

Abb. 20: DNA- und Aminosäuresequenz der humanen LRRK1 (Fortsetzung)

1531 aaatgaagatggactgaaaacgaagcgtatgacctttttcacacacagaggtcgcacagcgtcccgaggactgagggcagcaagtgctggaa
 511 N E D G L K T K R I A F F T T R G R Q R S G T E A A S V L E

LRR#12

1621 tttccggccttctctaaagtgaagcttttggagtcctttgacctgaaacgaacacacagtcctccctcggcttggcctactgaa
 541 F P A F L S E S L E V L C L N D N H L D T V P P S V C L L K

LRR#13

1711 agottatcagagctctacttgggaaa caa cctggcctccgggagctcccctctgagctgggcaagctgggcaacctctggcagctggac
 571 S L S E L Y L G N N P G L R E L P P E L G Q L G N L W Q L D

LRR#14**Linker1**

1801 actgaagacctgacatcagcaatgtgcctgcagaatc caaaaagaaggcccacagcaatgctgtctta cctcgtgctcagctggc
 601 T E D L T I S N V P A E I Q K E G P K A M L S Y L R A Q L R

1891 aaagcggaa aagtgcaagctgatgaaatgatcatcgtgggtcccgcgccaagggcaagtcacacccctcctggagatcttacaagcggg
 631 K A E K C K L M K M I I V G P P R Q G K S T L L E I L Q T G

Roc-Domäne

1981 agggc ccc ccagggtggtgca tggaga ggc cac catcaggac cac caa gttggagctcca gag gccggctggctcaagagc caa ggt tga
 661 R A P Q V V H G E A T I R T T K W E L Q R P A G S R A K V E

2071 tccgtggagtcaacgtctgggacatcggggaccggccagcatggccac tgt caa cca gttcttctt cac gga caa ggc cct gta cgt g
 691 S V E F N V W D I G P A S M A T V N Q C F F T D K A L Y V

2161 gtggtctggaa cctggcgtggggga gga ggc cgtggc caa cctcca gttctggctgct caa catcga ggc caa ggc cccaaa cgc cgt g
 721 V V W N L A L G E E A V A N L Q F W L L N I E A K A P N A V

2251 gtgctggtggtcgggacgca cctgga tttaat tgaagc caa gttccg tgtgga aaggat tgc aac gctgcg tgc cta tgt gct ggc act c
 751 V L V V G T H L D L I E A K F R V E R I A T L R A Y V L A L

2341 tgcgctc ccc ctc cggctc cagggc cac aggtctccc aga catcac cttcaa aca cttaca tgagat ttcctg caagagc cctgga aggt
 781 C R S P S G S R A T G F P D I T F K H L H E I S C K S L E G

2431 caggaagggtgcgacagctgat tttcca cgtcacgtgcagcatgaa gga cgtcggcagcac catcggctgccaacgactggcaggcgg
 811 Q E G L R Q L I F H V T C S M K D V G S T I G C Q R L A G R

2521 ctgatccc caggagctacctgagcctgca gga ggc cgtgctggcaga gca gca ggc cgc cag cgggga cga cga cgtgca gta cctgac g
 841 L I P R S Y L S L O E A V L A E O O R R S R D D D V O Y L T

COR-Domäne

2611 gacagcagctgga gca gctggtgga gca gac gcc cga caa cga catcaa gga cta cga gga cctgagtc agccatcagctt cctcata
 871 D R Q L E Q L V E Q T P D N D I K D Y E D L Q S A I S F L I

2701 gaaac cgg cac cctgctcca tttccc gga cac cagcca cggcctgaggaa cctcta cttcctcga ccttat tggctctcga atgtctg
 901 E T G T L L H F P D T S H G L R N L Y F L D P I W L S E C L

2791 cagagatctttaa tattaagggtctc gatcagtgcc caagaa tggggatgatcagagc aga agacctcag gatgctgctggg gggac t
 931 Q R I F N I K G S R S V A K N G V I R A E D L R M L L V G T

2881 ggcttcacgcagcagcagga agagcagta cttccagttcctggc caagtttgagatcgc cctgcc cgtcgc caa tga cagcta cctcctg
 961 G F T Q Q T E E O Y F Q F L A K F E I A L P V A N D S Y L L

2971 cccatctccttccatctaa acctggcctgga caccca cggtatggggca ccccac agc caa cac cat tca gagggtat taa gatgagc
 991 P H L L P S K P G L D T H G M R H P T A N T I O R V F K M S

Abb. 20: DNA- und Aminosäuresequenz der humanen LRRK1 (Fortsetzung)

3061 ttcgt tcc cgt tgg ctt ctg gca aag gtt tat agc acg gat gct gat cag cct ggc gga gat gga cct gca gct ttt tgaaa caagaa
 1021 F V P V G F W Q R F I A R M L I S L A E M D L O L F E N K K

COR-Domäne

3151 aa tactaa agcaggaa caggaa agt cac cat tta cag ttt tac aggaaa cca gagaaa tcg ctg tag cac att cag agt gaa aagaa t
 1051 N T K S R N R K V T I Y S F T G N Q R N R C S T F R V K R N

3241 cagac cat cta ttggcagga agggct cct ggt cac ttt tga tggggg cta cct cag tgt gga atc ttc cga cgt gaa ctg gaa aaa gaa
 1081 Q T I Y W Q E G L L V T F D G G Y L S V E S S D V N W K K K

3331 aaaagcggaggaat gaa aat tgt ttgcca atcaga agt gag gga ctt ctc agc cat ggc ttt cat cac gga cca cgt caa ttc ctt gat t
 1111 K S G G M K I V C Q S E V R D F S A M A F I T D H V N S L I

3421 gatcagtggtt tcc cgc cct gac ggc cac aga gag cga cgg gac gcc act cat gga gca gta cgt gcc ctg ccc ggt ctg cga gac agc c
 1141 D O W F P A L T A T E S D G T P L M E Q Y V P C P V C E T A

3511 tgggc cca gca cac gga ccc cag tga gaa atcaga gga tgt gca gta ctt cga cat gga aga ctg tgt cct gac ggc cat cga gcg gga c
 1171 W A Q H T D P S E K S E D V Q Y F D M E D C V L T A I E R D

Linker2

3601 ttcatctctgccc agaca ccc gac cct ccc cgt gcc gct gca gga gct ggt cct ga act gtt cat gac cga ctt ccc ggc cag gct c
 1201 F I S C P R H P D L P V P L Q E L V P E L F M T D F P A R L

3691 tt cct gga gaa cag caa gct gga gca cag cga gga cga ggg cag cgt cct ggg cca ggg cgg cag tgg cac cgt cat cta ccg ggc cgg
 1231 F L E N S K L E H S E D E G S V L G O G G S G T V I Y R A R

3781 ta ccaggg cca gcc tgt ggc cgt caa gcy ctt cca cat caa aaa att caa gaa ctt tgc taa cgt acc ggc aga cac cat gct gag gca c
 1261 Y O G O P V A V K R F H I K K F K N F A N V P A D T M L R H

Kinasedomäne

3871 ctg cgg ggc cac cga tgc cat gaagaa ctt ctc cga gtt ccg gca gga ggc cag cat gct gca cgc gct gca gca ccc ctg cat cgt ggc
 1291 L R A T D A M K N F S E F R O E A S M L H A L O H P C I V A

3961 ct cat cgg cat cag cat cca ccc gct ctg ctt cgc cct gga gct cgc gcc gct cag cag cct caa cac cgt gct gtc cga gaa cgc caga
 1321 L I G I S I H P L C F A L E L A P L S S L N T V L S E N A R

4051 gattcttcctt tat acc cct ggg aca cat gct cac ccaaaa atc agc cta cca gat cgc ctc ggg cct ggc cta cct gca caa gaa aac
 1351 D S S F I P L G H M L T O K I A Y O I A S G L A Y L H K K N

4141 at cat ctt ctg tga cct gaagtc gga caa cat tct ggt gtg gtc cct tga cgt caaggagca cat caa cat caa gct atc tga cta cggg
 1381 I I F C D L K S D N I L V W S L D V K E H I N I K L S D Y G

4231 atttcgaggcagtc att cca tgagggcgc cct agg cgt gga ggg cact cc tgg cta cca ggc ccc agagat cag gcc ctg cat tgt atat
 1411 I S R Q S F H E G A L G V E G T P G Y Q A P E I R P R I V Y

4321 gatgagaaggat agat at gtt ctc cta tgg aat ggt gct cta cga gtt gct gtc aggaca gcgccc tgc act ggg cca cca cca gct ccag
 1441 D E K V D M F S Y G M V L Y E L L S G Q R P A L G H H Q L Q

4411 at tgc caa gaagct gtc caa ggg cat ccg ccc ggt tct ggg gca gcc gga gga agt gca gtt ccg gcg act gca ggc gct cat gat gga
 1471 I A K K L S K G I R P V L G Q P E E V Q F R R L Q A L M M E

4501 tgctggga cac taagccaga gaagc gacc gct ggc cct gtc ggt ggt gag cca gat gaa gga ccc gact ttt ggc cac ctt cat gta tga
 1501 C W D T K P E K R P L A L S V V S O M K D P T F A T F M Y E

Abb. 21A: Multiples Sequenzalignment der Roc- und COR-Domänen von ROCO-Proteinen verschiedener Organismen

Multiple sequence alignment of LRR#12, LRR#13, LRR#14, Linker 1, and Roc-Domäne domains from various organisms including Chl. tep.TLS, Meth.bark, Meth.sarc.C2A, Tri.ery, Nos.punc, Magn.sp, C.elegans, D.melanogast, H.sapiens, D.discoideum, and H.sapiens.

Roc-Domäne

Multiple sequence alignment of the Roc-Domäne domain from various organisms including Chl. tep.TLS, Meth.bark, Meth.sarc.C2A, Tri.ery, Nos.punc, Magn.sp, C.elegans, D.melanogast, H.sapiens, D.discoideum, and H.sapiens.

Roc-Domäne

COR-Domäne

Multiple sequence alignment of the Roc-Domäne and COR-Domäne domains from various organisms including Chl. tep.TLS, Meth.bark, Meth.sarc.C2A, Tri.ery, Nos.punc, Magn.sp, C.elegans, D.melanogast, H.sapiens, D.discoideum, and H.sapiens.

COR-Domäne

Multiple sequence alignment of the COR-Domäne domain from various organisms including Chl. tep.TLS, Meth.bark, Meth.sarc.C2A, Tri.ery, Nos.punc, Magn.sp, C.elegans, D.melanogast, H.sapiens, D.discoideum, and H.sapiens.

COR-Domäne

Multiple sequence alignment of the COR-Domäne domain from various organisms including Chl. tep.TLS, Meth.bark, Meth.sarc.C2A, Tri.ery, Nos.punc, Magn.sp, C.elegans, D.melanogast, H.sapiens, D.discoideum, and H.sapiens.

Abb. 21B: Multiples Sequenzalignment der Linker 2 Region von ROCO-Proteinen verschiedener Organismen

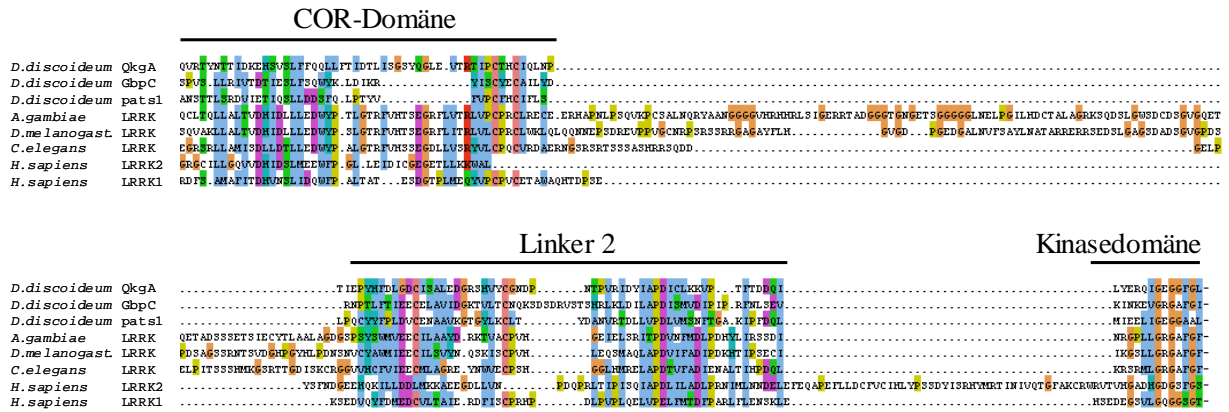


Abb. 21: Multiples Sequenzalignment der Roc- und COR-Domänen von ROCO-Proteinen verschiedener Organismen. Die Proteinnamen und Organismen sind links und die Domänenstruktur der ROCO-Proteine oberhalb des Sequenzalignments angegeben. Die Bezeichnung der Linker-1- und Linker-2-Region ist von Bosgraaf *et al.* [1] übernommen worden. **A** Vergleich der Sequenzen von ROCO-Proteinen des MASL1- und des LRRK-Typs über drei LRRs, Roc- und COR-Domäne. **B** ROCO-Proteine des LRRK-Typs enthalten einen zusätzlichen konservierten Bereich (Linker 2) N-terminal zur Kinasedomäne.

Alle konservierten Aminosäuren sind farblich unterlegt nach dem Jalview ClustalX-Schema [10]: hydrophob (hellblau), Tyrosin und Histidin (dunkelblau), positiv geladen (rot), negativ geladen (pink), Serin, Threonin und Asparagin (grün). Alle Glycine (braun) und Proline (gelb) sind farblich unterlegt.

Abb. 22A: Sequenzalignment der einzelnen WD-40-Repeats in den C-Termini von ROCO-Proteinen

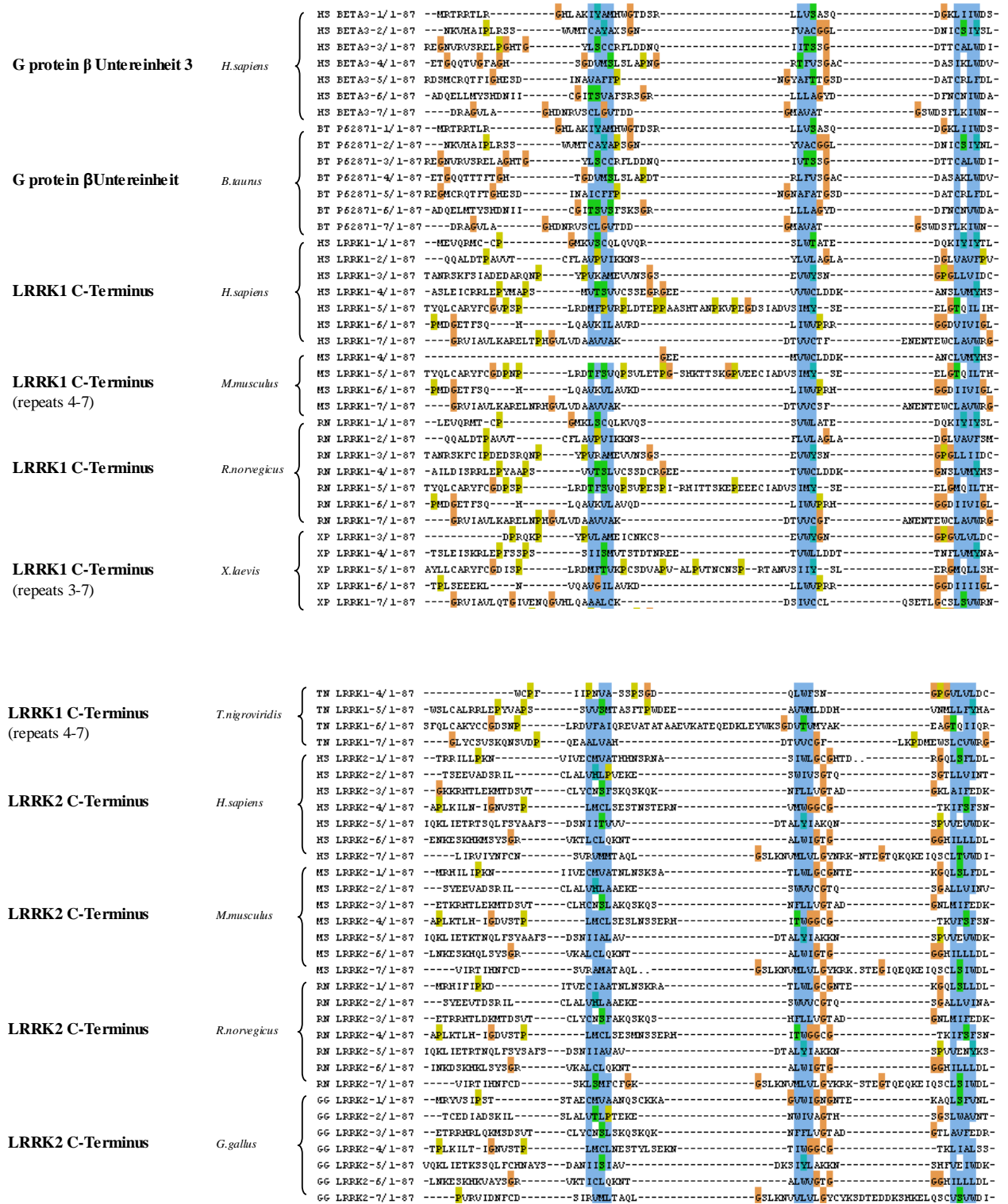


Abb. 22A: Sequenzalignment der einzelnen WD-40-Repeats in den C-Termini von ROCO-Proteinen (Fortsetzung)

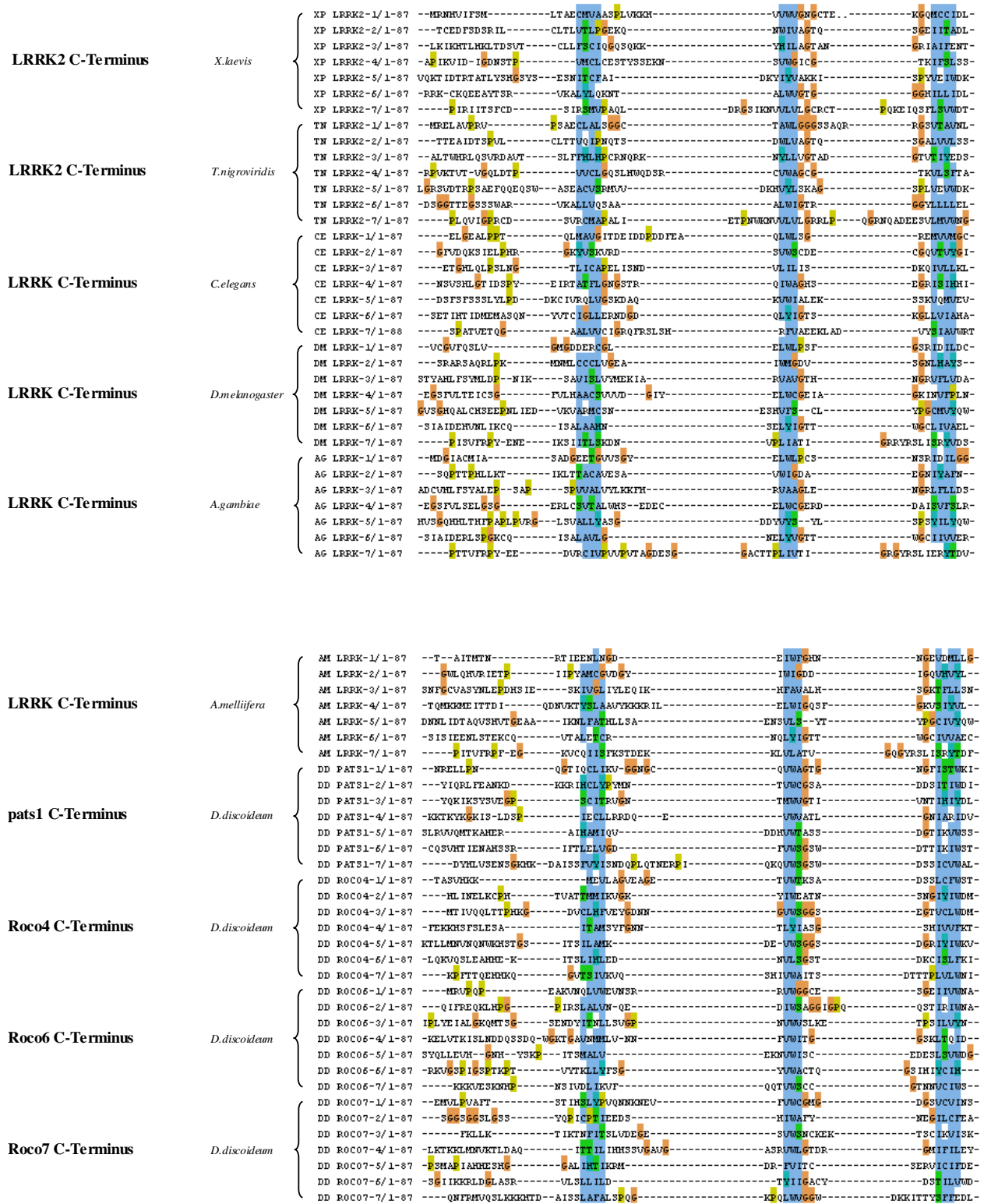


Abb. 22B: SekundärstrukturVorhersage der WD-40-Repeats in den C-Termini von ROCO-Proteinen

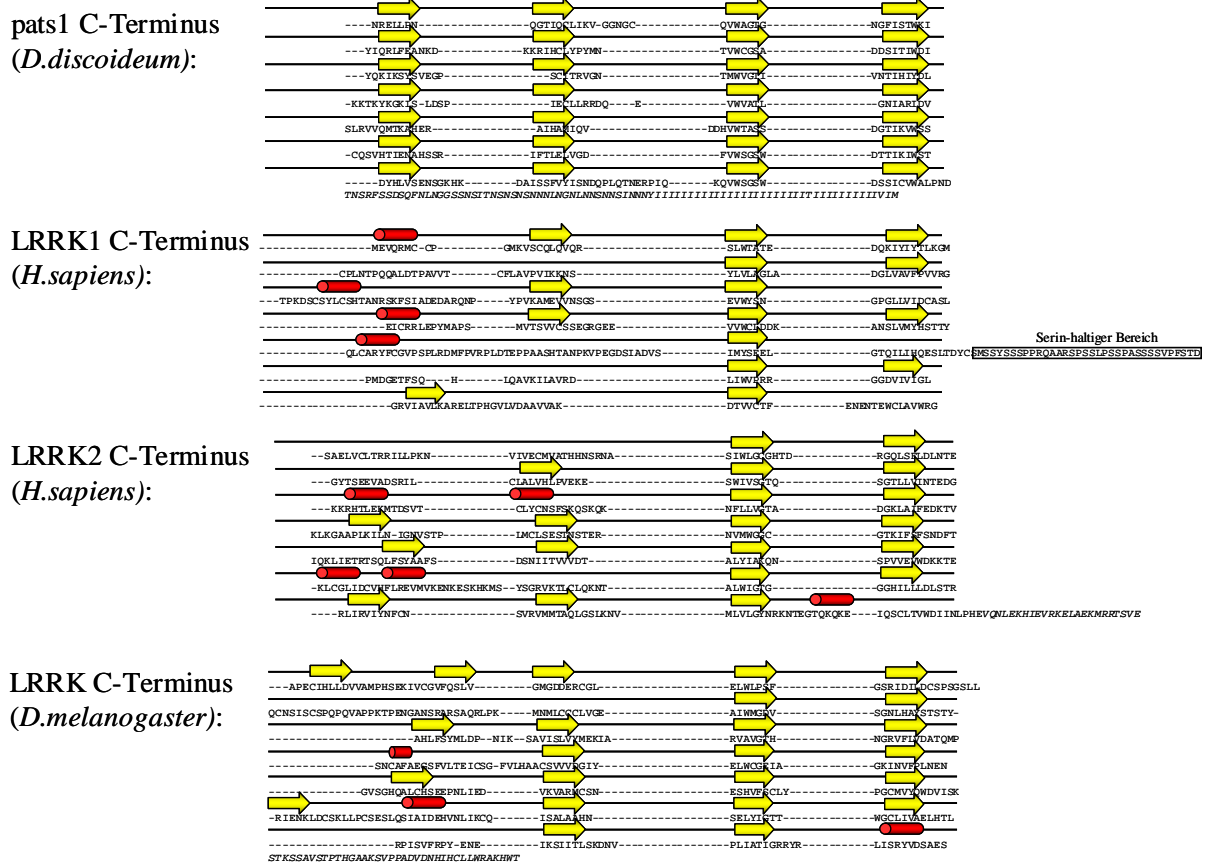


Abb. 22: WD-40-Repeats in den C-Termini von ROCO-Proteinen. **A** Sequenzalignment der einzelnen WD-40-Repeats in den C-Termini von ROCO-Proteinen. Dargestellt sind die β -Untereinheiten von G-Proteinen (HS beta 3: β -Untereinheit 3 aus *H. sapiens* und P62871: β -Untereinheit aus *B. taurus*) und die C-Termini der ROCO-Proteine. Alle Sequenzen sind in sieben Repeats unterteilt mit je drei konservierten hydrophoben Blöcken. Proteinbezeichnung, Organismus und Nummerierung der Repeats ist links angegeben. Die LRRK1 Sequenzen aus *M. musculus*, *X. laevis* und *T. nigroviridis* sind nur unvollständig in den Datenbanken vorhanden. **B** SekundärstrukturVorhersage der WD-40-Repeats in den C-Termini von ROCO-Proteinen. Beispielhaft sind die SekundärstrukturVorhersagen (berechnet mit PSIPRED [94]) von pats1 (*D. discoideum*), LRRK1, LRRK2 (beide *H. sapiens*) und LRRK (*D. melanogaster*) gezeigt. Die kalkulierten β -Faltblätter (gelbe Pfeile) und α -Helices (rote Zylinder) mit der größten Wahrscheinlichkeit (Neigungswert größer 5 in einer Skala von 0-10) sind über den Sequenzen aus den Sequenzalignments aus A dargestellt. Grauer Kasten: Serin-reicher Bereich **C** Gesamtdarstellung der WD-40-Repeats in den C-Termini von ROCO-Proteinen. Die Ergebnisse aus A und B sind in einem multiplen Sequenzalignment zusammengefasst. Mit einbezogen sind die kompletten Aminosäuresequenzen der β -Untereinheiten von G-Proteinen und den C-Termini der ROCO-Proteine. Theoretisch erwartete β -Faltblätter und WD-40-Repeats sind oberhalb der Sequenzen dargestellt. In der Sequenz des Roco6-Proteins wurde der Bereich einer Pleckstrin-homologen Domäne (PH-Domäne [144, 145], weisser Kasten) ausgespart. Die Farbkodierung der Sequenzalignments entspricht denen in Abbildung 21.

Abb. 22C: Gesamtdarstellung der WD-40-Repeats in den C-Termini von ROCO-Proteinen

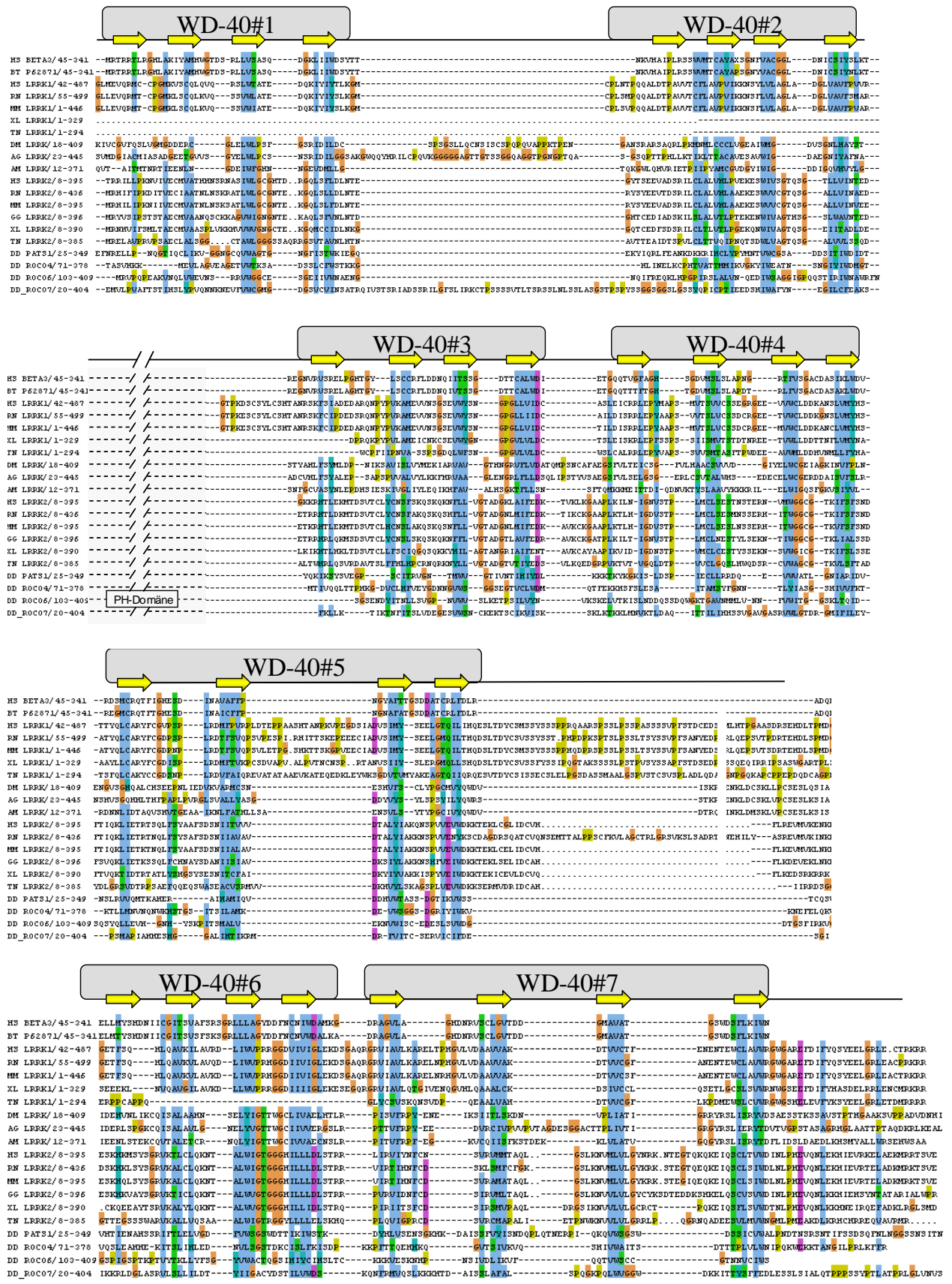


Tabelle 1: Auflistung der unterschiedlichen LRRK1-EST-Datenbankeinträge (*expressed sequence tags*)

	vorhanden	nicht vorhanden
71 bp Duplikation	Eileiter, Tumor; 8053058J1, Incyte Eileiter, Tumor; 7383958R8, Incyte Sequenz von Manning <i>et al.</i> Sequenz dieser Arbeit	Lymphknoten; CF137082, GenBank Lymphknoten; AW502884, GenBank Kopf/Hals; BE178723, GenBank
Exon 13	Sequenz von Manning <i>et al.</i>	Lymphknoten; BU431621, GenBank Darm; BG959015, GenBank Embr. Stammz.; CN365709 GenBank Sequenz dieser Arbeit
Verlängerung des Exon 16	Sequenz von Manning <i>et al.</i>	Hautmelanom; BU166882, GenBank Mandeln; AA282745, GenBank Mandeln; AA287802, GenBank Kopf; BI000807, GenBank Sequenz dieser Arbeit
	Exon 35	Exon 36
Alternative Stop-Kodons	Gebärmutterhals, Karz.; BC005408 Gebärmutterhals, Karz.; BC020379 (beide GenBank)	Darm; BF813830, GenBank Lymphknoten; BQ064084, GenBank Lymphknoten; CF142832, GenBank Lymphknoten; AW502376, GenBank Hoden; CD243942, GenBank Embryonales Karz.; BG388614, GenBank Fettgewebe; CB268975, GenBank Rhabdomyom, Karz.; BF303665, GenBank Sequenz von Manning <i>et al.</i> Sequenz dieser Arbeit

Tabelle 1: Auflistung der unterschiedlichen LRRK1-EST-Datenbankeinträge (*expressed sequence tags*). Die in Abschnitt 3.1.1 erwähnten Unterschiede in den EST-Datenbankeinträgen der LRRK1 sind aufgelistet. Die ESTs, die von Manning *et al.* [110] veröffentlichte und die dieser Arbeit zu Grunde liegende Sequenz sind nach Präsenz oder Fehlen einer 71 bp Insertion, des Exon 13 und einer Verlängerung des Exon 16 aufgelistet. Die ESTs der Exons 35 und 36 enthalten beide Stop-Kodons und kodieren für alternative C-Termini der LRRK1.