740-7.

Drago, J., McColl, C.D., Horne, M.K., Finkelstein, D.I. and Ross, S.A. (2003). "Neuronal nicotinic receptors: insights gained from gene knockout and knockin mutant mice." Cellular and Molecular Life Sciences **60**: 1267-1280.

Drisdel, R. C., Green, W.N. (2000). "Neuronal alpha-bungarotoxin receptors are alpha7 subunit homomers." J. Neurosci. **20**(1): 133-139.

Du, Q., Melnikova, I.N., Gardner, P.D. (1998). "Differential effects of heterogeneous nuclear ribonucleoprotein K on Sp1- and Sp3-mediated transcriptional activation of a neuronal nicotinic acetylcholine receptor promoter." J Biol Chem **273**(31): 19877-83.

Dynan, W. S. and R. Tjian (1983). "The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter." Cell **35**(1): 79-87.

Elgoyhen, A. B., Johnson, D.S., Boulter, J., Vetter, D.E., Heinemann, S. (1994). "Alpha9: an acetylcholine receptor with novel pharmacological properties expressed in rat cochlear hair cells." Cell **79**(4): 705-715.

Elgoyhen, A. B., Vetter, D.E., Katz E., Rothlin, C.V., Heinemann, S.F., Boulter, J. (2001). "Alpha10: a determinant of nicotinic cholinergic receptor function in mammalian vestibular and cochlear mechanosensory hair cells." Proc. Natl. Acad. Sci. **98**(6): 3501-6.

Faraone, S. V., Su, J., Taylor, L., Wilcox, M., Van Eerdewegh, P., Tsuang, M.T. (2004). "A novel permutation testing method implicates sixteen nicotinic acetylcholine receptor genes as risk factors for smoking in schizophrenia families." Hum Hered. **57**(2): 59-68.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., and Karch, F. (1994). "The *Trithorax-like* gene encodes the *Drosophila* GAGA factor." Nature **371**: 806-808.

Ferrari, R., P. Pedrazzi, S. Algeri, L. F. Agnati and M. Zoli (1999). "Subunit and region-specific decreases in nicotinic acetylcholine receptor mRNA in the aged rat brain." Neurobiol Aging **20**(1): 37-46.

Fischbach, K. F., and Dittrich, A.P.M. (1989). "The optic lobe of *Drosophila melanogaster*. I.A Golgi analysis of wild-type structure." Cell Tissue Res **258**: 441-475.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fournier, D., F. Karch, J. M. Bride, L. M. Hall, J. B. Berge and P. Spierer (1989). "*Drosophila melanogaster* acetylcholinesterase gene. Structure, evolution and mutations." J Mol Biol **210**(1): 15-22.

Freedman, R., Hall, M., Adler, L.E., Leonard, S. (1995). "Evidence in postmortem brain tissue for decreased numbers of hippocampal nicotinic receptors in schizophrenia." Biol Psychiatry **38**(1): 22-33.

Frengen, E., D. Weichenhan, B. Zhao, K. Osoegawa, M. van Geel and P. J. de Jong (1999). "A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites." Genomics **58**(3): 250-3.

Fucile, S., Matter, J.M., Erkman, L., Ragozzino, D., Barabino, B., Grassi, F., Alema, S., Ballivet, M. and Eusebi, F. (1998). "The neuronal alpha6 subunit forms functional heteromeric acetylcholine receptors in human transfected cells." Eur. J. Neurosci. **10**(1): 172-178.

Fucile, S., Renzi, M., Lax, P., Eusebi, F. (2003). "Fractional Ca^{2+} current through human neuronal α7 nicotinic acetylcholine receptors." Cell Calcium **34**: 205-209.

Garza, D., J. W. Ajioka, D. T. Burke and D. L. Hartl (1989). "Mapping the Drosophila genome with yeast artificial chromosomes." Science **246**(4930): 641-6.

Gault, J., Hopkins, J., Berger, R., Drebing, C., Logel, J., Walton, C., Short, M., Vianzon, R., Olincy, A., Ross, R.G., Adler, L.E., Freedman, R., Leonard, S. (2003). "Comparison of polymorphisms in the alpha7 nicotinic receptor gene and its partial duplication in schizophrenic and control subjects." Am J Med Genet. **123B**(1): 39-49.

Gojobori, k. T. a. T. (1998). "Evolution of nicotinic acetylcholine receptor subunits." Mol. Biol. Evol **15**(5): 518-527.

Goodman, C.S., Spitzer, N.C. (1979). "Embryonic development of identified neurones: differentiation from neuroblast to neurone." Nature **280**(5719): 208-214.

Gorczyca, M. G., Budnik, V., White, K., Wu, C.F. (1991). "Dual muscarinic and nicotinic action on a motor program in *Drosophila*." J. Neurobiol. **22**(4): 391-404.

Grauso,M., Reenan, R.A., Culetto, E., Sattelle, D.B. (2002). "Novel putative nicotinic acetylcholine receptor subunit genes, Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster identify a new and highly conserved target of adenosine deaminase acting on RNA-mediated A-to-I pre-mRNA editing." Genetics **160**(4):

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1519-1533.

Gray, R., Rajan, A.S., Radcliffe, K.A., Yakehiro, M., Dani, J.A. (1996). "Hippocampal synaptic transmission enhanced by low concentration of nicotine." Nature **383**(6602): 713-716.

Greger, I. H., Khatri, L. and Ziff, E.B. (2002). "RNA editing at arg607 controls AMPA receptor exit from the endoplasmic reticulum." Neuron: 759-772.

Gundelfinger, E.D., Hess, N. (1992). "Nicotinic acetylcholine receptors of the central nervous system of *Drosophila*." Biochim Biophys Acta. **1137**(3): 299-308.

Gundelfinger, E.D. (1995). "Evolution and desensitization of LGIC receptors." Trends Neurosci **18**(7): 297; author reply 298-9.

Haenlin, M., H. Steller, V. Pirrotta and E. Mohier (1985). "A 43 kilobase cosmid P transposon rescues the fs(1)K10 morphogenetic locus and three adjacent Drosophila developmental mutants." Cell **40**(4): 827-37.

Halevi, S., J.McKay, M. Palfreyman, L. Yassin, M.Eshel, E. Jorgensen and M. Treinin (2002). "The C. elegans *ric*-3 gene is required for maturation of nicotinic acetylcholine receptors." EMBO J **21**(5): 1012-20.

Hanrahan, C.J., M.J. Palladino, B.Ganetzky and R.A. Reenan (2000). "RNA editing of the *Drosophila* para Na(+) channel transcript. Evolutionary conservation and developmental regulation." Genetics **155**(3): 1149-60.

Harborne, A. J. and M. E. Smith (1988). "Effect of soluble factors from nerve and muscle on alpha-bungarotoxin binding sites in isolated sarcolemmal membranes of the rat." Br J Pharmacol **94**(3): 968-76.

Hermans-Borgmeyer, I., Zopf, D., Ryseck, R-P., Hovemann, B., Betz, H. and Gundelfinger, E.D. (1986). "Primary structure of a developmentally regulated nicotinic acetylcholine receptor protein from *Drosophila*." EMBO J. **5**(7): 1503-1508.

Hermans-Borgmeyer, I., Sabine Hoffmeister, Erich Sawruk, Heinrich Betz, Bertram Schmitt, and Eckart D. Gundelfinger (1989). "Neuronal acetylcholine receptors in *Drosophila*: mature and immature transcripts of the *ard* gene in the developing central nervous system." Neuron **2**: 1147-1156.

Hernandez, M.C., Erkman, L., Matter-Sadzinski, L., Roztocil, T., Ballivet,M. and Matter, J.M. (1995). "Characterization of the nicotinic acetylcholine receptor beta 3 gene." J. Biol. Chem **270**(7): 3224-3233.

Hess,N., Merz,B., Gundelfinger, E.D. (1994). "Acetylcholine receptors of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Drosophila brain: a 900 bp promoter fragment contains the essential information for specific expression of the *ard* gene *in vivo*." FEBS Lett. **346**(2-3): 135-140.

Hiel, H., Luebke, A.E. and Fuchs, P.A. (2000). "Cloning and expression of the alpha9 nicotinic acetylcholine receptor subunit in cochlear hair cells of the chick." Brain Res. **858**(1): 215-225.

Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R. and Seeburg, P.H. (2000). "Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2." Nature: 78-81.

Hiromi,Y. and W. J. Gehring (1987). "Regulation and function of the *Drosophila* segmentation gene fushi tarazu." Cell **50**(6): 963-74.

Hirose,S., Okada,M., Kaneko,S., Mitsudome,A. (2002). "Molecular genetics of human familial epilepsy syndromes." Epilepsia **43**(Suppl 9): 21-25.

Hoopengardner,B., T.Bhalla, C.Staber and R.Reenan (2003). "Nervous system targets of RNA editing identified by comparative genomics." Science **301**(5634): 832-6.

Horch, H.L., Sargent, P.B. (1995). "Perisynaptic surface distribution of multiple classes of nicotinic acetylcholine receptors on neurons in the chicken ciliary ganglion." J Neurosci. **15**(12): 7778-95.

Hoskins, R.A., C.R.Nelson, B.P.Berman, T.R. Laverty, R.A.George, L.Ciesiolka, M. Naeemuddin, A.D. Arenson, J. Durbin, R.G. David, P.E. Tabor, M.R. Bailey, D. R. DeShazo, J. Catanese, A.Mammoser, K. Osoegawa, P.J. de Jong, S.E. Celniker, R.A. Gibbs, G.M. Rubin and S.E. Scherer (2000). "A BAC-based physical map of the major autosomes of *Drosophila melanogaster*." Science **287**(5461): 2271-4.

Hu, M., C.B. Bigger and P.D. Gardner (1995). "A novel regulatory element of a nicotinic acetylcholine receptor gene interacts with a DNA binding activity enriched in rat brain." J Biol Chem **270**(9): 4497-502.

Huang, Y., Williamson, M.S., Devonshire, A.L., Windass, J.D., Lansdell, S.J., Millar, N.S. (2000). "Cloning, heterologous expression and co-assembly of Mpbeta1, a nicotinic acetylcholine receptor subunit from the aphid *Myzus persicae*." Neurosci. Lett. **284**(1-2): 116-120.

Ian W. Jones, J. B., Michael J. O'Neill, Susan Wonnacott (2004). "Alpha bungarotoxin-1.4nm gold: a novel conjugate for visualising the precise subcellular distribution of alpha 7 nicotinic acetylcholine receptors." Journal of Neuroscience Methods **134**: 65-74.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ihara,M., Matsuda,K., Otake,M., Kuwamura,M., Shimomura,M., Komai,K., Akamatsu,M., Raymond,V., Sattelle,D.B. (2003). "Diverse actions of neonicotinoids on chicken α7, α4β2 and *Drosophila*-chicken SADβ2 and ALSβ2 hybrid nicotinic acetylcholine receptors expressed in *Xenopus laevis* oocytes." Neuropharmacology **45**: 133-144.

Jacob, M.H., Berg, D.K. (1983). "The ultrastructural localization of alpha-bungarotoxin binding sites in relation to synapses on chick ciliary ganglion neurons." J Neurosci. **3**(2): 260-271.

Jacob, M.H., Berg, D.K., Lindstrom, J.M. (1984). "Shared antigenic determinant between the electrophorus acetylcholine receptor and a synaptic component on chick ciliary ganglion neurons." Proc. Natl. Acad. Sci. **81**(10): 3223-7.

Jeanclos, E. M., L. Lin, M. W. Treuil, J. Rao, M. A. DeCoster and R. Anand (2001). "The chaperone protein 14-3-3eta interacts with the nicotinic acetylcholine receptor alpha 4 subunit. Evidence for a dynamic role in subunit stabilization." J Biol Chem **276**(30): 28281-90.

Jonas, P.E., Phannavong, B., Schuster, R., Schroder, C., Gundelfinger, E.D. (1994). "Expression of the ligand-binding nicotinic acetylcholine receptor subunit D alpha 2 in the *Drosophila* central nervous system." J Neurobiol **25**(12): 1494-508.

Jones, A.K., Elgar, G. and Sattelle, D.B. (2003). "The nicotinic acetylcholine receptor gene family of the pufferfish, *Fugu rubripes*." Genomics **82**: 441-451.

Karlin, A., Akabas, M.H. (1995). "Toward a structural basis for the function of nicotinic acetylcholine receptors and their cousins." Neuron **15**: 1231-1244.

Katsani, K.R., Hajibagheri, M.A.N and Verrijzer, C.P, (1999). "Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology." EMBO J. **18**: 698-708.

Kawasaki, F., Collins, S.C., and Ordway, R.W. (2002). "Synaptic calcium-channel function in Drosophila: analysis and transformation rescue of temperature-sensitive paralytic and lethal mutations of cacophony." J. Neurosci.: 5856-5864.

Kimmerly, W., K. Stultz, S. Lewis, K. Lewis, V. Lustre, R. Romero, J. Benke, D. Sun, G. Shirley, C. Martin and M. Palazzolo (1996). "A P1-based physical map of the *Drosophila* euchromatic genome." Genome Res **6**(5): 414-30.

Kuryatov,A., V. Gerzanich, M. Nelson, F. Olale and J. Lindstrom (1997). "Mutation causing autosomal dominant nocturnal frontal lobe epilepsy alters Ca^{2+} permeability, conductance, and gating of human alpha4beta2 nicotinic

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

acetylcholine receptors." J Neurosci **17**(23): 9035-47.

Lansdell, S.J., Schmitt, B., Betz, H., Sattelle, D.B. and Millar, N.S. (1997). "Temperature-sensitive expression of *Drosophila* neuronal nicotinic acetylcholine receptors." Journal of Neurochemistry **68**(5): 1812-1819.

Lansdell, S.J., Millar, N.S. (2000). "Cloning and heterologous expression of Dalpha4, a Drosophila neuronal nicotinic acetylcholine receptor subunit: identification of an alternative exon influencing the efficiency of subunit assembly." Neuropharmacology **39**(13): 2604-2614.

Lansdell, S.J., Millar, N.S. (2000). "The influence of nicotinic receptor subunit composition upon agonist, alpha-bungarotoxin and insecticide (imidacloprid) binding affinity." Neuropharmacology **39**(4): 671-679.

Lansdell, S.J., Millar, N.S. (2002). "Dbeta3, an atypical nicotinic acetylcholine receptor subunit from *Drosophila*: molecular cloning, heterologous expression and coassembly." J Neurochem **80**(6): 1009-18.

Lansdell, S.J., Millar, N.S. (2004). "Molecular characterization of Dalpha6 and Dalpha7 nicotinic acetylcholine receptor subunits from *Drosophila*: formation of a high-affinity alpha-bungarotoxin binding site revealed by expression of subunit chimeras." J Neurochem **90**(2): 479-489.

Larsson, A., Engel, J.A. (2004). "Neurochemical and behavioral studies on ethanol and nicotine interactions." Neurosci Biobehav Rev. **27**(8): 713-720.

Lauder, J.M. (1993). "Neurotransmitters as growth regulatory signals: role of receptors and second messengers." Trends in Neurocience **16**(6): 233-240.

Le Novere, N. and J.P. Changeux (1995). "Molecular evolution of the nicotinic acetylcholine receptor: an example of multigene family in excitable cells." J Mol Evol **40**(2): 155-72.

Le Novere, N., M. Zoli, C. Lena, R. Ferrari, M.R. Picciotto, E. Merlo-Pich and J.P. Changeux (1999). "Involvement of alpha6 nicotinic receptor subunit in nicotine-elicited locomotion, demonstrated by in vivo antisense oligonucleotide infusion." Neuroreport **10**(12): 2497-501.

Leech, C.A. and D.B. Sattelle (1992). "Multiple conductances of neuronal nicotinic acetylcholine receptors." Neuropharmacology **31**(5): 501-7.

Lees, G., Beadle, D.J., Botham, R.P. (1983). "Cholinergic receptors on cultured neurons from the central nervous system of embryonic cockroaches." Brain Res. **288**(1-2): 49-59.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Levey, M. S., C. L. Brumwell, S. E. Dryer and M. H. Jacob (1995). "Innervation and target tissue interactions differentially regulate acetylcholine receptor subunit mRNA levels in developing neurons in situ." Neuron **14**(1): 153-62.

Lindstrom, J. (1997). "Nicotinic acetylcholine receptors in health and disease." Mol Neurobiol **15**(2): 193-222.

Lis, J.T., J. A. Simon and C.A. Sutton (1983). "New heat shock puffs and beta-galactosidase activity resulting from transformation of *Drosophila* with an hsp70-lacZ hybrid gene." Cell **35**(2 Pt 1): 403-10.

Littleton, J. T., Ganetzky, B. (2000). "Ion Channels and synaptic organization: analysis of the *Drosophila* genome." Neuron **26**(1): 35-43.

Loscher, W., Potschka, H., Wlaz, P., Danysz, W., Parson, C.G. (2003). "Are neuronal nicotinic receptors a target for antiepileptic drug development? Studies in different seizure models in mice and rats." Eur J Pharmacol. **466**(1-2): 99-111.

Macallan, D. R., Lunt, G.G., Wonnacott, S., Swanson, K.L., Rapoport, H., Albuquerque, E.X. (1988). "Methyllycaconitine and (+)-anatoxin-a differentiate between nicotinic receptors in vertebrate and invertebrate nervous systems." FEBS Lett. **226**(2): 357-363.

Mandelzys, A., B. Pie, E. S. Deneris and E. Cooper (1994). "The developmental increase in ACh current densities on rat sympathetic neurons correlates with changes in nicotinic ACh receptor alpha-subunit gene expression and occurs independent of innervation." J Neurosci **14**(4): 2357-64.

Maneu, V., Rojo, J., Mulet, J., Valor, L.M., Sala, F., Craido,M., Garcia,A.G. and Gandia, L. (2002). "A single neuronal nicotinic receptor alpha3alpha7beta4 is present in the bovine chromaffin cell." Ann. N. Y. Acad. Sci. **971**: 165-167.

Marks, M.J., Pauly, J.R., Gross, S.D., Deneris, E.S., Hermans-Borgmeyer, I., Heinemann, S.F., Collins, A.C. (1992). "Nicotine binding and nicotinic receptor subunit RNA after chronic nicotine treatment." J Neurosci. **12**(7): 2765-2784.

Martin-Ruiz, C.M., Lee,M., Perry,R.H., Baumann,M., Court, J.A., Perry, E.K. (2004). "Molecular analysis of nicotinic receptor expression in autism." Brain Res Mol Brain Res. **123**(1-2): 81-90.

Matsuda,K., Buckingham, S.D., Kleier, D., Rauh, J.J., Grauso,M. and Sattelle, D.B. (2001). "Neonicotinoids: insecticides acting on insect nicotinic acetylcholine receptors." Trends in Pharmacological Sciences **22**(11): 573-580.

Matsushima, N., Hirose, S.,and Sugiyama, H., et al. (2002). "Mutation (Ser284Leu) of neuronal nicotinic acetylcholine receptor alpha4 subunit associated with frontal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lobe epilepsy causes faster desensitization of the rat receptor expressed in oocytes." Epilepsy Res. **48**(3): 181-186.

Matter-Sadzinski, L., Hernandez, M.C., Roztocil, T., Ballivet,M. and Matter, J.M. (1992). "Neuronal specificity of the alpha 7 nicotinic acetylcholine receptor promoter develops during morphogenesis of the central nervous system." EMBO J. **11**(12): 4529-4538.

McDonough, J. and E.Deneris (1997). "beta43': An enhancer displaying neural-restricted activity is located in the 3'-untranslated exon of the rat nicotinic acetylcholine receptor beta4 gene." J Neurosci **17**(7): 2273-83.

McDonough,J., N. Francis, T. Miller and E. S. Deneris (2000). "Regulation of transcription in the neuronal nicotinic receptor subunit gene cluster by a neuron-selective enhancer and ETS domain factors." J Biol Chem **275**(37): 28962-70.

McGehee, D.S., Role, L.W. (1995). "Physiological diversity of nicotinic acetylcholine receptors expressed by vertebrate neurons." Annu Rev Physiol. **57**(521-546).

Meisler, M.H., Kearney, J., Ottman, R., Escayg, A. (2001). "Identification of epilepsy genes in human and mouse." Annu Rev Genet. **35**: 567-588.

Mongan, N.P., Baylis, H.A., Adcock, C., Smith, G.R., Sansom, M.S. and Sattelle, D.B. (1998). "An extensive and diverse gene family of nicotinic acetylcholine receptor alpha subunits in Caenorhabditis elegans." Recept. Channels **6**(3): 213-228.

Mongan, N.P., Jones, A.K., Smith, G.R., Sansom, M.S. and Sattelle, D.B. (2002). "Novel alpha7-like nicotinic acetylcholine receptor subunits in the nematode Caenorhabditis elegans." Protein Sci. **11**(5): 1162-1171.

Moran, L., M.E.Mirault, A.Tissieres, J.Lis, P.Schedl, S. Artavanis-Tsakonas and W.J. Gehring (1979). "Physical map of two *D. melanogaster* DNA segments containing sequences coding for the 70,000 dalton heat shock protein." Cell **17**(1): 1-8.

Morley, B.J. and J. F. Rodriguez-Sierra (2004). "A phenotype for the alpha7 nicotinic acetylcholine receptor null mutant." Brain Res **1023**(1): 41-7.

Moss, B.L., Role, L.W. (1993). "Enhanced ACh sensitivity is accompanied by changes in ACh receptor channel properties and segregation of ACh receptor subtypes on sympathetic neurons during innervation *in vivo*." J. Neurosci. **13**(1): 13-28.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Mulholland, N.M., I.F.King and R.E. Kingston (2003). "Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins *Zeste* and GAGA." Genes Dev **17**(22): 2741-6.

Nai, Q., McIntosh, J.M., Margiotta, J.F. (2003). "Relating neuronal nicotinic acetylcholine receptor subtypes defined by subunit composition and channel function." Mol Pharmacol **63**(2): 311-324.

Narahashi, T., Marszalec, W., Moriguchi, S., Yeh, J.Z., Zhao, X. (2003). "Unique mechanism of action of Alzheimer's drugs on brain nicotinic acetylcholine receptors and NMDA receptors." Life Sci. **74**(2-3): 281-291.

Nef, P., Oneyser, C., Alliod, C., Couturier, s. and Ballivet, M. (1988). "Genes expressed in the brain define three distinct neuronal nicotinic acetylcholine receptors." EMBO J. **7**(3): 595-601.

Nguyen, L., Rigo, J-M., Rocher, V., Belachew, S., Malgrange, B., Rogister, B., Leprince, P., Moonen, G. (2001). "Neurotransmitters as early signals for central nervous system development." Cell Tissue Res **305**: 187-202.

O'Connell, M.A. (1997). "RNA editing: rewriting receptors." Curr Biol **7**(7): R437-9.

Okada, M., Hirose, S. (1998). "Chromatin remodeling mediated by *Drosophila* GAGA factor and ISWI activates *fushi tarazu* gene transcription in vitro." Mol Cell Biol **18**: 2455-2461.

Ortells, M.O., Lunt, G.G. (1995). "Evolutionary history of the ligand-gated ion-channel superfamily of receptors." Trends in Neurocience **18**(3): 121-127.

Ortells, M.O. and G.G. Lunt (1995). "Evolutionary history of the ligand-gated ion-channel superfamily of receptors." Trends Neurosci **18**(3): 121-7.

Patton, D. E., T. Silva and F. Bezanilla (1997). "RNA editing generates a diverse array of transcripts encoding squid Kv2 K+ channels with altered functional properties." Neuron **19**(3): 711-22.

Picciotto, M.R. (2003). "Nicotine as a modulator of behavior: beyond the inverted U." Trends in Pharmacological Sciences **24**(9): 493-499.

Picciotto, M.R., M.Zoli and J.P.Changeux (1999). "Use of knock-out mice to determine the molecular basis for the actions of nicotine." Nicotine Tob Res **1 Suppl 2**: S121-5; discussion S139-40.

Ptashne, M. and A. Gann (1997). "Transcriptional activation by recruitment." Nature **386**(6625): 569-77.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Qian, S., B.Varjavand and V. Pirrotta (1992). "Molecular analysis of the zeste-white interaction reveals a promoter-proximal element essential for distant enhancer-promoter communication." Genetics **131**(1): 79-90.

Raff, J., Kellum, R. and Alberts, B. (1994). "The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle." EMBO J. **13**: 5977-5083.

Raff, J. W., R. Kellum and B. Alberts (1994). "The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle." EMBO J **13**(24): 5977-83.

Raimondi, E., F. Rubboli, D. Moralli, B. Chini, D. Fornasari, P. Tarroni, L. De Carli and F. Clementi (1992). "Chromosomal localization and physical linkage of the genes encoding the human alpha 3, alpha 5, and beta 4 neuronal nicotinic receptor subunits." Genomics **12**(4): 849-50.

Ramiro Salas, A. O.-U., Ron S. Broide, Arthur Beaudet, Richard Paylor, and Mariella De Biasi (2003). "The nicotinic acetylcholine receptor subunit $\alpha 5$ mediates short-term effects of nicotine *in vivo*." Molecular Pharmacology **63**(5): 1059-1066.

Reenan, R. A., Hanrahan, C.J. and Barry, G. (2000). "The mle(napts) RNA helicase mutation in *Drosophila* results in a splicing catastrophe of the para Na+ channel transcript." Neuron **25**: 139-149.

Rezvani, A.H., Levin, E.D. (2001). "Cognitive effects of nicotine." Biol Psychiatry **49**(3): 258-267.

Role, L.W., Berg, D.K.(1996). "Nicotinic receptors in the development and modulation of CNS synapses." Neuron **16**: 1077-1085.

Rosenberg, M.M., R.C. Blitzblau, D.P. Olsen and M.H. Jacob (2002). "Regulatory mechanisms that govern nicotinic synapse formation in neurons." J Neurobiol **53**(4): 542-55.

Rubin, G.M. (1988). "*Drosophila melanogaster* as an experimental organism." Science **240**(4858): 1453-9.

Rubin, G.M. and A.C. Spradling (1982). "Genetic transformation of *Drosophila* with transposable element vectors." Science **218**(4570): 348-53.

Rubin, G. M. and A. C. Spradling (1983). "Vectors for P element-mediated gene transfer in *Drosophila*." Nucleic Acids Res **11**(18): 6341-51.

Saragoza, P.A., Modir, J.G., Goel, N., French, K.L., Li, L., Nowak, M.W., Stitzel,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

J.A. (2003). "Identification of an alternatively processed nicotinic receptor $\alpha7$ subunit RNA in mouse brain." Molecular Brain Research **117**: 15-26.

Sargent, P.B. (1993). "The diversity of neuronal nicotinic acetylcholine receptors." Annu Rev Neurosci. **16**: 403-443.

Sattelle, D.B. (1992). "Receptors for L-glutamate and GABA in the nervous system of an insect (Periplaneta americana)." Comp Biochem Physiol C **103**(3): 429-38.

Sawruk, E., Hermans-Borgmeyer, I., Betz, H., Gundelfinger, E.D. (1988). "Characterization of an invertebrate nicotinic acetylcholine receptor gene: the *ard* gene of *Drosophila melanogaster*." FEBS Lett. **235**(1-2): 40-46.

Sawruk,E., Udri, C., Betz, H., Schmitt, B. (1990). "SBD, a novel structural subunit of the *Drosophila* nicotinic acetylcholine receptor, shares its genomic localization with two alpha-subunits." FEBS Lett. **273**(1-2): 177-181.

Sawruk,E., Schloss,P., Betz, H., and Schmitt, B. (1990). "Heterogeneity of *Drosophila* nicotinic acetylcholine receptors: SAD, a novel developmentally regulated α-subunit." EMBO J. **9**(9): 2671-2677.

Schaub,M. and W. Keller (2002). "RNA editing by adenosine deaminases generates RNA and protein diversity." Biochimie **84**(8): 791-803.

Schloss,P., Hermans-Borgmeyer, I., Betz,H., Gundelfinger, E.D. (1988). "Neuronal acetylcholine receptors in *Drosophila*: the ARD protein is a component of a high-affinity α-bungarotoxin binding complex." EMBO J. **7**(9): 2889-2894.

Schloss,P., Betz, H., Schroder,C., Gundelfinger,E.D. (1991). "Neuronal nicotinic acetylcholine receptors in *Drosophila*: antibodies against an alpha-like and a non-alpha-subunit recognize the same high-affinity alpha-bungarotoxin binding complex." J Neurochem **57**(5): 1556-1562.

Schloss,P., Mayser, W., Gundelfinger, E.D., Betz, H. (1992). "Cross-linking of ^{125}I-alpha-bungarotoxin to *Drosophila* head membrane identifies a 42 kDa toxin binding polypeptide." Neurosci. Lett. **145**(1): 63-66.

Schoepfer,R., Conroy, W.G., Whiting, P., Gore, M. and Lindstrom, J. (1990). "Brain alpha-bungarotoxin binding protein cDNAs and MAbs reveal subtypes of this branch of the ligand-gated ion channel gene superfamily." Neuron **5**(1): 35-48.

Schulz, R., Sawruk, E., Mulhardt, C., Bertrand, S., Baumann, A., Phannavong, B., Betz, H., Bertrand, D., Gundelfinger, E.D., Schmitt, B. (1998). "Dalpha3, a new functional alpha subunit of nicotinic acetylcholine receptors from *Drosophila*." J

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Neurochem **71**(2): 853-863.

Schulz, R., Bertrand, S., Chamaon, K., Smalla, K.H., Gundelfinger, E.D., Bertrand, D. (2000). "Neuronal nicotinic acetylcholine receptors from *Drosophila*: two different types of alpha subunits coassemble within the same receptor complex." J Neurochem **74**(6): 2537-46.

Schuster, R., Phannavong, B., Schroder, C., Gundelfinger, E.D. (1993). "Immunohistochemical localization of a ligand-binding and a structural subunit of nicotinic acetylcholine receptors in the central nervous system of *Drosophila melanogaster*." J Comp Neurol. **335**(2): 149-162.

Scott, J. (1997). "RNA editing. Message change for a fat controller." Nature **387**(6630): 242-3.

Seguela,P., Wadiche, J., Dineley-Miller, K., Dani, J.A., and Patrick, J.W. (1993). "Molecular cloning, functional properties and distribution of rat brain α7: a nicotinic cation channel highly permeable to calcium." J. Neurosci. **13**(2): 596-604.

Semenov, E.P. and W.L. Pak (1999). "Diversification of *Drosophila* chloride channel gene by multiple posttranscriptional mRNA modifications." J Neurochem **72**(1): 66-72.

Sgard, F., Obosi, L.A., King, L.A., Windass, J.D. (1993). "ALS and SAD-like nicotinic acetylcholine receptor subunit genes are widely distributed in insects." Insect Mol Biol. **2**(4): 215-223.

Sgard, F., Fraser, S.P., Katkowska, M.J., Djamgoz, M.B., Dunbar, S.J., Windass, J.D. (1998). "Cloning and functional characterisation of two novel nicotinic acetylcholine receptor alpha subunits from the insect pest *Myzus persicae*." J. Neurochem **71**(3): 903-912.

Shizuya, H., B. Birren, U. J. Kim, V. Mancino, T. Slepak, Y. Tachiiri and M. Simon (1992). "Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector." Proc Natl Acad Sci U S A **89**(18): 8794-7.

Shoop,R.D., Martone,M.E., Yamada,N., Ellisman, M.H., Berg,D.K. (1999). "Neuronal acetylcholine receptors with alpha7 subunits are concentrated on somatic spines for synaptic signaling in embryonic chick ciliary ganglia." J. Neurosci. **19**(2): 692-704.

Siden-Kiamos, I., R.D.Saunders, L.Spanos, T.Majerus, J.Treanear, C.Savakis, C. Louis, D.M.Glover, M.Ashburner and F.C.Kafatos (1990). "Towards a physical map of the *Drosophila melanogaster* genome: mapping of cosmid clones within defined genomic divisions." Nucleic Acids Res **18**(21): 6261-70.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Soeller, W.C., S.J. Poole and T.Kornberg (1988). "*In vitro* transcription of the *Drosophila* engrailed gene." Genes Dev **2**(1): 68-81.

Sommer,B., M.Kohler, R. Sprengel and P.H. Seeburg (1991). "RNA editing in brain controls a determinant of ion flow in glutamate-gated channels." Cell **67**(1): 11-9.

Spradling, A.C. and G.M.Rubin (1982). "Transposition of cloned P elements into *Drosophila* germ line chromosomes." Science **218**(4570): 341-7.

Strutt,H., G.Cavalli and R.Paro (1997). "Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression." EMBO J **16**(12): 3621-32.

Tchoubrieva, E.B. and J.B. Gibson (2004). "Conserved (CT)n.(GA)n repeats in the non-coding regions at the Gpdh locus are binding sites for the GAGA factor in *Drosophila melanogaster* and its sibling species." Genetica **121**(1): 55-63.

Thompson, J.D., D.G. Higgins and T.J.Gibson (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**(22): 4673-80.

Thummel, C.S. (1989). "The *Drosophila* E74 promoter contains essential sequences downstream from the start site of transcription." Genes Dev **3**(6): 782-92.

Thummel, C.S., A. M. Boulet and H. D. Lipshitz (1988). "Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection." Gene **74**(2): 445-56.

Tornoe,C., D.Bai, L. Holden-Dye, S.N.Abramson and D.B. Sattelle (1995). "Actions of neurotoxins (bungarotoxins, neosurugatoxin and lophotoxins) on insect and nematode nicotinic acetylcholine receptors." Toxicon **33**(4): 411-24.

Tribollet,E., Bertrand,D., Marguerat, A., and Raggenbass, M. (2004). "Comparative distribution of nicotinic receptor subtypes during development, adulthood and aging: an autoradiographic study in the rat brain." Neuroscience.

Tsunoyama, K. and T. Gojobori (1998). "Evolution of nicotinic acetylcholine receptor subunits." Mol Biol Evol **15**(5): 518-27.

Ullian, E. M., McIntosh, J.M., Sargent, P.B. (1997). "Rapid synaptic transmission in the avian ciliary ganglion is mediated by two distinct classes of nicotinic receptors." J Neurosci. **17**(19): 7210-9.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Vassilatis, D. K., Elliston, K.O., Paress, P.S., Hamelin, M., Arena, J.P., Schaeffer, J.M., Van der Ploeg, L.H.T., Cully, D.F. (1997). "Evolutionary relationship of the ligand-gated ion channels and the avermerctin-sensitive, glutamate-gated chloride channels." J Mol Evol **44**: 501-508.

Venter, J. C., M. D. Adams, G. G. Sutton, A. R. Kerlavage, H. O. Smith and M. Hunkapiller (1998). "Shotgun sequencing of the human genome." Science **280**(5369): 1540-2.

Vetter, D. E., Liberman, M.C., Mann, J., Barhanin, J., Boulter, J., Brown, M.C., Saffiote-Kolman, J., Heinemann, S.F., Elgoyhen, A.B. (1999). "Role of alpha9 nicotinic ACh receptor subunits in the development and function of cochlear efferent innervation." Neuron **23**(1): 93-103.

Wadsworth, S. C., Rosenthal, L.S., Kammermeyer, K.L., Potter, M.B., and Nelson, D.J. (1988). "Expression of a *Drosophila melanogaster* acetylcholine receptor-related gene in the central nervous system." Mol Cell Biol **8**: 778-785.

Whitehouse, P. J., Martino, A.M., Marcus, K.A., Zweig, R.M., Singer, H.S., Price, D.L., Kellar, K.J. (1988). "Reductions in acetylcholine and nicotine binding in several degenerative diseases." Arch Neurol. **45**(7): 722-724.

Wonnacott, S. (1997). "Presynaptic nicotinic ACh receptors." Trends in Neurocience **20**: 92-98.

Xie, J. and D. P. McCobb (1998). "Control of alternative splicing of potassium channels by stress hormones." Science **280**(5362): 443-6.

Yang, L.-C. (1996). "Identification of *cis*-acting regulatory sequences required for transcriptional control of the *ard* gene of a neuronal nicotinic acetylcholine receptor from *Drosophila*."

Yang, X., Y. Kuo, P. Devay, C. Yu and L. Role (1998). "A cysteine-rich isoform of neuregulin controls the level of expression of neuronal nicotinic receptor channels during synaptogenesis." Neuron **20**(2): 255-70.

Zirger, J. M., Beattie, C.E., McKay, D.B. and Thomas Boyd, R. (2003). "Cloning and expression of zebrafish neuronal nicotinic acetylcholine receptors." Gene Expr. Patterns **3**(6): 747-754.

Zoli, M., M. R. Picciotto, R. Ferrari, D. Cocchi and J. P. Changeux (1999). "Increased neurodegeneration during ageing in mice lacking high-affinity nicotine receptors." EMBO J **18**(5): 1235-44.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.