

## Molecular Evolution of Three Avian Neurotrophin Genes: Implications for Proregion Functional Constraints

Ravinder N.M. Sehgal,<sup>1</sup> Irby J. Lovette<sup>2</sup>

<sup>1</sup> Center for Tropical Research, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA

<sup>2</sup> Evolutionary Biology Program, Cornell Laboratory of Ornithology, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA

Received: 6 May 2002 / Accepted: 17 April 2003

**Abstract.** Neurotrophin proteins are essential for the survival, differentiation, and maintenance of neurons in the peripheral and central nervous systems. Recent studies have shown that the unprocessed proforms of the neurotrophins are preferential high-affinity ligands for p75<sup>NTR</sup> and potent inducers of p75<sup>NTR</sup>-mediated cell death. Here, we explore differences in the selective constraints acting on the proregions of the three avian neurotrophin genes—NT-3, BDNF, and NGF—in an explicit phylogenetic context. We found a 50-fold difference in levels of constraint as estimated by  $d_N/d_S$  ratios, with the NGF proregion showing the lowest degree of constraint and BDNF the highest. These patterns suggest that the high conservation exhibited by the BDNF proregion results from intense functional constraints that are relaxed in NGF and somewhat relaxed in NT-3. The proregion of BDNF is likely to have a function that differentiates it from the corresponding regions of the NGF and NT-3 genes, suggesting that BDNF is the avian neurotrophin most likely to be used both in its precursor and mature forms *in vivo*.

**Key words:** Neurotrophin — NT-3 — BDNF — NGF — Functional constraints — Proregion

### Introduction

Neurotrophin proteins are fundamental developmental regulators for the vertebrate nervous system that promote the survival, growth, differentiation, and plasticity of neurons (e.g., Chao 2000; Patapoutian and Reichardt 2001). The neurotrophin gene family comprises six known constituents, the three genes included in the present study—nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), and neurotrophin-3 (NT-3)—and three other genes that are less well characterized: neurotrophin-4/5 (NT-4/5; also known as NT-4), which is absent from chickens (Carlos Ibañez and Finn Hallböök, personal communication) and presumably from other birds; and two that have been found in fish but not in other vertebrates, NT-6 and NT-7 (Hallböök 1999; Heinrich and Lum 2000). The differentiation of these now specialized neurotrophins occurred in concert with increases in complexity of the early vertebrate nervous system, probably via several gene duplication events from a single precursor neurotrophin found in the common ancestor of the Gnathostomata (Hallböök et al. 1998).

All known neurotrophin genes share a common structure. *In vivo*, they are translated as a pre-protein which is proteolytically processed at a conserved arginine motif to yield the mature form. The presequence mediates the translocation of the proteins to the endoplasmic reticulum, and the propeptide region has been shown to promote folding of the mature region (Rattenholl et al. 2001a). Recent *in vitro* experiments have shown that a dimeric

## BDNF

|             | I          |             |            |            |            |           |            |            |  |  |  |
|-------------|------------|-------------|------------|------------|------------|-----------|------------|------------|--|--|--|
| chicken     | SVRGHGSLAY | PGLRTHGTLE  | SLTGNPAGSR | GLTSLADTFE | HVIEELLEDD | QDIQPSEEN | KDADLYTSRV | MLSSQVPLEP |  |  |  |
| human       | NI..Q.G... | ..V.....    | ..VN..K... | .....      | .....      | .....     | ..KVR..N.. | .....      |  |  |  |
| alligator   | ..Q.....   | ..S.S...    | .....      | .....      | .....      | .....     | .....      | .....      |  |  |  |
| ostrich     | G..Q.....  | ..N..T...   | .....      | .....      | .....      | ..M....   | .....      | .....      |  |  |  |
| duck        | ..Q.....   | ..N.....    | .....      | .....      | .....      | .....     | .....      | .....      |  |  |  |
| kestrel     | ..Q.....   | ..N..S...   | .....      | .....      | .....      | .....     | .....      | .....      |  |  |  |
| kingfisher  | ..Q.....   | ..VN..S...  | .....      | .....      | .....      | .....     | .....      | .....      |  |  |  |
| hummingbird | ..Q.....   | ..N.N..S... | .....      | .....      | .....      | ..R....   | .....      | .....      |  |  |  |
| bulbul      | ..L..Q...  | ..IN..S...  | .....      | .....      | .....      | ..G....   | .....      | .....      |  |  |  |
| sunbird     | ..L..Q...  | ..IN..S...  | .....      | .....      | .....      | ..G....   | .....      | .....      |  |  |  |
| manakin     | ..L..Q...  | ..N..S...   | .....      | .....      | .....      | ..G....   | .....      | .....      |  |  |  |
| flycatcher  | ..L..Q...  | ..N..S...   | .....      | .....      | .....      | ..G....   | .....      | .....      |  |  |  |

|             | II         |            |            |            |            |           |             |            |  |  |
|-------------|------------|------------|------------|------------|------------|-----------|-------------|------------|--|--|
| chicken     | PILLFLEEYK | NYLDDAABSM | AVRRHEDPAR | RGELSVCDSI | SEWVTAEEKK | TAVMGSATV | TVLEKVPVPEK | QGLKQYFFET |  |  |
| human       | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| alligator   | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| ostrich     | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| duck        | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| kestrel     | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| kingfisher  | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| hummingbird | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| bulbul      | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| sunbird     | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| manakin     | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |
| flycatcher  | .....      | .....      | .....      | .....      | .....      | .....     | .....       | .....      |  |  |

|             | III        |            | IV        |            | V     |  |
|-------------|------------|------------|-----------|------------|-------|--|
| chicken     | KCNPKGATKE | GCRGIDKRHH | NSGCKTQSY | VRLATMDNKK | R     |  |
| human       | ..N.....   | .....      | .....     | .....      | ..... |  |
| alligator   | .....      | .....      | .....     | .....      | ..... |  |
| ostrich     | .....      | .....      | .....     | .....      | ..... |  |
| duck        | .....      | .....      | .....     | .....      | ..... |  |
| kestrel     | .....      | .....      | .....     | .....      | ..... |  |
| kingfisher  | .....      | .....      | .....     | .....      | ..... |  |
| hummingbird | .....      | .....      | .....     | .....      | ..... |  |
| bulbul      | .....      | .....      | .....     | .....      | ..... |  |
| sunbird     | .....      | .....      | .....     | .....      | ..... |  |
| manakin     | .....      | .....      | .....     | .....      | ..... |  |
| flycatcher  | .....      | .....      | .....     | .....      | ..... |  |

## NT-3

|             | I          |              |            |             |            |           |            |            |  |  |
|-------------|------------|--------------|------------|-------------|------------|-----------|------------|------------|--|--|
| chicken     | SLIIKLIRAD | ILKNKLSKQV   | MDVKENYQNI | VQKVEDHQEM  | DGDEVKVSDF | QPVISMDTL | LRQQRRYNSP | RVLLSDNTEP |  |  |
| human       | .....      | ..M.V.....   | ..ST       | LP.A.APR.P  | ERGGPA..A. | .....     | ..A..E..   | .....      |  |  |
| alligator   | .....      | ..M.V.I..... | ..T        | ..E.AQ.D.   | ..A.....   | .....     | ..E.....   | .....      |  |  |
| ostrich     | .....      | ..M.V.I..... | ..HH.T     | ..M..A..RD. | ..V.....   | .....     | ..PV.....  | .....      |  |  |
| duck        | .....      | ..I.V.I..... | ..T        | ..K..AQ.DV  | .....      | .....     | .....      | .....      |  |  |
| kestrel     | .....      | ..M.V.I..... | ..T        | ..ADTQ.D.   | .....      | .....     | .....      | .....      |  |  |
| kingfisher  | .....      | ..M.V.I..... | ..T        | ..DTQ.D.    | NE.....    | .....     | ..V.....   | .....      |  |  |
| hummingbird | .....      | ..M.V.I..... | ..T        | ..ADTQ.D.   | .....      | .....     | .....      | .....      |  |  |
| bulbul      | .....      | ..M.V.I..... | ..T        | ..TDAQ.D.   | ..G.....   | .....     | ..V.....   | .....      |  |  |
| sunbird     | .....      | ..M.V.I..... | ..T        | ..TDTQ.DV   | ..E.....   | .....     | ..V.....   | .....      |  |  |
| manakin     | .....      | ..M.V.I..... | ..T        | ..ADTQ.D.   | ..E.....   | .....     | ..V.....   | .....      |  |  |
| flycatcher  | .....      | ..M.V.I..... | ..T        | ..ADTQ.DV   | ..E.....   | .....     | ..V.....   | .....      |  |  |

|             | II        |             | NH2        |            | I          |            | II         |            |
|-------------|-----------|-------------|------------|------------|------------|------------|------------|------------|
| chicken     | EPPLYLITD | YVGSVVVLR   | TSRRKRFAEH | KSHRGEYSVC | DSESLWVTDK | SSAIDIRGHQ | VTVLGEIKTG | NSPVKQYFFE |
| human       | ..M.....  | ..P.A.      | .....      | .....      | .....      | .....      | .....      | .....      |
| alligator   | .....     | ..I.....    | .....      | .....      | .....      | .....      | .....      | .....      |
| ostrich     | .....     | ..M.V.....  | .....      | .....      | .....      | .....      | .....      | .....      |
| duck        | .....     | ..M.I.N..V. | .....      | .....      | .....      | .....      | .....      | .....      |
| kestrel     | .....     | ..M.I.N..V. | .....      | .....      | .....      | .....      | .....      | .....      |
| kingfisher  | .....     | ..M.I.N..V. | .....      | .....      | .....      | .....      | .....      | .....      |
| hummingbird | .....     | ..M.I.N..V. | .....      | .....      | .....      | .....      | .....      | .....      |
| bulbul      | .....     | ..M.I.N..V. | .....      | RG..       | .....      | .....      | .....      | .....      |
| sunbird     | .....     | ..M.I.N..V. | .....      | RG..       | .....      | .....      | .....      | .....      |
| manakin     | .....     | ..M.I.N..V. | .....      | G..        | .....      | .....      | .....      | .....      |
| flycatcher  | .....     | ..M.I.N..V. | .....      | G..        | .....      | .....      | .....      | .....      |

|             | III        |            | IV         |        |
|-------------|------------|------------|------------|--------|
| chicken     | TRCKEAKPVK | NGCRGIDDKH | WNSGCKTSQT | YVRALT |
| human       | ..R.....   | .....      | .....      | .....  |
| alligator   | .....      | .....      | .....      | .....  |
| ostrich     | .....      | .....      | .....      | .....  |
| duck        | .....      | .....      | .....      | .....  |
| kestrel     | .....      | .....      | .....      | .....  |
| kingfisher  | .....      | .....      | .....      | .....  |
| hummingbird | .....      | .....      | .....      | .....  |
| bulbul      | .....      | .....      | .....      | .....  |
| sunbird     | .....      | .....      | .....      | .....  |
| manakin     | .....      | .....      | .....      | .....  |
| flycatcher  | .....      | .....      | .....      | .....  |

## NGF

|             | I          |            |            |            |            |           |            |              |  |  |
|-------------|------------|------------|------------|------------|------------|-----------|------------|--------------|--|--|
| chicken     | PLEYPAEHSI | PSTQSQNGQH | IAKAAPQTH  | GRFAWMPDGT | EDLNIAHQDN | FFKKKFRSS | RVLFPSTQPP | VSRKQGSTGF   |  |  |
| human       | .AGHTIPQAH | WTKL.HSLDT | ALRR.RS-AP | AAAAAARVAG | QTAH.TV.PK | L..R.L.P  | .....      | EAADT.DLD.   |  |  |
| ostrich     | ..G...P.   | SN.HR..H.  | P...HAS.   | SHST.TI.R. | AV.TV.PK   | L..R..P   | .....      | ..EQ.HM..    |  |  |
| duck        | .....      | LN.HR..RH. | PQ.....S.  | SHST.TIGRR | ..I.T.PS   | L..R.H.P  | .....      | ..GQ.NM.Q    |  |  |
| kestrel     | .....      | LNRHP..RH. | P.....S.   | HST.TIGRR  | ..A.T.PV   | V..R..P   | .....      | ..A.GH..V.   |  |  |
| kingfisher  | .....      | LNSH..RH.  | P.....S.   | SNTT.TRS.R | ..AI.T.PK  | I..RH..P  | .....      | ..A.GQ..N.E. |  |  |
| hummingbird | .....      | LDAHR..RH. | PE.....P.  | HA.S-RR    | ..AI.TV.PK | I.R.R..P  | .....      | ..AGQ..L.    |  |  |
| manakin     | .....      | NARR.DRS.  | LPE..R.P.  | HT.SAGRR   | ..AI.TV.PK | I.R.R..P  | .....      | ..VGQ..GLE.  |  |  |
| flycatcher  | .....      | .....      | .....      | .....      | .....      | .....     | .....      | .....        |  |  |

|             | II         |            | NH2        |            | I         |            | II         |            | III |  |
|-------------|------------|------------|------------|------------|-----------|------------|------------|------------|-----|--|
| chicken     | -LSSAVSLNR | TARTKR-TAH | PVLHRGEFSV | CDSVSMVGGD | KTATDIKQK | EVTVLGEVNI | NNNVFKQYFF | ETKCRDPRPV |     |  |
| human       | ..G...P.   | H.S..SSS.  | ..IF.....  | .....      | .....     | .....      | .....      | .....      |     |  |
| ostrich     | ..N.D...   | .....      | .....      | .....      | .....     | .....      | .....      | .....      |     |  |
| duck        | .....      | .....      | .....      | .....      | .....     | .....      | .....      | .....      |     |  |
| kestrel     | ..G...P.   | .....      | .....      | .....      | .....     | .....      | .....      | .....      |     |  |
| kingfisher  | ..G.GP.    | .....      | .....      | .....      | .....     | .....      | .....      | .....      |     |  |
| hummingbird | ..G...P.   | .....      | .....      | .....      | .....     | .....      | .....      | .....      |     |  |
| manakin     | ..GG.G.    | .....      | .....      | .....      | .....     | .....      | .....      | .....      |     |  |
| flycatcher  | ..GG.G.    | .....      | .....      | .....      | .....     | .....      | .....      | .....      |     |  |

|             | IV         |            | V          |          |
|-------------|------------|------------|------------|----------|
| chicken     | SSGCRGIDAK | HWNSYCTITH | TFVKALTMEG | KQAAWRFI |
| human       | .....      | .....      | .....      | .....    |
| ostrich     | .....      | .....      | .....      | .....    |
| duck        | .....      | .....      | .....      | .....    |
| kestrel     | .....      | .....      | .....      | .....    |
| kingfisher  | .....      | .....      | .....      | .....    |
| hummingbird | .....      | .....      | .....      | .....    |
| manakin     | .....      | .....      | .....      | .....    |
| flycatcher  | .....      | .....      | .....      | .....    |

recombinant human NGF-protein maintains biological activity (Rattenholl et al. 2001b), but less is known about the general function of the preproregions in vivo. Neurotrophin proteins bind to two classes of receptors, the p75 neurotrophin receptor (p75<sup>NTR</sup>) to which all neurotrophins bind and members of the Trk receptor family for which there is more ligand specificity (Kaplan and Miller 2000). The p75<sup>NTR</sup> receptor has been shown to mediate the apoptosis of neurons (Dobrowsky and Carter 2000) and oligodendrocytes (Beattie et al. 2002). The Trk receptors are high-affinity tyrosine kinase receptors that mediate growth and survival responses to the neurotrophins (Friedman and Greene 1999).

We sequenced the three avian neurotrophin loci from a set of 10 avian taxa that spans a broad range of evolutionary diversity, from representatives of the most ancient extant avian lineages to recently derived members of several suborders. Because the phylogenetic relationships of these taxa are known with reasonable certainty, we were able to reconstruct nucleotide and amino acid changes in the neurotrophins using an explicit phylogenetic context. This allows us to compare rates and patterns of nucleotide and amino acid evolution in the pro and mature regions of the three loci to assess whether their large differences in proregion variation stem from differences in mutation rate or from functional constraints. Recent experiments have shown that the proform, compared to the mature form, of NGF preferentially activates the p75<sup>NTR</sup> receptor (Lee et al. 2001; Chao and Bothwell 2002; Ibáñez 2002; Beattie et al. 2002). The results presented here show that among the three avian neurotrophins, the proregion of BDNF has the highest degree of functional constraint and thus is likely to have a function not shared with the proregion of NGF and NT-3.

## Materials and Methods

### Laboratory Methods

We sequenced 587 nucleotides of the NT-3 gene corresponding to bases 172–759 of the human NT-3 sequence (Jones and Reichardt 1990 [GenBank M37763]), 602 nucleotides of the BDNF gene corresponding to bases 148–750 of the human BDNF sequence (Jones and Reichardt 1990 [GenBank M37762]), and 587 nucleotides of the NGF gene corresponding to bases 227–814 of the chicken NGF sequence (Ebendal et al. 1986 [GenBank X04003])

from up to 10 avian taxa and from the nonavian outgroup taxon *Alligator mississippiensis*.

Additional sequences were obtained from GenBank: *Homo sapiens* NGF (NM\_002506), *Homo sapiens* BDNF (M37762), *Homo sapiens* NT-3 (M37763), *Gallus gallus* NGF (X04003), *Gallus gallus* BDNF (M83377), and *Gallus gallus* NT-3 (Z30092). In addition to obtaining these chicken sequences from GenBank, we sequenced the three *Gallus gallus* loci using the protocols described below, and in each case the nucleotide sequence we obtained was identical to the sequenced deposited in GenBank by previous investigators.

The taxa we sequenced and included in phylogenetic analyses (and their vernacular names used for convenience hereafter) were *Alligator mississippiensis* (alligator), *Struthio camelus* (ostrich), *Gallus gallus* (chicken), *Anas platyrhynchos* (mallard), *Falco sparverius* (kestrel), *Halcyon malimbica* (kingfisher), *Coeligena torquata* (hummingbird), *Chloropipo holochlora* (manakin), *Mionectes striatocollis* (flycatcher), *Andropodus latirostris* (bulbul), and *Nectarinia olivacea* (sunbird). Muscle samples from the alligator, ostrich, and chicken were obtained from commercial breeders; blood for the kestrel was kindly provided by L.A. Tell; DNA for the mallard was kindly provided by R.E. Ricklefs; and blood samples from the hummingbird, the kingfisher, and four passerines were kindly provided by T.B. Smith. Our phylogenetic analyses therefore included one nonavian outgroup, the alligator; one representative Palaeognathae, the ostrich; and nine Neognathae. These Neognathae represent six avian orders in the most widely used traditional avian classifications (e.g., Wetmore 1960). The order Passeriformes was represented by four taxa, two suboscines (manakin and flycatcher) and two oscines (bulbul and sunbird).

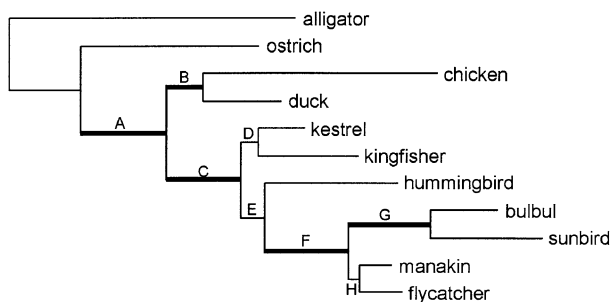
Purified genomic DNA was obtained from each sample using standard phenol–chloroform extraction protocols from blood or muscle tissue. Amplified DNA templates used for sequencing reactions were obtained via gene-specific PCRs using primers we designed from the published chicken sequences. All 5' primers include the start codon of the protein-coding sequence, and all 3' primers include the stop codon of the protein-coding sequence. NT-3 primers were ChickNT3-5' (ATGTCATCTGTGTTTATGTG) and ChickNT3-3' (GTTCTTCCTATTTTTCTTGAC), BDNF primers were ChickBDNF5' (ATGACCATCCTTTCTTACTATG) and ChickBDNF3' (TCTTCCCCTTTAATGGTTAATGTAC), and NGF primers were AllNGF5' (GGTGCATAGCGTAATGCCATG) and AllNGF3' (ATAATTTACAGGCTGAGGTAG).

Each 25- $\mu$ l PCR reaction mixture contained 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 4.0 mM MgCl<sub>2</sub>, 0.001% gelatin, 200  $\mu$ M of each dNTP, each primer at 1.25  $\mu$ M, and 1 unit Amplitaq (Cetus–Perkin–Elmer). Reaction conditions were also invariant: an initial 5-min denaturation at 94°C; 35 cycles of 30 s of denaturation at 94°C, 30 s of annealing at 55°C, and 30 s of extension at 72°C; followed by 7 min of extension at 72°C. PCR products were checked via electrophoresis in 2% agarose and were purified using Qiagen spin-columns. Cycle sequencing reactions (with dye-terminator fluorescent labeling) were conducted using the amplification primers followed by electrophoresis using an ABI model 377 automated DNA sequencer (Applied Biosystems).

Strong PCR bands and clean chromatograms were readily obtained from all taxa for BDNF and NT-3 and from most taxa for

**Fig. 1.** Amino acid sequences of avian, alligator, and human BDNF, NT-3, and NGF neurotrophin genes. *Dark vertical lines with arrowheads* indicate the proteolytic processing sites that separate the prepro and mature region of each protein. *Shaded* portions of mature region sequences indicate the variable regions (NH<sub>2</sub>, I, II, III, IV, V, and COOH) of the neurotrophins as described by Ibáñez et al. (1993). *Darker-shaded* regions in the proregions of the neurotrophins (I and II) indicate the two conserved

domains necessary and sufficient for the biosynthesis of biologically active NGF as described by Suter et al. (1991). *Asterisks* identify sites known to be important for receptor binding (Ibáñez et al. 1992; Urfer et al. 1994). Indels are indicated by hyphens. The protein coding sequences of BDNF, NT-3, and NGF shown here begin at amino acid numbers 25, 26, and 6, respectively, of the published human sequences. The human sequences are included here for reference but were not used in phylogenetic analyses.



**Fig. 2.** Phylogenetic relationships among 10 avian taxa based on analyses of combined BDNF and NT-3 nucleotide sequences. The tree shown is a phylogram from a Bayesian maximum likelihood search; *thick lines* indicate portions of this topology that were recovered across a variety of reconstruction techniques. Support indices for lettered internodes under these various reconstruction methods are given in Table 1.

NGF using the protocols described above. We encountered NGF amplification difficulties, however, for the alligator and the two oscine passerines, the bulbul and sunbird, and did not obtain NGF sequences from these taxa. This more variable amplification success for NGF is probably related to the much greater magnitude of proregion amino acid variation at that locus. Owing to the absence of the alligator, bulbul, and sunbird NGF sequences, most of our phylogenetic analyses were based only on BDNF and NT-3 sequences.

### Sequence Alignment

Complementary sequences from each individual were aligned using Sequencer 3.1.1 (Gene Codes Corporation) and the chromatograms scanned by eye. All sequences used in phylogenetic analysis were confirmed by double-stranded sequencing, and all have been deposited in GenBank or EMBL nucleotide databases. Accession numbers listed here are in the order BDNF, NT-3, NGF: *Struthio camelus*, AF416632, SCA316235, SCA316265; *Anas platyrhynchos*, AF416633, APL316236, APL316266; *Falco sparverius*, AF416634, FSP316237, FSP316267; *Halcyon malimbica*, AF416635, HMA316238, HMA316261; *Coeligena torquata*, AF416636, CTO316239, CTO316262; *Chloropipo holochlora*, AF416637, CHO316240, CHO316263; and *Mionectes striaticollis*, AJ416888, AJ416887, MST316264. For the following three species, NGF was not sequenced (the order is BDNF, NT-3): *Alligator mississippiensis*, AF416634, AMI316234; *Andropadus latirostris*, AF416639, ALA316242; and *Nectarinia olivacea*, AF416640, NOL316243.

As described below, indels were rare and the sequences were aligned manually. In comparisons that included the alligator and all 10 avian taxa, no indels were present in the 600 nucleotides of the BDNF coding sequence, whereas two indels were present within the 591 nucleotides of NT-3 coding sequences. One of these NT-3 indels was a one-codon deletion restricted to the single coraciiforme taxon, the kingfisher *H. malimbica*. The other NT-3 indel was a one-codon deletion found exclusively in the two suboscine passerines, the manakin and flycatcher. In comparisons of NGF sequences that included eight avian taxa, we found only a single indel, a two-codon deletion present only in the suboscine passerine manakin *C. holochlora*.

### Phylogenetic Relationships

To explore whether phylogenetic reconstructions based on the neurotrophin gene sequences were similar to those generated using

other types of markers, we used maximum likelihood (ML) and maximum parsimony (MP) techniques to reconstruct relationships among taxa. Except where otherwise noted, analyses were conducted using PAUP\**b.8* (Swofford 1999). These reconstructions were rooted using the alligator outgroup sequences.

ML analyses were conducted using MrBayes 1.11 (Huelsenbeck 2000), in which a Markov chain Monte Carlo search was parameterized using the general time reversible model ( $nst = 6$ ), with site-specific rate variation partitioned by codon. Bayesian analyses were run for 250,000 generations and sampled every 1000 generations; inspection of the resulting ML scores suggested that stationarity was reached at approximately 10,000 generations in all searches, and we therefore conservatively discarded the topologies sampled from the first 20,000 generations. Consensus topologies from the remaining 2300 sampled trees were generated in PAUP\*. MP reconstructions were generated in PAUP\* via exhaustive searches, with all nucleotide substitutions weighted equally and with length variable sites excluded; MP bootstrap analyses were conducted for 1000 replications.

As discussed below, these reconstructions identified a single topology that was congruent with previous hypotheses of avian relationships based on independent molecular markers (e.g., DNA–DNA hybridization [Sibley and Ahlquist 1990]).

### Comparisons of Rate Variation

We used three complementary methods to estimate levels of functional constraint in the pro and mature regions of the three neurotrophin loci. In all cases, comparisons were restricted to the eight avian taxa from which sequences of all three loci were obtained. All three analytical methods gave broadly congruent results.

First, we employed a simple but likely biased parsimony-based approach by importing the topology shown in Fig. 2 into the program MacClade (Maddison and Maddison 1992) and determining the minimum number of steps required to map the observed synonymous substitutions onto that tree. We quantified amino acid variation in the pro and mature regions of each gene in a similar way by determining the minimum number of steps required to map the observed amino acid changes onto that same topology. These parsimony-based character reconstructions provide estimates of the minimum number of nucleotide or amino acid changes that have occurred since these eight extant lineages shared a common ancestor, but this method is likely to underestimate the actual number of changes, particularly for classes of silent substitutions where some sites have likely undergone multiple changes. Second, we used the topology-based, maximum likelihood approach implemented in the CODEML program in the PAML package (Yang 1997) to estimate the mean  $d_N/d_S$  ratio ( $w$ ) for each gene region. The search model was parameterized using the topology shown in Fig. 2, and we assumed a single  $d_N/d_S$  for each locus. Third, we used MEGA 2.01 (Kumar et al. 2001) to calculate  $d_N/d_S$  ratios from the aligned sequences using the Nei and Gojobori (1986) method as modified by Ina (1995) to include differing rates of transversion and transition substitutions.

## Results and Discussion

### Nucleotide and Amino Acid Variation in Avian BDNF and NT-3

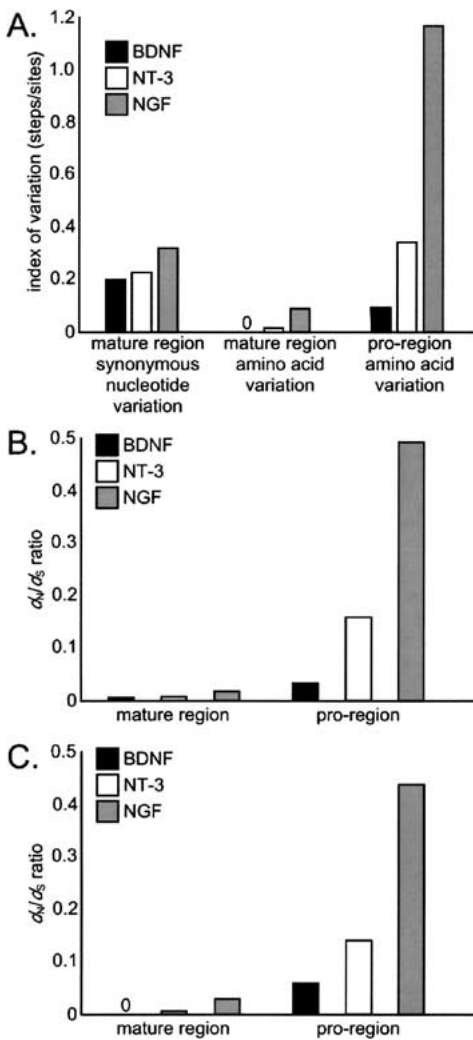
An analysis of the nucleotide sequences of each neurotrophin locus showed high levels of similarity among avian species and between the birds and the

### Support for Phylogenetic Framework

As our comparisons of locus-specific patterns of molecular variation depend upon the phylogeny used to estimate the minimum number of nucleotide or amino acid substitutions at each locus, we examined whether phylogenetic reconstructions based on the neurotrophin sequences are congruent with one another and with previous reconstructions based on other markers. Our analyses of the combined NT-3 and BDNF sequences from 10 avian taxa suggest that nucleotide sequences of these loci are useful phylogenetic markers that provide broadly congruent estimates of avian phylogenetic relationships.

Phylogenetic reconstructions that included neurotrophin sequences from 10 avian taxa and the alligator outgroup all identified broadly congruent and generally well-supported trees (Fig. 2, Table 1). This topology was also obtained in reconstructions based on single gene sequences (Table 1). Most nodes in the tree shown in Fig. 2 also received consistently high support across reconstructions employing different subsets of the NT-3 and BDNF sequences and across different reconstruction techniques (Table 1). Topological features receiving universally high support included (1) the placement of the chicken/duck "galloanseriform" lineage basal to the remaining Neognathae, a relationship consistent with studies based on other molecular markers (e.g., Sibley and Ahlquist 1990; Caspers et al. 1997; Groth and Barrowclough 1999; Mindell et al. 1999; van Tuinen et al. 2000); (2) the monophyly of the four passerine taxa; and (3) the sister relationships of the bulbul/sunbird and the manikin/flycatcher within the passerine clade. The only region of the tree receiving consistently weak support involved the relationships among the four neognath lineages represented by the kestrel, kingfisher, hummingbird, and passerines. These lineages are each representatives of different avian orders in traditional taxonomic classifications, and there is growing molecular phylogenetic evidence (e.g., Sibley and Ahlquist 1990; Groth and Barrowclough 1999; van Tuinen et al. 2000) that these and other Neognathae lineages arose during a relatively short period of extensive radiation subsequent to the split between the galloanseriformes and the remaining neognaths.

Comparisons of phylogenetic analyses that included all nucleotide sites with those that included only amino acid-invariant codons indicated that nucleotide substitutions at synonymous codon positions contributed most of the phylogenetic signal to the reconstructions. In particular, changes at synonymous sites do not appear to be more phylogenetically noisy than changes at nonsynonymous sites, as would be expected if synonymous sites had high levels of



**Fig. 3.** Contrasting levels of amino acid constraint at three avian neurotrophin loci. **A** A parsimony-based index of variation calculated by dividing the number of steps required to map the observed nucleotide or amino acid changes onto the topology depicted in Fig. 2, divided by the total number of nucleotide sites (*left-hand* comparison) or amino acids (two *right-hand* comparisons) within a given protein region. **B** Nonsynonymous:synonymous substitution ratios ( $d_N/d_S$ ) estimated via a topology-based, maximum likelihood approach (Yang 1997). **C**  $d_N/d_S$  ratios estimated via the modified (Ina 1995) method of Nei and Gojobori (1986). In the *bottom* two panels, high  $d_N/d_S$  ratios suggest relaxed selective constraint, whereas low ratios indicate strong selective constraint.

alligator outgroup. Figure 1 illustrates the alignment of the predicted protein sequences from the nucleotide sequences. The protein sequences of the mature regions of BDNF are identical among all the birds and the alligator, with only six amino acid substitutions separating the birds and human in the region we sequenced. Similarly, in the mature NT-3, the predicted protein sequence of alligator is identical to that of the chicken, but there are some conservative amino acid changes among the four passerine birds. NGF exhibited much greater variation among the avian taxa we examined (Fig. 1).

**Table 1.** Support indices for the topology shown in Fig. 2 under different reconstruction techniques

| Node | Bayesian ML |                  |                |                | Heuristic ML | NJ  | MP  |
|------|-------------|------------------|----------------|----------------|--------------|-----|-----|
|      | All sites   | Synonymous sites | NT-3 gene only | BDNF gene only |              |     |     |
| A    | 100         | 99               | 100            | 62             | 95           | 99  | 99  |
| B    | 100         | 100              | 78             | 96             | 95           | 97  | 86  |
| C    | 100         | 100              | 100            | 84             | 99           | 97  | 99  |
| D    | 69          | —                | 72             | —              | 50           | 83  | —   |
| E    | 87          | 82               | —              | 82             | —            | 57  | 69  |
| F    | 100         | 100              | 100            | 97             | 100          | 100 | 100 |
| G    | 100         | 100              | 100            | 99             | 100          | 100 | 100 |
| H    | 88          | 96               | 74             | 65             | 88           | 100 | 91  |

nucleotide saturation and exhibited concomitantly high levels of homoplasy. For example, posterior probabilities in the ML analysis restricted to amino acid-invariant codons were very similar to those in the ML analysis that included all codons, despite the potential loss of information caused by excluding the 43 variable codons. Similarly, in the MP analyses, measures of character consistency for the shortest trees based on the complete sequences (RCI = 0.423) and on the invariant amino acids were nearly identical (RCI = 0.400).

#### *Molecular Evolution of Avian Neurotrophin Genes*

Within each of the three neurotrophin loci, we noted large but expected differences in relative amino acid variation between the proregions and the mature regions: for all three loci, a much higher proportion of codons varied in the proregion than in the mature region (Fig. 1). Amino acid conservation was strict in all three mature regions: the mature region of the BDNF gene showed no amino acid variation, two codons of the NT-3 gene varied, and seven codons of the NGF gene varied (see also Fig. 3). In contrast, much greater variation in amino acid constraint was present among the proregions of the three loci: the BDNF proregion varied at 12 codons; the NT-3 proregion was intermediate, with variation at 29 codons; and NGF proregion had the highest variation, at 53 codons.

This pattern across loci of differing variability in the pro and mature regions is interesting primarily because it provides evidence for differences in the functions of the protein sequences encoded by the three proregions. The high amino acid conservation in the mature regions of each locus is likely to result from strong functional selection on the mature protein, but it is less obvious why the three proregions would show different levels of functional constraint. One explanation is that the three loci could be subject to different underlying mutation rates, but the similarity in the rates of synonymous nucleotide substitutions across all three loci (Fig. 3) argues against this possibility. The similarities in structure and composition

of these three neurotrophin genes make them an unusually tractable system for comparing rates and patterns of molecular evolution among loci. In particular, we can test for intrinsic mutation rate differences among loci by comparing locus-specific rates of nucleotide substitutions at synonymous sites, because those sites are presumably free from selection on function at the amino acid level. We chose to base our estimates of underlying mutation rates on nucleotide changes at synonymous nucleotide positions within the mature regions of each protein. These degenerate sites are highly comparable across loci because the mature regions of the three loci have similar general structures, including a number of blocks of amino acids that are semiconserved across loci (Fig. 1); have similar codon use and therefore similar proportions of degenerate nucleotide sites; are nearly identical in length; and have nearly identical and highly even base compositions. All three loci have very similar rates of synonymous substitution (Fig. 3A).

All three mature regions also have similar (and very low)  $d_N/d_S$  ratios (Fig. 3B and C), indicating that they are all subject to high levels of selective constraint. In contrast, we found an order-of-magnitude difference in the  $d_N/d_S$  ratios at the three proregions (Fig. 3): the BDNF proregion is nearly as conserved as the NGF mature region, whereas the highly variable NGF proregion is under very relaxed constraint. We speculate that the BDNF proregion (and perhaps the intermediately variable NT-3 proregion) may have a function in addition to its role in protein folding. For example, the conservation of the proregions inversely correlates with the efficiency of proteolytic cleavage of the neurotrophins (NGF proteolytic processing is considerably more efficient than BDNF proteolytic processing [Mowla et al. 2001; Carlos Ibañez, personal communication]). Thus, in vivo the ratio of proBDNF to mature BDNF is likely to be higher than the corresponding pro:mature NGF ratio. In addition, recent studies have shown that proBDNF mediates TrkB phosphorylation and is thus biologically active, and that BDNF differs from NGF and NT-3 in its secretory

pathway (Mowla et al. 2001). It seems likely that proBDNF may have a function that differentiates it from mature BDNF, with proteolytic processing being a possible mechanism that regulates the disparate biological activities. Lee et al. (2001) found that the proforms of NGF and BDNF are expressed *in vivo*, secreted, and cleaved extracellularly and that the proform of NGF preferentially activates the p75<sup>NTR</sup>. In addition, recent studies have shown that proNGF induces p75<sup>NTR</sup>-mediated apoptosis of oligodendrocytes following spinal cord injury (Beattie et al. 2002). Based on the patterns of variation among the three avian loci, we suggest that BDNF, with its apparent high degree of proregion functional constraints, may be the most likely candidate to operate in its proform and to be regulated by proteolytic processing *in vivo*. The apparent absence of the NT-4 protein in chickens (C. Ibañez and F. Hallböök, personal communication) further suggests that BDNF may play a compensatory role in the development of the avian nervous system.

**Acknowledgments.** We thank Tom Smith for samples and support at the Center for Tropical Research, Robert Ricklefs and Lisa Tell for additional samples, and Sonya Clegg, Finn Hallböök, and Enrique Lessa, for comments on an early version of the manuscript. We thank Carlos Ibañez for insightful suggestions and Rasmus Nielsen for advice on data analysis. R.N.M.S. was supported by a Research Infrastructure in Minority Institutions award from the National Center for Research Resources with funding from the Office of Research on Minority Health, National Institutes of Health Grant 5 P20 RR11805. R.N.M.S. was also supported by an NIGMS MORE Institutional Research and Academic Career Development Award to UC Davis and San Francisco State University, (<http://prof.ucdavis.edu/>; Grant K12GM00679). This work was also funded by Grants NSF-DEB9726425 and NSF-IRCEB9977072.

## References

- Beattie MS, Harrington AW, Lee R, Kim JY, Boyce SL, Longo FM, Bresnahan JC, Hempstead BL, Yoon SO (2002) ProNGF induces p75-mediated death of oligodendrocytes following spinal cord injury. *Neuron* 36:375–386
- Caspers G-J, Uit de Weerd U, Wattel J, de Jong WW (1997) Alpha-crystallin sequences support a Galliform/Anseriform clade. *Mol Phylogenet Evol* 2:185–188
- Chao MV (2000) Trophic factors: An evolutionary cul-de-sac or door into higher neuronal function? *J Neurosci* 59:353–355
- Chao MV, Bothwell M (2002) Neurotrophins, to cleave or not to cleave. *Neuron* 33:9–12
- Dobrowsky RT, Carter BD (2000) p75 Neurotrophin receptor signaling: Mechanisms for neurotrophic modulation of cell stress? *J Neurosci* 61:237–243
- Ebendal T, Larhammar D, Persson H (1986) Structure and expression of the chicken beta nerve growth factor gene. *EMBO J* 5:1483–1487
- Friedman WJ, Greene A (1999) Neurotrophin signaling via Trks and p75. *Exp Cell Res* 25:131–142
- Goldman N, Anderson JP, Rodrigo AG (2000) Likelihood-based tests of topologies in phylogenetics. *Syst Biol* 49:652–670
- Groth JG, Barrowclough GF (1999) Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol Phylogenet Evol* 12:115–123
- Hallböök F (1999) Evolution of the vertebrate neurotrophin and Trk receptor gene families. *Curr Opin Neurobiol* 9:616–621
- Hallböök F, Lundin L-G, Kullander K (1998) *Lamtera fluviatilis* neurotrophin homolog, descendent of a neurotrophin ancestor, discloses the early molecular evolution of neurotrophins in the vertebrate subphylum. *J Neurosci* 18:8700–8711
- Heinrich G, Lum T (2000) Fish neurotrophins and Trk receptors. *Int J Dev Neurosci* 18:1–27
- Huelsenbeck JP (2000) MrBayes: Bayesian inference of phylogeny. Distributed by the author. Department of Biology, University of Rochester, Rochester, NY
- Ibañez CF (2002) Jekyll-Hyde neurotrophins: The story of proNGF. *Trends Neurosci* 25:284–286
- Ibañez CF, Ebandal T, Barbany G, Murray-Rust J, Blundell TL, Persson H (1992) Disruption of the low affinity receptor-binding site in NGF allows neuronal survival and differentiation by binding to the *trk* gene product. *Cell* 69:329–341
- Ibañez CF, Ilag L, Murray-Rust J, Persson H (1993) An extended surface of binding to Trk tyrosine kinase receptors in NGF and BDNF allows the engineering of a multifunctional pan-neurotrophin. *EMBO J* 12:2281–2293
- Ina Y (1995) New methods for estimating the numbers of synonymous and nonsynonymous substitutions. *J Mol Evol* 40:190–226
- Jones KR, Reichardt LF (1990) Molecular cloning of a human gene that is a member of the nerve growth factor family. *Proc Natl Acad Sci USA* 87:8060–8064
- Kaplan DR, Miller D (2000) Neurotrophin signal transduction in the nervous system. *Curr Opin Neurobiol* 10:381–391
- Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics* 17:1244–1245
- Lee R, Kermani P, Teng KK, Hempstead BL (2001) Regulation of cell survival by secreted proneurotrophins. *Science* 294:1945–1948
- Maddison WP, Maddison DR (1992) MacClade: Analysis of phylogeny and character evolution, version 3.06. Sinauer Associates, Sunderland, MA
- Mindell DP, Sorenson MD, Dimcheff DE, Hasegawa M, Ast JC, Yuri T (1999) Interordinal relationships of birds and other reptiles based on whole mitochondrial sequences. *Syst Biol* 48:138–152
- Mowla SJ, Farhadi HF, Pareek S, Atwal JK, Morris SJ, Seidah NG, Murphy RA (2001) Biosynthesis and post-translational processing of the precursor to brain-derived neurotrophic factor. *J Biol Chem* 276:12660–12666
- Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3:418–426
- Patapoutian A, Reichardt LF (2001) Trk receptors: Mediators of neurotrophin action. *Curr Opin Neurobiol* 11:272–280
- Rattenholl A, Ruoppolo M, Flagiello A, Monti M, Vinci F, Marino G, Lilie H, Schwarz E, Rudolph R (2001a) Pro-sequence assisted folding and disulfide biformation of human nerve growth factor. *J Mol Biol* 305:523–533
- Rattenholl A, Lilie H, Grossmann A, Stern A, Schwarz E, Rudolph R (2001b) The pro-sequence facilitates folding of human nerve growth factor from *Escherichia coli* inclusion bodies. *Eur J Biochem* 268:3296–3303
- Sibley CG, Ahlquist JE (1990) Phylogeny and classification of birds. Yale University Press, New Haven, CT
- Suter U, Heymach JV Jr, Shooter EM (1991) Two conserved domains in the NGF propeptide are necessary and sufficient for

- the biosynthesis of correctly processed and biologically active NGF. *EMBO J* 10:2395–2400
- Swofford DL (1999) PAUP\*: Phylogenetic analysis using parsimony (\*and other methods), version 4.0b2. Sinauer, Sunderland, MA
- Urfer R, Twoulfas P, Soppet D, Escandon E, Parada LF, Presta LG (1994) The binding epitopes of neurotrophin-3 to its receptors trkC and gp75 and the design of a multifunctional human neurotrophin. *EMBO J* 13:5806–5909
- Van Tuinen M, Sibley CG, Hedges SB (2000) The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. *Mol Biol Evol* 17:451–457
- Wetmore A (1960) A classification for the birds of the world. *Smithson Misc Collect* 139(11)
- Yang Z (1997) PAML: A program for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13:555–556