

## Product datasheet for **RC402139**

### NOTCH3 (NM\_000435) Human Mutant ORF Clone

#### Product data:

Product Type:	Mutant ORF Clones
Product Name:	NOTCH3 (NM_000435) Human Mutant ORF Clone
Mutation Description:	C440R
Affected Codon#:	440
Affected NT#:	1318
Nucleotide Mutation:	NOTCH3 Mutant (C440R), Myc-DDK-tagged ORF clone of Homo sapiens notch 3 (NOTCH3) as transfection-ready DNA
Effect:	CADASIL
Symbol:	NOTCH3
Synonyms:	CADASIL; CADASIL1; CASIL; IMF2; LMNS
E. coli Selection:	Kanamycin (25 ug/mL)
Mammalian Cell Selection:	Neomycin
Vector:	pCMV6-Entry (PS100001)
Tag:	Myc-DDK
ACCN:	NM_000435
ORF Size:	6963 bp
Restriction Sites:	SgfI-MluI
ORF Nucleotide Sequence:	>RC402139 representing NM_000435 Red=Cloning site Blue=ORF Green=Tags(s)

TTTTGTAATACGACTCACTATAGGGCGGCCGGGAATTCGTCGACTGGATCCGGTACCGAGGAGATCTGCC  
GCC**CGATCGCC**

ATGGGGCCGGGGCCCGTGGCCGCCGCCCGCGTCGCCGATGTCGCCGCCACCGCCACCGCCACCCG  
TGCGGGCGCTGCCCTGCTGCTGCTGCTAGCGGGCCGGGGGCTGCAGCCCCCTTGCCTGGACGGAAG  
CCCGTGTGCAAATGGAGGTCGTTGCACCCAGCTGCCCTCCGGGAGGCTGCCTGCCTGTGCCCGCTGGC  
TGGGTGGGTGAGCGGTGTCAGCTGGAGGACCCCTGCTCACTCAGGCCCTGTGCTGGCCGTGGTGTCTGCC  
AGAGTTCAGTGGTGGCTGGCACCGCCGATTCTCATGCCGGTGCCCCGTGGCTCCGAGGCCCTGACTG  
CTCCCTGCCAGATCCCTGCCTCAGCAGCCCTTGTGCCACGGTGCCCGCTGCTCAGTGGGGCCCGATGGA  
CGCTTCTCTGCTCCTGCCACCTGGTACCAGGGCCGACGCTGCCGAAGCGACGTGGATGAGTGCCGGG



[View online »](#)

TGGGTGAGCCCTGCCGCCATGGTGGCACCTGCCTCAACACACCTGGCTCCTCCGCTGCCAGTGTCCAGC  
TGGCTACACAGGGCCACTATGTGAGAACCCCGCGGTGCCCTGTGCACCCTCACCATGCCGTAACGGGGG  
ACCTGCAGGCAGAGTGGCGACCTCACTACGACTGTGCCTGTCTTCTGGGTTTGGGGTTCAGAAATTGTG  
AAGTGAACGTGGACGACTGTCCAGGACACCGATGTCTCAATGGGGGACATGCGTGGATGGCGTCAACAC  
CTAACTGCCAGTGCCTCCTGAGTGGACAGGCCAGTTCGCACGGAGGACGTGGATGAGTGTGAGTGTG  
CAGCCCAACGCCTGCCACAATGGGGTACCTGCTTCAACACGCTGGGTGGCCACAGCTGCGTGTGTGCA  
ATGGCTGGACAGGCGAGAGCTGCAGTCAGAATATCGATGACTGTGCCACAGCCGTGCTTCCATGGGGC  
CACCTGCCATGACCGCGTGGCTTCTTTCTACTGTGCCTGCCCATGGGCAAGACTGGCCTCCTGTGTAC  
CTGGATGACGCCTGTGTGAGCAACCCCTGCCACGAGGATGCTATCTGTGACACAAATCCGGTGAACGGC  
GGGCCATTTGCACCTGTCTCCCGGCTTACGGGTGGGCATGTGACCAGGATGTGGACGAGTGTCTAT  
CGGCGCAACCCCTGCGAGCACTGGGCAGGTGCGTGAACACGACAGGGCTCCTTCTGTGCCAGTGGGT  
CGTGGCTACACTGGACCTCGCTGTGAGACCGATGTCAACGAGTGTCTGTGCGGGCCCCGCCAAACAGG  
CCACGTGCCTCGACCGCATAGGCCAGTTCACCTGTATCTGTATGGCAGGCTTACAGGAACCTATTGCGA  
GGTGGACATTGACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG  
TTCAGTGCACCTGCCCTCGGGCTTACAGCGCTCCACGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG  
CGCCCTGCAGGAATGGCGCAAAATGCGTGGACAGCCCGATGGCTACGAGTGGCCTGTGCCGAGGGCTT  
TGAGGGCACGCTGTGTGATCGCAACGTGGACGACTGCTCCCTGACCCATGCCACCATGGTGCCTGCGT  
GATGGCATCGCCAGCTTCTCATGTGCCTGTGCTCCTGGCTACACGGGCACACGCTGCGAGAGCCAGGTGG  
ACGAATGCCCGACCCAGCCCTGCCGCCATGGCGGCAAAATGCCTAGACCTGGTGGACAAGTACCTTGCCG  
CTGCCCTTCTGGGACCACAGGTGGAATGCGAAGTGAACATTGACGACTGTGCCAGCAACCCCTGCACC  
TTTGGAGTCTGCCGTGATGGCATCAACCGCTACGACTGTGTCTGCCAACCTGGCTTACAGGGCCCTTT  
GTAACGTGGAGATCAATGAGTGTGCTTCCAGCCATGCGGCGAGGGAGGTTCTGTGTGGATGGGAAAA  
TGCTTCCGCTGCCTTGCCTCGCTGGCTCCTTGCCTCCTGCTCCTGCTCCTGCTCCTGCTCCTGCTCCTG  
CATGAGCCCTGCAGTACGGCATCTGCTATGATGCACCTGGCGGGTCCGCTGTGTGTGAGCCTGGCT  
GGAGTGGCCCCGCTGCAGCCAGAGCCTGGCCCCGAGACGCTGTGAGTCCCAGCCGTGCAGGGCCGTGG  
GACATGCAGCAGCGATGGAATGGGTTTCCACTGCACCTGCCCGCTGGTGTCCAGGGACGTGAGTGTGAA  
CTCCTCTCCCCCTGCACCCGAACCCCTGTGAGCATGGGGCCGCTGCGAGTCTGCCCTGGCCAGCTGC  
CTGTCTGCTCCTGCCCCAGGGTGGCAAGGCCACGATGCCAGCAGGATGTGGACGAGTGTGCTGGCC  
CGCACCTGTGGCCCTCATGGTATCTGCACCACTGGCAGGGAGTTTCCAGTGCACCTGCCATGGAGGG  
TACACTGGCCCTTCTGCGATCAGGACATCAATGACTGTGACCCCAACCCATGCCTGAACGGTGGCTCGT  
GCCAAGACGGCGTGGCTCCTTTCTGCTCCTGCCTCCTGGTTTCGCCGGCCACGATGCGCCCGCA  
TGTGGATGAGTGCCTGAGCAACCCCTGCGGCCCGGGCACCTGTACCGACACAGTGGCCTCCTTACCTGC  
ACCTGCCCGCCAGGCTACGGAGGCTTCCACTGCGAACAGGACCTGCCCGACTGCAGCCCCAGCTCCTGCT  
TCAATGGCGGGACCTGTGTGGACGGCGTGAACCTCGTTCAGCTGCCTGTGCCGTCCCGGCTACACAGGAGC  
CCACTGCCAACATGAGGCAGACCCCTGCCTCTCGCGCCCTGCCTACACGGGGGGCGTCTGACGCGCCGC  
CACCTGGCTTCCGCTGCACCTGCCTCGAGAGCTTACGGGCCCGCAGTGCAGACGCTGGTGGATTGGT  
GCAGCCGCCAGCCTTGTCAAACGGGGTGCCTGCGTCCAGACTGGGGCTATTGCCCTTGTCCCCTGG  
ATGGAGCGGACGCTTGTGACATCCGAAGCTTGCCTGCAGGGAGGCCGAGCCAGATCGGGGTGCGG  
CTGGAGCAGCTGTGTGAGCGGGTGGCAGTGTGTGGATGAAGACAGCTCCCACTACTGCGTGTGCCAG  
AGGGCCGTAAGTGGTACCCACTGTGAGCAGGAGGTGGACCCTGCTTGGCCAGCCCTGCCAGCATGGGGG  
GACCTGCCGTGGCTATATGGGGGCTACATGTGTGAGTGTCTTCTGGTACAATGGTGATAACTGTGAG  
GACGAGTGGACGAGTGTGCCTCCAGCCCTGCCAGCACGGGGTTTATGCATTGACCTCGTGGCCCGCT  
ATCTCTGCTCCTGTCCCCAGGAACGCTGGGGTGTCTGCGAGATTAATGAGGATGACTGCGGCCAGG  
CCCACCGTGGACTCAGGGCCCCGCTGCCTACACAATGGCACCTGCGTGGACCTGGTGGTGGTTTCCGC  
TGCACCTGTCCCCAGGATACACTGGTTTGCCTGCGAGGCAGACATCAATGAGTGTGCTCAGTGCCT  
GCCACGCGGCACACCCGGGACTGCCTGCAGGACCCAGGCGGAGGTTTCCGTTGCCCTTGTGATGCTGG  
CTTCTCAGGCTCCTGCTGTGAGTGTGCTGCTCCTGCGAGTCCCAGCATGCCAGCATGGAGGCCAG  
TGCCGCTTAGCCCGGTCTGGGGTGGGCTGACCTTACCTGTCACTGTGCCAGCCGTTCTGGGGT  
CGCGTTGCGAGCGGGTGGCGGCTCCTGCCGGGAGCTGCAGTGCCCGGTGGGCGTCCCATGCCAGCAGAC  
GCCCGCGGGCCGCGCTGCGCCTGCCCCAGGGTTGTGCGGACCCCTCCTGCCGAGCTTCCCGGGTGC  
CCGCGGGGGCCAGCAACGCCAGTGCAGCGGCCCCCTGTCTCCAGGGGGCTCCTGCCGCCCGCGC  
CGCTCGCGCCCTTCTCCGCTGCGCTTGCAGCAGGGCTGGACCGGGCCGCGCTGCGAGGCGCCCGCCG

GGCACCCGAGGTCTCGGAGGAGCCGCGGTGCCCGCGCCGCCTGCCAGGCCAAGCGCGGGGACCAGCGC  
 TGCGACCCGAGTGCAACAGCCAGGCTGCGGCTGGGACGGCGGCGACTGCTCGTGAGCGTGGGCGACC  
 CCTGGCGGCAATGCGAGGCGCTGCAGTGTGGCGCCTTTCAACAACAGCCGCTGCGACCCCGCCTGCAG  
 CTCGCCCGCTGCCTCTACGACAACCTCGACTGCCACGCCGGTGGCCGCGAGCGCACTTGAACCCGGTG  
 TACGAGAAGTACTGCGCCGACCACTTTGCCAGCGGCGCTGCGACCAGGGCTGCAACACGGAGGAGTGCG  
 GCTGGGATGGGCTGGATTGTGCCAGCGAGGTCCCGCCCTGCTGGCCCGCGCGTGTGGTGTCTCACAGT  
 GCTGTGCCGCCAGAGGACTACTGCGTTCAGCGCCGACTTTCTGCAGCGGCTCAGCGCCATCTCGCGC  
 ACCTCGCTGCGCTTCCGCTGGACGCGCACGGCCAGGCCATGGTCTTCCCTTACCACCGCCTAGTCCTG  
 GCTCCGAACCCCGGGCCCGTCCGGAGCTGGCCCCGAGGTGATCGGCTCGGTAGTAATGTGGAGATTGA  
 CAACCGGCTCTGCCTGCAGTGCCTGAGAATGATCACTGCTTCCCGATGCCAGAGCGCCGCTGACTAC  
 CTGGGAGCGTTGTGAGCGGTGGAGCGCTGGACTTCCCGTACCCACTGCGGGACGTGCGGGGGAGCCGC  
 TGGAGCCTCCAGAACCAGCGTCCCGTGTGCCACTGCTAGTGGCGGGCGTGTCTGTGTGGTGTGAT  
 TCTCGTCTGGGTGTGATGGTGGCCCGCGCAAGCGGAGCACAGCACCCCTGTGGTTCCTGAGGGCTTC  
 TCACTGCACAAGGACGTGGCCTCTGGTACAAGGGCCGGCGGGAACCCGTGGGCCAGGACGCGTGGGCA  
 TGAAGAACATGGCCAAGGGTGAAGCCTGATGGGGAGGTGGCCACAGACTGGATGGACACAGAGTGCC  
 AGAGGCCAAGCGCTAAAGGTAGAGGAGCCAGGCATGGGGCTGAGGAGGCTGTGGATTCCGCTCAGTGG  
 ACTCAACACCATCTGGTGTGTGACTCCCGTGGCACCAGCCATGGCACTGACACCACCACAGGGCG  
 ACGCAGATGCTGATGGCATGGATGTCAATGTGCGTGGCCAGATGGCTTACCCCGCTAATGCTGGCTTC  
 CTTCTGTGGGGGGCTCTGGAGCCAATGCCAAGTGAAGAGGATGAGGCAGATGACACATCAGTACATC  
 ATCTCCGACCTGATCTGCCAGGGGGCTCAGCTTGGGGCACGGACTGACCGTACTGGCGAGACTGCTTTGC  
 ACCTGGCTGCCGTTATGCCCGTGTGATGCAGCAAGCGGCTGCTGGATGCTGGGGCAGACCAATGCG  
 CCAGGACCACTCAGGCCGCACTCCCTGCACACAGCTGTCACAGCCGATGCCAGGGTGTCTTCCAGATT  
 CTCATCCGAAACCGCTTACAGACTTGGATGCCCGCATGGCAGATGGCTCAACGGCACTGATCTGGCG  
 CCGCCTGGCAGTAGAGGGCATGGTGAAGAGCTCATGCCAGCCATGCTGATGTAATGCTGTGGATGA  
 GCTTGGGAAATCAGCCTTACACTGGGCTGCGGCTGTGAACAACGTGGAAGCCACTTTGGCCCTGTCAA  
 AATGGAGCCAATAAGGACATGCAGGATAGCAAGGAGGAGACCCCTATTCTGGCCGCCGCGAGGGCA  
 GCTATGAGGCTGCCAAGCTGCTGTTGGACCACTTTGCCAACCGTGAATCACCAGCACCTGGACAGGCT  
 GCCCGGGGACGTAGCCAGGAGAGACTGCACCAGGACATCGTGCCTTGTGGATCAACCAGTGGGCC  
 CGCAGCCCCCGGTCCACAGGCTGGGGCTCTGCTCTGTCTCCAGGGCCCTTCTCCCTGGCCTCA  
 AAGCGCACAGTCGGGGTCCAAGAAGAGCAGGAGGCCCGGGAAGCGGGGCTGGGGCCGAGGGCC  
 CCGGGGCGGGCAAGAAGCTGACGCTGGCCTGCCGGGCCCTGGCTGACAGCTCGGTACGCTGTGCG  
 CCGTGGACTCGCTGGACTCCCCGCGGCTTTCGGTGGGCCCTGCTTCCCCTGGTGGCTTCCCCTTG  
 AGGGGCCATATGCAGCTGCCACTGCCACTGCAGTGTCTTGGCACAGCTTGGTGGCCAGGCCGGGCGG  
 TCTAGGGCGCCAGCCCCCTGGAGGATGTGTAAGTACTCAGCCTGGGCTGCTGAACCTGTGGCTGTGCCCTC  
 GATTGGGCCCGGCTGCCCCACCTGCCCTCCAGGCCCTCGTTCTGCTGCCACTGGCGCCGGGACCC  
 AGCTGTCAACCCAGGGACCCCGTCTCCCGCAGGAGCGGCCCGCCTTACCTGGCAGTCCCAGGACA  
 TGGCGAGGAGTACCCGGCGGCTGGGGCACACAGCAGCCCCAAAGGCCCGCTTCTGCGGGTCCCAGT  
 GAGCACCTTACCTGACCCATCCCCGAATCCCCTGAGCACTGGGCCAGCCCCACCTCCCTCCCTCT  
 CAGACTGGTCCGAATCCAGCCTAGCCAGCCACTGCCACTGGGGCCATGGCCACCACCCTGGGGCACT  
 GCCTGCCAGCCACTTCCCTGTCTGTTCCAGCTCCCTGCTCAGGCCAGACCAGCTGGGGCCCCAG  
 CCGGAAGTTACCCCAAGAGGCAAGTGTGGCC

AGCGGACCGACGCGTACGCGCCGCTCGAGCAGAACTCATCTCAGAAGAGGATCTGGCAGCAAATGATATCC  
 TGGATTACAAGGATGACGACGA TAAGTTTAA

**Protein Sequence:** >RC402139 representing NM\_000435  
 Red=Cloning site Green=Tags(s)

MGPGARRRRRRRPMSPPPPPPVRLPLLLLLAGPGAAAPPCLDGGSPCANGGRCTQLPSREAACLPPG  
 WVGERCQLEDPCSHGSPCAGRVCQSSVAVGTARFSCRCPRGFRGPDCSLPDPCLSSPCAHGARCSVGPDG  
 RFLCSCPPGYQGRSCRSDVDECRVGEPCRHHGTCLNTPGSFRCQCPAGYTGPLCENPAVPCAPSPCRNGG  
 TCRQSGDLTYDCACLPGFEGQNCENVVDDCPGHRCLNGGTCTVDGVNTYNCQCPPEWTGQFCTEDVDECQL  
 QPNACHNGGTCTFNTLGGHSCVCVNGWTGESCSQNIIDDCATAVCFHGATCHDRVASFYCACPMGKTGLLCH  
 LDDACVSNPCHEDAICDTPVNGRAICTCPPGFTGGACDQDVDECSIGANPCEHLGRCVNTQGSFLCQCG  
 RGYTGPRCETDVNECLSGPRRNQATCLDRIGQFTICIMAGFTGTyceVDIDEQSSPCVNGGVCKDRVNG  
 FSCTCPSGFSGSTCQLDVDECASTPCRNGAKCVDQPDGYECRAEGFEGTLCDRNVDSCDPDCHHGRCV  
 DGIASFSCACAPGYTGTRCESQVDECRSQPCRHHGKCLDLVDKYLRCRPSGTTGVNCEVNIIDDCASNPT  
 FGVCRDGINRYDCVCQPFTGPLCNVEINECASSPCGEGGSCVDGENFRCLCPPGSLPPLCLPPSHPCA  
 HEPCHSHGICYDAPGGFRVCPEPWSGPRCSQSLARDACESQPCRAGGTCSSDGMGFHCTCPPGVQGRQCE  
 LLSPCTPNPCEHGRCESAPGQLPVCSCPQGWQPRCQDQVDECAGPAPCGPHGICTNLAGFSCTCHGG  
 YTGPSCDQDINDCDPNPCLNGGSCQDGVGSFSCSCLPGFAGPRCARDVDECLSNPCGPGTCTDHVASFTC  
 TCPPGYGGFHCEQDLPDCSPSSCFNGGTCTVDGVNSFSCLCRPGYTGHAHCQHEADPCLSRPCLHGGVCSAA  
 HPGFRCTCLESFTGPQCQTLVDWCSRQCQNGGRCVQTGAYCLCPPGWSGRLCDIRSLPCREAAAQIGVR  
 LEQLCQAGGQCQVDESSHYCVCEPGRGTSHEQEVDPCLAQPCQHGGRGCRGMYGGYMCECLPGYNGDNCE  
 DDVDECASQPCQHGGSIDL VARYLCSPPGTLGVLCEINEDDCGPGPPLDSGPRCLHNGTCDVLDVGGFR  
 CTCPPGYTGLRCEADINECRSGACHAAHTRDCLQDPGGGFRCLCHAGFSGPRCQTVLSPCESQPCQHGQ  
 CRPSPGPGGGLTFTCHCAQPFWGPRCERVARSCRELQCPVGVPCQQTTPRGPRCACPPGLSGPSCRSPFGS  
 PPGASNASCAAAPCLHGGSCRPAPLAPFFRCACAQGWTPRCEAPAAAPEVSEEPRCRAACQAKRGDQR  
 CDRECNSPGCGWDGGDCSLSVGDPWRQCEALQCWRLFNNRCDPACSSPACLYDNFDCHAGGRERTCNPV  
 YEKYCADHFADGRCDQGCNTEECGWDLDCASEVPALLARGVLLVTVLLPPEELLRSSADFLQRLSAILR  
 TSLRFRLDAHQAAMVFPYHRPSPGSEPRARRELAPEVIGSVVMLEIDNRLCLQSPENDHCFDAQSAADY  
 LGALSAVERLDFPYPLRDVRGELEPEPEPSVPLPLLVAGAVLLLVLVLMVARRKREHSTLWFPEGF  
 SLHKDVASGHKGRREPVGQDALGMKNMAKGESLMGEVATDWMDECEPEAKRLKVEEPMGAEAAVDCRQW  
 TQHHLVAADIRVAPAMALTPPQGDADADGMDVNVVRGPDGFTPLMLASFCCGGALEPMPTEEDEADTSASI  
 ISDLICQGAQLGARTDRTGETALHLAARYARADAAKRLLDAGADTNAQDHSGRTPLHTAVTADAQGVFQI  
 LIRNRSTDL DARMADGSTALILAAARLAVEGMVEELIASHADVNAVDELGKSALHWAAVNNVEATLALLK  
 NGANKDMQDSKEETPLFLAAREGSYEAAKLLLDHFANREITDHLDRLPDVAQERLHQDIVRLLDQPSGP  
 RSPPGPHGLPPLLCPGAFPLGLKAAQSGSKSRPPGKAGLGPQGRGRGKLLTACPGPLADSSVTL  
 PVDSLDSRPFPGPPASPGGFLEGPYAAATATAVSLAQLGGPGRAGLGRQPPGGCVLSLGLLNPVAVPL  
 DWARLPPAPPGPSFLLPLAPGQQLLNPGTVPSPQERPPPYLAVPGHGEEYPAAGAHSSPPKARFLRVPS  
 EHPYLTPSPESPEHWASPPSLSDWSESTPSPATATGAMATTTGALPAQPLPLSVPSLAQAQTQLGPQ  
 PEVTPKRQVLA

SGP TRTRRLEQKLI SEEDLAANDILDYKDDDDKV

**Restriction Sites:** SgfI-MluI

**Cloning Scheme:**

**OTI Disclaimer:**

Due to the inherent nature of this plasmid, standard methods to replicate additional amounts of DNA in E. coli are highly likely to result in mutations and/or rearrangements. Therefore, OriGene does not guarantee the capability to replicate this plasmid DNA. Additional amounts of DNA can be purchased from OriGene with batch-specific, full-sequence verification at a reduced cost. Please contact our customer care team at [custsupport@origene.com](mailto:custsupport@origene.com) or by calling 301.340.3188 option 3 for pricing and delivery.

The molecular sequence of this clone aligns with the gene accession number as a point of reference only. However, individual transcript sequences of the same gene can differ through naturally occurring variations (e.g. polymorphisms), each with its own valid existence. This clone is substantially in agreement with the reference, but a complete review of all prevailing variants is recommended prior to use. [More info](#)

**OTI Annotation:**

This clone was engineered to express the complete ORF with an expression tag. Expression varies depending on the nature of the gene.

**Components:**

The ORF clone is ion-exchange column purified and shipped in a 2D barcoded Matrix tube containing 10ug of transfection-ready, dried plasmid DNA (reconstitute with 100 ul of water).

**RefSeq:**

[NP\\_000426](#)

**RefSeq Size:**

6963 bp

**RefSeq ORF:**

6966 bp

**Locus ID:**

4854

**Cytogenetics:**

19p13.12

**Domains:**

NL, EGF\_CA, ANK, EGF, EGF

**Protein Families:**

Druggable Genome

**Protein Pathways:**

Dorso-ventral axis formation, Notch signaling pathway

**MW:** 255.3 kDa

**Gene Summary:** This gene encodes the third discovered human homologue of the *Drosophila melanogaster* type I membrane protein notch. In *Drosophila*, notch interaction with its cell-bound ligands (delta, serrate) establishes an intercellular signalling pathway that plays a key role in neural development. Homologues of the notch-ligands have also been identified in human, but precise interactions between these ligands and the human notch homologues remains to be determined. Mutations in NOTCH3 have been identified as the underlying cause of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL). [provided by RefSeq, Jul 2008]