AD_____

Award Number: W81XWH-04-1-0509

TITLE: The Role of Drosophila Merlin in the Control of Mitosis Exit and Development

PRINCIPAL INVESTIGATOR: Long-Sheng Chang, Ph.D.

CONTRACTING ORGANIZATION: Children's Research Institute
Columbus, OH 43205-2696

REPORT DATE: July2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-07-2006	Annual	1 JUL 2005 - 30 JUN 2006

4. TITLE AND SUBTITLE
The Role of Drosophila Merlin in the Control of Mitosis Exit and Development

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-04-1-0509

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Long-Sheng Chang, Ph.D.

E-Mail: lchang@chi.osu.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Children's Research Institute
Columbus, OH 43205-2696

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT: Presently, the mechanism by which Merlin functions as a tumor suppressor is not understood. By utilizing *Drosophila* genetics, we have found a role of Merlin in the control of mitosis exit. Merlin mutations lead to two types of mitosis exit asynchrony, the asynchronous anaphase-telophase figures and the asynchronous telophase-interphase figures. Also, we show that cells lacking Merlin possess greater ability to overcome vein restriction in the wing. The Merlin protein is colocalized with the Wingless morphogen in the cells at the dorsal/ventral compartment border of the wing imaginal disc. Merlin inactivation may lead to an alteration on the determination/maintenance of Wg stripe expression. We have found potential genetic interactions between the *Merlin* and *porcupine* genes and between the *Merlin* and *shibire* genes. We also discover an interaction between the *Merlin* and *lap* (*like-AP180*), which is important for clathrin-mediated endocytosis of synaptic vesicles, was identified. Our results suggest that Merlin counteracts with Lap and through Lap, Merlin may regulate the EGFR pathway required for vein fate determination in the wing. In addition, by analyzing the evolution, diversity, and overall distribution of merlin among different taxa, we demonstrate a monophyletic origin of the merlin proteins with their root in the early metazoa. The overall similarity among the primary and secondary structures of all merlin proteins and the conservation of several functionally important residues suggest a universal role for merlin in a wide range of metazoa.

15. SUBJECT TERMS
Neurofibromatosis 2 (NF2), *NF2* Gene, Merlin, ezrin-radixin-moesin (ERM), *Drosophila melanogaster*, mitosis exit, development, imaginal disc, morphogen, vesicular trafficking, wingless, porcupine, shibire, Lap, spermatogenesis, and evolution

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	42	**19b. TELEPHONE NUMBER** *(include area code)*

TABLE OF CONTENTS

INTRODUCTION:

Neurofibromatosis type 2 (NF2) is a hereditary disorder characterized by the development of bilateral vestibular schwannomas and is associated with mutations in the tumor suppressor gene called the *neurofibromatosis type 2* (*NF2*) gene (Chang et al., 2005; Neff et al., 2006). The *NF2* gene encodes a protein named Merlin for moesin-ezrin-radixin like protein (Trofatter et al., 1993) Merlin shares a great deal of homology with the ezrin, radixin, and moesin (ERM) proteins, which belong to the protein 4.1 superfamily of cytoskeleton-associated proteins that link cell surface glycoproteins to the actin cytoskeleton. Presently, the mechanism by which Merlin functions as a tumor suppressor is poorly understood.

Drosophila melanogaster provides a genetic and developmental system, which is amenable to experimental manipulation, and has been very valuable to the study of tumor genetics. The *Drosophila* homolog of Merlin shares sequence similarity to the human Merlin protein (McCartney and Fehon, 1996; Fehon et al., 1997). In addition, the human *NF2* gene could rescue the lethal Merlin mutant allele in *Drosophila*, implying a functional conservation (LaJeunesse et al., 1998). Molecular genetic analysis reveals that Merlin is essential for regulation of proliferation and differentiation in the imaginal disc. However, understanding the tumor-suppressor function of Merlin requires additional knowledge about specific cell-cycle points where Merlin regulates proliferation and coordinates it with morphogenesis.

We have found that cells in the wing imaginal disc from the fly larva with Merlin mutation (Mer^4) displayed abnormalities in the control of mitosis exit. Cytological images of mutant cells frequently showed asynchronous anaphase and telophase. We have also isolated adult Mer^4 pharates. Interestingly, these *Mer* mutant adults showed abnormal leg morphology. Some of them displayed duplication of the wing disc, and in some cases, abnormalities were seen in the dorsal/ventral compartment border of the *Mer* mutant wing disc. These results suggest that Merlin is important not only for the control of mitosis exit but also for the determination/maintenance of morphogen gradients in the wing imaginal disc.

The goal of our proposed research is to examine the novel role of Merlin in the control of mitosis and development. Specifically, we plan to confirm the role of Merlin in the control of mitosis and determine whether there are any additional points in the cell cycle where Merlin executes its activity. We will examine the role of Merlin in wing imaginal disc development and the effect of Merlin mutation on specific regulatory protein expression within the wing imaginal disc. In addition, we will attempt to investigate whether the abnormalities in mitosis observed in Merlin mutant fly can also be seen in mouse and human schwannoma cells lacking *NF2* function. From this study, we hope to a better understanding of how Merlin executes regulation of proliferation and how it coordinates proliferation, mitosis, and morphogenesis. Future investigation of the signaling pathways that link Merlin to intracellular signals regulating cell division may enable designs for novel therapeutic regiments to cure NF2 schwannomas and other associated tumors.

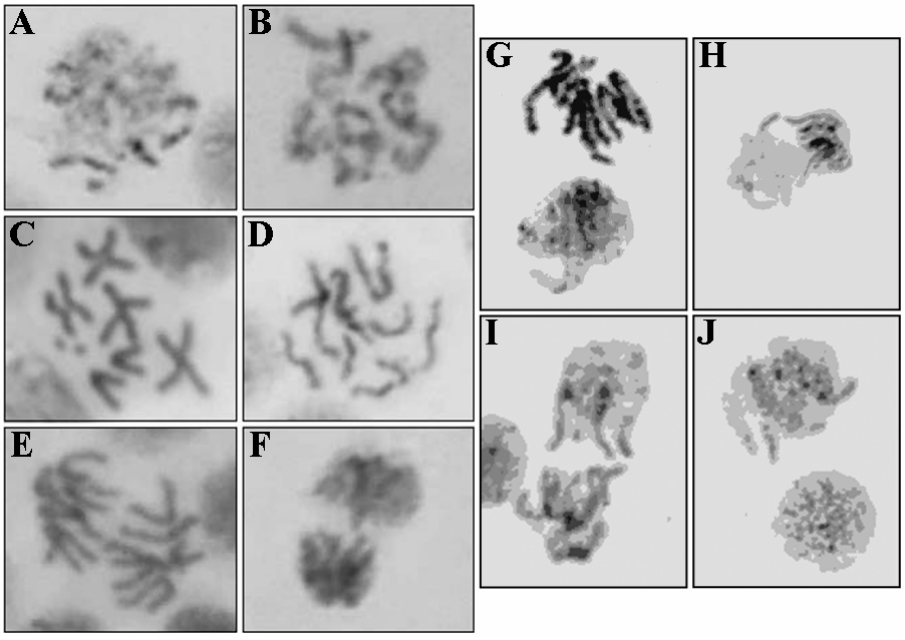
BODY:

Aim 1: <u>To conduct cytological analysis on additional Merlin mutant alleles and allelic combinations for the control of mitosis exit and morphogenesis</u>

Task 1: Accomplished in Year One.

Task 2: We have completed the cytological analysis on tissues isolated from three wild-type strains, various *Merlin* mutants [Mer^{1} ($Gln^{324\rightarrow stop}$), Mer^{2} ($Gln^{318\rightarrow stop}$), Mer^{3} ($Met^{177\rightarrow Ile}$), and Mer^{4} ($Gln^{170\rightarrow stop}$)], as well as the $Mer^{4};Mer^{+}$ strain that contains a wild-type *Merlin* transgene in the *Mer4* background. We observed that *Merlin mutants displayed* two types of mitosis exit asynchrony more frequently than the three wild-types strains. The first is the asynchronous anaphase-telophase figures with one sister chromosome set in anaphase and the other in telophase (Figure 1A-B) and the second is the asynchronous telophase-interphase figures with one sister chromosome set in telophase and the other in interphase (Figure 1C-D).

*Figure 1. Images of normal mitosis in neural brain cells from the third-instar wild-type larvae (A-F) as compared with those of asynchronous mitosis exit in neural brain cells from the Mer^{4} mutant larvae. During normal mitosis, chromosome condensation at heterochromatin regions begins at early prophase P1 (A). Condensed homologous chromosomes are paired at late prophase P2 (B). Sister chromatids are connected at precentromeric heterochromatin regions in metaphase (C). Chromosomes are separated along their lengths at late metaphase (D). Two sister chromosome sets are separating from each other and oriented toward the opposite poles in anaphase (E). Fully separated sister chromosome sets begin to undergo chromosome decondensation in telophase (F). Unlike normal mitosis, the Mer^{4} mutant cells frequently display asynchronous anaphase-telophase figures. While tw*o sister chromosome sets are separating from each other and oriented toward the opposite poles, one set of chromosomes already undergoes chromosome decondensation (G,H). *The Mer^{4} mutant cells also frequently show asynchronous telophase-interphase figures (I,J).*



As shown in Table 1, in contrast to those from the three wild-type strains, neural brain cells from all four *Merlin* mutants showed both types of mitosis exit asynchrony at higher frequencies. The frequency of asynchronous figures was lower in the weak Mer^3 (missense mutation) allele than the other three truncated *Merlin* alleles, Mer^1, Mer^2, and Mer^4. Importantly, when a wild-type *Merlin* transgene was introduced back to the Mer^4 genetic background, the asynchronous mitosis exit phenotype was substantially diminished. These results suggest that Merlin is important for the control of mitosis exit.

Table 1. Mitotic exit asychrony in various Merlin alleles, compared with the wild-type Lausenne, Hikkone A-W, and Oregon R strains.

Strain	No. of cells in anaphase analyzed[a]	No. of cells with asynchronous anaphase-telophase figures[b]	% of cells in anaphase with asynchronous anaphase-telophase figures	No. of cells in telophase analyzed[c]	No. of cells with asynchronous telophase-interphase figures[d]	% of cells in telophase with asynchronous telophase-interphase figures
Lausenne	298	0	0	11	0	0
Hikkone A-W	93	0	0	14	0	0
Oregon R	101	0	0	26	0	0
Mer^4	118	14	11.9	40	16	40
$Mer^4;Mer^+$	257	7	2.7	18	1	5.6
Mer^3	161	16	9.9	29	5	17.2
Mer^2	80	14	17.5	10	3	30
Mer^1	38	5	13.2	7	2	28.6

[a]Cells with at least one sister chromosome set in anaphase
[b]One sister chromosome set in anaphase while the other in telophase.
[c]Cells with at least one sister chromosome set in telophase but none in anaphase
[d]One sister chromosome set in telophase while the other in interphase.

Task 3: The task was proposed for the 3rd year study.

Aim 2: <u>To examine and compare the duration of the cell cycle and mitosis phases using various Merlin mutants and to study subcellular localization of Merlin at various phases of mitosis</u>

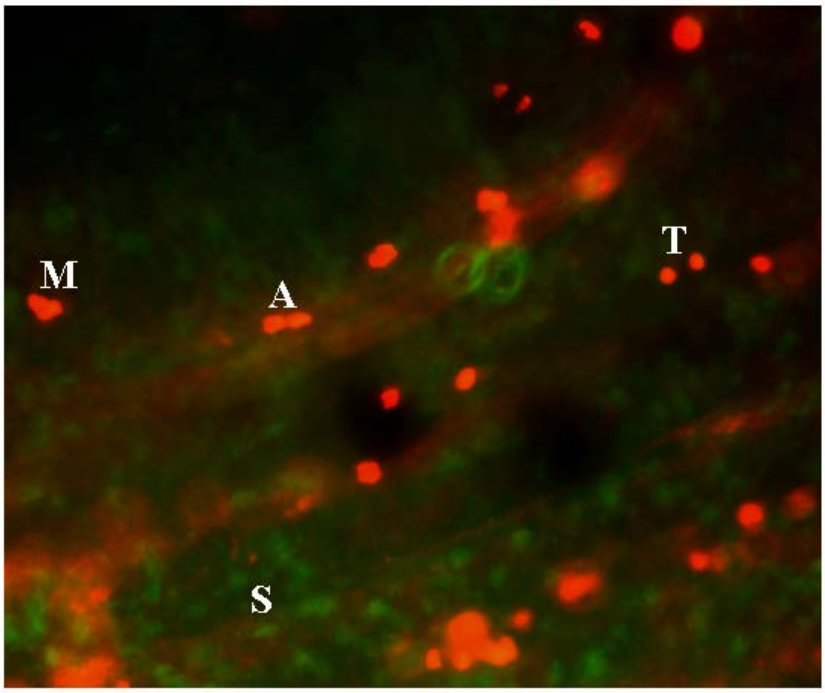
Task 4: This task was completed in year one. Our results indicate that the G_2 phase of Mer^4 cells is about two times shorter than that of wild-type cells.

Task 5: We have made an attempt to cultivate the Mer^4 and wild-type wing imaginal discs using the Robb's complete tissue culture medium (Ashburner, 1989). We found that while we were able to cultivate wing imaginal discs in this medium for about 2-3 hr, cultivation for a longer period led to cell detachment from the discs. Subsequently, we used the Schneider's tissue culture medium (Sigma Chemicals Co.), and we checked the metabolic activity of the discs by BrdU incorporation, followed by anti-BrdU antibody staining. With this approach, we were able to cultivate the discs for 10-13 hr.

We have performed a double labeling experiment with BrdU and anti-phospho-histone H3 antibody on the cultured wing imaginal disc. BrdU incorporation allowed the detection of S-phase cells, while anti-phospho-histone H3 antibody labeled cells in mitosis. Figure 2

displays an area of the double labeled wing imaginal disc. The green label illustrates the BrdU incorporated cells, which represented the cells at the S phase. The red label represents the cells at various stages of mitosis as determined by chromatin morphology. We are in the process of culturing BrdU-labeled cells for various times and then staining them with anti-phospho-histone H3 antibody. By counting the number of doubly labeled cells from different time points, we hope to determine the durations of various cell cycle phases.

Figure 2. Double labeling of the wing imaginal disc with BrdU and anti-phospho-histone H3 antibody allows examination of cells at various stages of the cell cycle. Examples of cells at the S phase (S), metaphase (M), anaphase (A), and telophase (T) were indicated.



Task 6: As indicated in the year-One report, we found that the overall cell-cycle duration in the wing disc was not significantly affected by *Merlin* mutation. To study the effect of *Merlin* mutation on cell proliferation in different regions of the wing imaginal disc, mosaic clones of *multiple wing hairs* (*mwh*) were induced in the *+/mwh* larvae with or without Mer^3 mutation at 96 hr after egg laying (AEL) by 1000R of γ-rays irradiation. Irradiated larvae were grown to the adult stage and male wings were removed. The *mwh/mwh* clone spots in the wings were photographed. Clone spots from 70 Mer^3 and 76 wild-type wings were projected onto the map of an adult wing. Figure 3 illustrates the distribution of mosaic mwh clones in the wild-type and Mer^3 adult wings. Consistent with the previous finding (Gonzalez-Gaitan et al., 1994), the wild-type mosaic clones induced at such a late developmental stage (96 hr AEL) respected the vein restriction and resided within the boundary of veins. However, the Mer^3 mutant clones did not follow such a rule and some of them crossed the vein boundary (arrows in Figure 3 point to these clones). In addition to crossing the vein, the mosaic clones that abut on the vein were also more frequently found in the Mer^3 mutant wing than those in the wild-

type control (Table 1). These results further support the role of Merlin in cell motility, cell adhesion and cell proliferation. Cell lacking Merlin may possess greater ability to overcome vein restriction.

Figure 3. Mer^3 mutation delimits the normal clonal restriction of mosaic clones to cross veins. Arrows point to the mer^3 mutant clones that cross the vein boundary. Colors distinguish clone overlaps with the brown color for two overlapping clones, violet for three overlapping clones, and blue for four overlapping clones. The red color denotes uni-cell clones and the green color indicates two-cell clones.

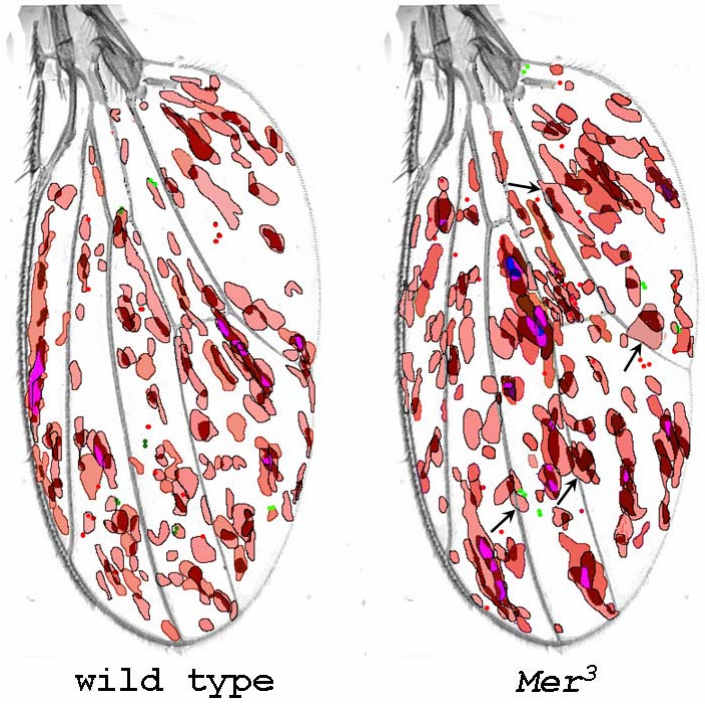
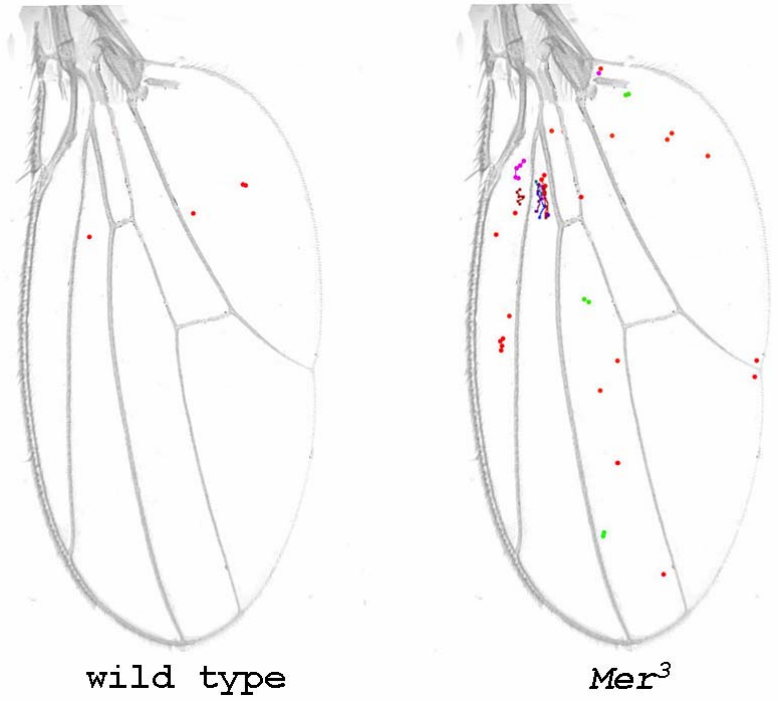


Table 1. Vein boundary restriction of mosaic clones in the wild-type and Mer^3 genetic background. Mosaic clones induced in each wing were carefully examined on an enlarged images. Data were collected from the analysis of 70 Mer^3 and 76 wild-type wings.

Genetic background of mosaic clones	No. of clones abutting on the wing vein	No. of clones crossing the vein	Total No. of clones analyzed
Mer^3	16	6	240
Wild type	6	0	193

We also induced mosaic clones in the wild-type and *Mer^3* wings late in development (48h after puparium formation). Consistent with previous observation showing that no cell proliferation occurred in the wing disc at this time (Garcia-Bellido and Merriam, 1971), only three uni-cell clones were found in 62 wild-type wings examined (Figure 4). In contrast, we identified 26 mosaic clones in 47 *Mer^3* wings analyzed; among them, three were two-cell clones and five others contained multiple cells. These results indicated that significant cell proliferation took place 48 hr after puparium formation in the *Mer^3* wings. Our results further argue the possibility for a temporal difference in developmental timing between the wild-type and *Mer^3* mutant larvae.

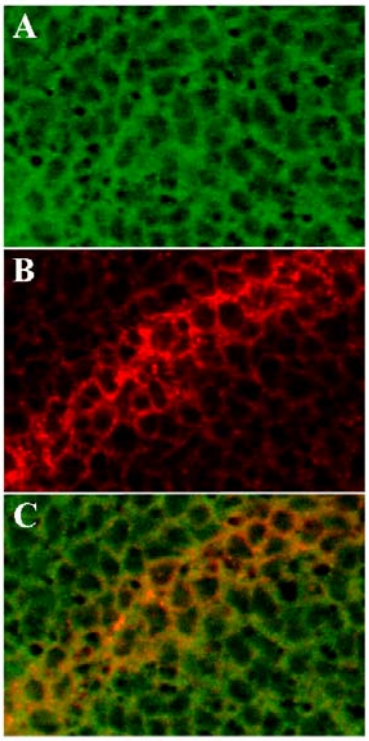
Figure 4. Mosaic clone induction in the wild-type and Mer^3 wing 48h after puparium formation. Each cell is marked by a color dot. Two-cell clones are indicated in green. The cells in multi-cell clones are connected by a line. The picture represents the summary of mosaic clones found on 62 wild-type and 47 Mer^3 wings.



Task 7: We acquired a small aliquot of anti-Merlin antibody from Dr. Rick Fehon at the University of Chicago and performed an immunostaining experiment. As reported previously (McCartney and Fehon, 1996), the Merlin protein displayed cortical localization and some cytoplasmic staining throughout the wing imaginal disc tissue (Figure 5A). Johnston and Edgar (1998) previously showed that the Wingless (Wg)-expressing cells at the dorsal/ventral (D/V) compartment border of the wing imaginal disc were arrested at the G_1 phase of the cell cycle. Furthermore, in the anterior compartment, the stripe of G1-arrested, Wg-expressing cells at the D/V border was surrounded by the two stripes of G2-arrested cells. Consistently, we detected strong Wg staining in the cells at the D/V compartment border of the wing imaginal disc, and the majority of Wg protein staining was found in the plasma membrane (Figure 5B). By superimposing the Merlin and Wg staining images together, we found that the Merlin and Wg proteins appeared to be colocalized in the cells at the D/V compartment border (Figure 5C). Together with previous findings, these results indicate that the majority of Merlin protein is localized in the plasma membrane of cells in both the G1 and G2 phases. The colocalization between Merlin and the Wg morphogen further suggest a role of Merlin in Wg trafficking.

Figure 5. Colocalization of Merlin with the Wg morphogen at the D/V compartment borde. Wing imaginal discs from the 3rd instar wild-type larva were fixed and stained for Merlin (green) and Wg (red) using the guinea pig anti-Merlin polyclonal antibody (McCartney BM and Fehon RG, 1996) and the anti-Wg monoclonal antibody 4D4 (Brook and Cohen, 1994), respectively. (A) The Merlin protein was localized

to the plasma membrane and some within the cytoplasm. (B) The Wg protein was synthesized in the cells at the D/V compartment border. Within this stripe, Wg was associated with the plasma membrane. The Wg protein spread from the stripe of expressing cells to the adjacent celles in a gradient fashion by vesicle trafficking (punctuate structures). (C) Merlin and Wg were colocalized at the plasma membrane of cells at the D/V compartment border.



Aim 3: To further examine the role of Merlin in the determination/maintenance of the D/V compartment border in the *Drosophila* wing imaginal disc and to investigate how Merlin mutation affects the expression of proteins important for the determination of the compartment border

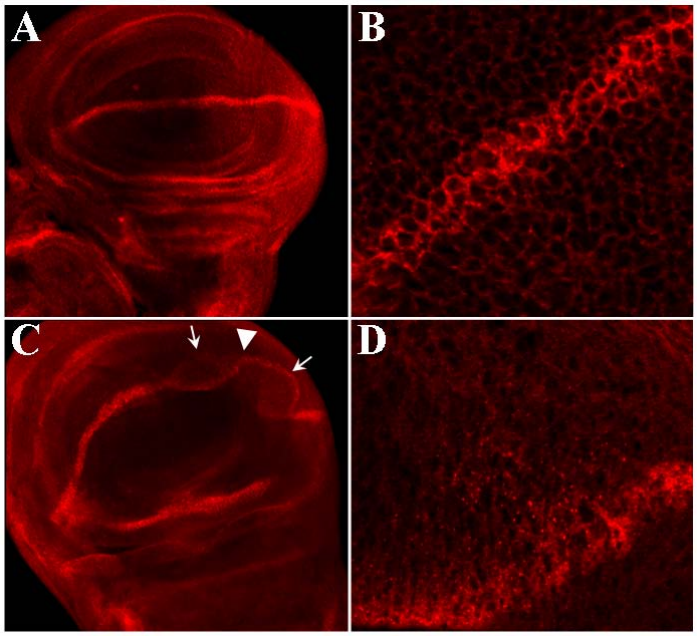
Task 8: The task was accomplished in year one and the results confirmed that Merlin plays an important role in the determination of the wing morphology.

Task 9: To examine the role of Merlin in the determination of the D/V compartment border in the wing imaginal disc, we examined the distribution pattern of the Wg morphogen in the wild-type and Mer^4 mutant wing imaginal discs. It is well-documented that the Wg protein is synthesized by the cells at the D/V compartment boundary and then forms spatial concentration gradients while moving away from the D/V border (Seto and Bellen, 2004). In contrast to the tight stripe pattern of Wg expression in the wild-type wing disc (Figure 6A-B), the Wg stripe in the mer^4 mutant disc was frequently more diffuse, and in about 40% of the cases some diffuse regions of the Wg stripe were of particular notable (indicated with arrows in Figure 6C-D). These results suggest that Merlin inactivation may lead to an alteration on the determination/maintenance of Wg stripe expression.

Fehon and colleagues previously found that Merlin and Expanded, another member of the protein 4.1 superfamily, were associated with

the apical junctional region in imaginal epithelia and with endocytic vesicles in cultured cells (McCartney and Fehon, 1996; McCartney et al., 2000). Together, they regulated the abundance, localization, and turnover of cell-surface receptors. Loss of these proteins affected abundance, cell-surface localization, and endocytic trafficking of Notch, EGFR, and other signaling and adhesion receptors in epithelial cells. As mentioned above, we found that the Merlin protein co-localized with the Wg protein in the plasma membrane of cells at the D/V compartment border of the wing imaginal disc (Figure 5). We also observed that the *Mer*^{4} mutant displayed a slightly more diffuse Wg stripe at the D/V compartment border. In light of the fact that the Wg morphogen is distributed through the tissues by vesicular trafficking (Entchev and Gonzalez-Gaitan, 2002), we hypothesize that *Merlin* mutation may affect this process.

Figure 6. Wg protein distribution in the wild-type and Mer4 mutant wing imaginal discs. The wing imaginal discs were stained with an anti-Wg antibody. The Wg stripe (red) was readily detected at the D/V compartment border of the wild-type wing pouch (A). At a higher magnification, the Wg protein showed a cortical localization in the cell (B). However, the Wg stripe at the D/V border of the mer^{4} *mutant disc appeared more diffuse (C), particularly in some regions of the Wg stripe (arrows). An enlarged area of a diffused region (arrowhead in C) showed that while the cortical localization of the Wg protein was still seen, Wg granules were diffusely distributed (D).*

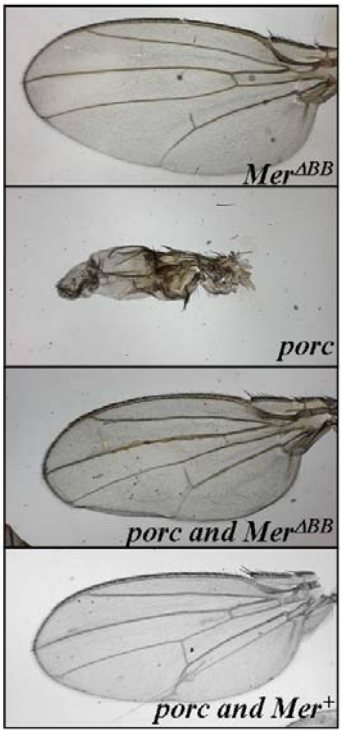


To test this hypothesis, we performed experiments to examine potential genetic interactions between Merlin and the proteins involved in vesicular trafficking. We first searched FlyBase for available strains carrying a UAS construct with a gene participating in vesicular trafficking, or having an EP-element insertion near or in the genes of interest. We identified fly strains carrying the UAS construct for *sgl*, *frc*, *Rab5*, *Rab7*^{DN}, *Csp*, or *shi*^{K44A}, as well as strains with an EP-insertion for *Damp*, *garnet*, *α-Adaptin*, *Scamp*, *Cirl*, *Gdi*, *Rop*, *lap*, or *AP-47*. We used the wing-pouch specific *Gal4* driver 1096 to ectopically express these constructs and examined any change in the wing morphology. Among the strains examined, only

ectopic expression of *porc*, *shi*^{K44A}, a dominant negative *shi* allele (*shi*^{DN}), and *lap* gave rise to abnormal wing morphology.

The *porcupine* (*porc*) is a segment polarity gene and encodes a putative multi-pass transmembrane protein belonging to the membrane-bound *O*-acyltransferase superfamily (Hofmann, 2000). Genetic and immunocytochemical studies suggest that *porc* is required for the secretion of active Wg ligand. The Porcupine (Porc) protein stimulates the posttranslational N-glycosylation of Wg in the endoplasmic reticulum. Ectopic expression of Porc stimulates the N-glycosylation of both endogenously and exogenously expressed Wg (Tanaka et al., 2002). The *shibire* (*shi*) encodes a dynamin protein, a GTPase essential for cytokinesis and endocytosis (van der Bliek and Meyerowitz, 1991; Kitamoto, 2002). The Shibire (Shi) protein is known to be involved in vesicle trafficking of various neurotransmitters and the Wg protein during early embryogenesis (Strigini and Cohen, 2000). The *lap* (*like-AP180*) gene encodes a presynaptically enriched clathrin adaptor protein. The Lap protein plays an important role in clathrin-mediated endocytosis of synaptic vesicles (SVs) and regulates the size of SVs (Zhang et al. 1998).

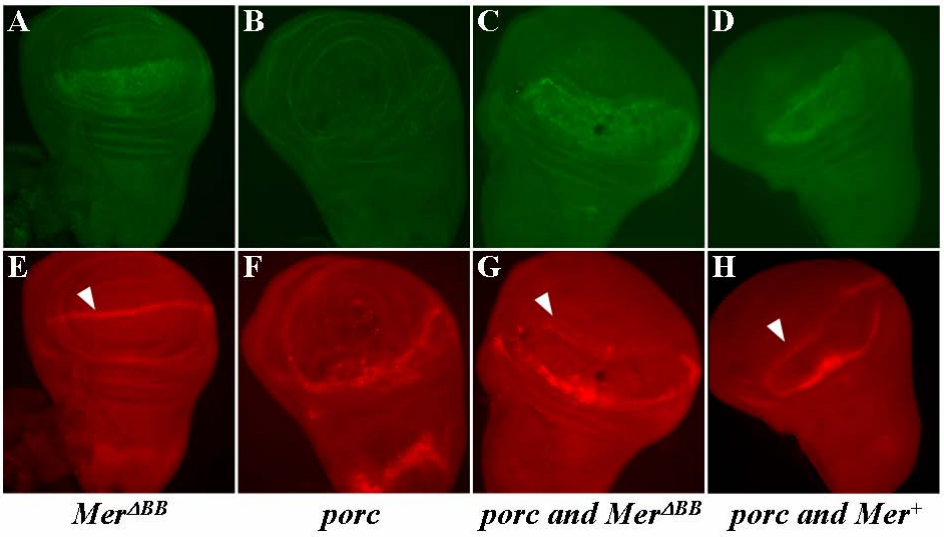
Figure 7. Potential genetic interaction between Mer and porc. (A) Overexpression of Mer$^{\Delta BB}$ *in the wing pouch using the 1096 driver did not affect the overall morphology of the wing. (B) Overexpression of porc in the wing pouch completely disrupted the wing morphology. (C) Overexpression of porc together with Mer*$^{\Delta BB}$ *resulted in some flies with normal wing morphology. (D) Over-expression of porc together with Mer*$^{+}$ *completely rescued the overall wing morphology.*



Last year, we briefly reported the potential genetic interaction between *Mer* and *porc*, and between *Mer* and *shi*^{DN} (*shi*^{K44A}). We have extended our analysis and showed that over-expression of *Mer*$^{\Delta BB}$, a dominant-negative *Merlin* mutant with a deletion of the Blue-Box in the FERM domain of Merlin (LaJeunesse et al., 1998), in the wing

pouch using the 1096 driver resulted in loss of the posterior cross vein and a slightly enlarged wing, but did not affect the overall wing structure (Figure 7A). In contrast, over-expression of *porc* in the wing pouch completely disrupted the wing structure, yielding abnormally small wing with no veins (Figure 7B). Interestingly, when *porc* was over-expressed together with $Mer^{\Delta BB}$ using the same 1096 driver, flies with normal wing morphology were recovered; however, the defect in the posterior cross vein was still present, presumably due to the effect of $Mer^{\Delta BB}$ over-expression (Figure 7C). Similarly, when *porc* was over-expressed together with a wild-type *Mer* (Mer^{+}), flies with completely normal wing morphology including cross veins were obtained (Figure 7D).

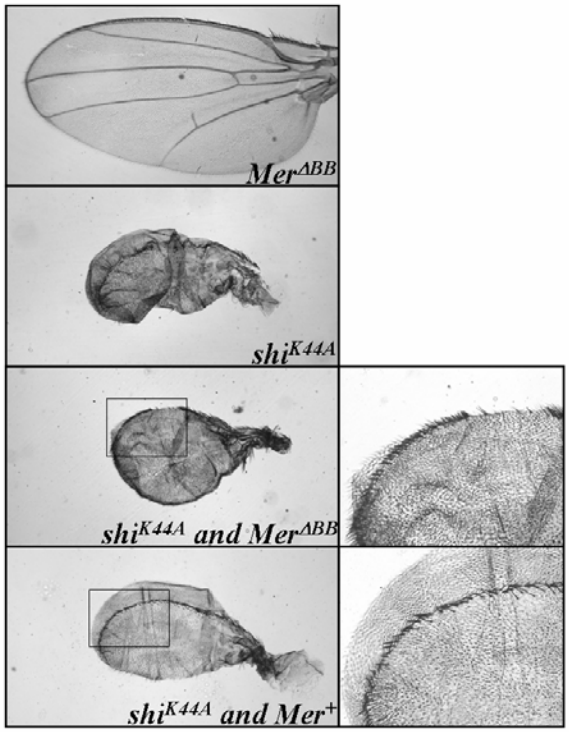
Figure 8. The effect of over-expression of $Mer^{\Delta BB}$ (A and E), porc (B and F), porc plus $Mer^{\Delta BB}$ (C and G), and porc plus Mer^{+} (D and H) on Wg expression at the D/V compartment border. (A-D) Immunostaning for Merlin. (E-H) Immunostaining for Wg. White arrowhead show the Wg stripe of expression at D/V compartment border.



As mentioned above, the Wg morphogen normally is synthesized in the cells at the D/V compartment border of the wing imaginal disc. Wg is secreted by these cells and spreads over the entire wing disc tissue in a gradient fashion. Wg is a long-range morphogen and its gradient activates target genes as a function of distance from the source. Wg is essential for embryonic development; however, strong viable alleles of *wg* have been isolated and could result in flies with no wings. We examined the effect of overexpression of $Mer^{\Delta BB}$, *porc*, $Mer^{\Delta BB}$ plus *porc*, and *porc* plus Mer^{+} in the wing pouch on Wg expression. Immunostaining with an anti-Merlin antibody revealed that the 1096 driver, as predicted, was active in the wing pouch, resulting in overexpression of the $Mer^{\Delta BB}$ and the wild-type Merlin protein (Figure 8A-D). When $Mer^{\Delta BB}$ was overexpressed in the wing pouch, no obvious effect on the Wg stripe at the D/V compartment border was seen (Figure 8E). Intriguingly, overexpression of *Porc* resulted in a complete loss of the Wg stripe (Figure 8F). Simultaneous overexpression of $Mer^{\Delta BB}$ or *Mer+* (the wild-type Merlin) with Porc led to re-appearance of the Wg stripe at the D/V

compartment border; however, a broad Wg stripe was observed particularly in the case of simultaneous over-expression of *Porc* and $Mer^{\Delta BB}$ (Figures 8G and 8H). Together, these results suggest a potential interaction between the *porc* and *Mer* genes. Such a genetic interaction has a profound effect on the determination of the Wg stripe at the D/V compartment border.

Figure 9. Potential genetic interaction between Mer and shi. Over-expression of $Mer^{\Delta BB}$ in the wing pouch using the 1096 driver did not affect the overall morphology of the wing. Over-expression of shi^{K44A} in the wing pouch completely disrupted the wing morphology. Over-expression of shi^{K44A} together with $Mer^{\Delta BB}$ partially restored the formation of staut bristles in the wing margin. Simultaneous over-expression of shi^{K44A} and Mer^{+} also partially restored the formation of triple rows including the medial triple row of staut bristles and ventral and dorsal triple rows of sensory bristles.



We have also extended our analysis on the potential genetic interaction between *Mer* and shi^{K44A}, a dominant negative *shi* allele. Overexpression of shi^{K44A} in a wing pouch using the 1096 driver resulted in flies with abnormally small wings, which did not have any veins or staut bristles, but had some disorganized sensory bristles (Figure 9). Staut bristles originate from the Wg-expressing cells at the D/V compartment border of the wing imaginal disc. The disappearance of staut bristles may imply that these cells did not undergo normal differentiation processes during wing development. Interestingly, when $Mer^{\Delta BB}$ was over-expressed together with shi^{K44A} in the wing disc using the 1096 driver, a partial restoration of triple rows including the medial triple row of staut bristles was observed; however, the wing remained small in size and had no vein (Figure 9). Similarly, simultaneous over-expression of Mer^{+} and shi^{K44A} in the wing pouch partially restored the formation of triple rows including the medial triple row of staut bristles and ventral and dorsal triple

rows of sensory bristles. However, the spacing and organization of ventral and dorsal triple rows still appeared abnormal, and there were no veins in the wing. These results suggest a potential genetic interaction between *Merlin* and *shibire*.

Furthermore, we found that over-expression of *lap*, which encodes a clathrin adaptor protein, in the wing pouch using the 1096 driver resulted in wings with ectopic vein formation, frequently seen at the distal part of vein V (Figure 10). When *lap* was over-expressed together with $Mer^{\Delta BB}$, excessive ectopic vein materials, which was even more than that with *lap* over-expression alone, appeared in many parts of the wing blade. In contrast, simultaneous over-expression of Mer^{+} and *lap* in the wing pouch resulted in wings with normal or almost normal vein formation (Figure 10). These results indicate that the dominant-negative Merlin $Mer^{\Delta BB}$ enhances the effect of Lap over-expression on the vein fate determination, while the wild-type Merlin suppresses it. LeJeunesse et al. (2001) have proposed that Merlin functions to antagonize the epidermal growth factor receptor (EGFR) pathway which control vein fate determination in the wing. Maitra et al. (2006) further show that Merlin and Expanded function cooperatively to modulate receptor endocytosis and signaling including EGFR. Taken together, our data suggest that Merlin counteracts with Lap, which is important for receptor endocytosis. Merlin may interact with Lap directly or indirectly. Through Lap, Merlin regulates the EGFR pathway required for vein fate determination in the wing (Figure 11).

Figure 10. Potential genetic interaction between Mer and lap. Over-expression of $Mer^{\Delta BB}$ in the wing pouch using the 1096 driver did not affect the overall morphology of the wing. Over-expression of lap in the wing pouch completely disrupted the wing morphology. Over-expression of lap together with $Mer^{\Delta BB}$ resulted in some flies with normal wing morphology. Over-expression of lap together with Mer^{+} completely rescued the overall wing morphology. Black arrowheads denote the sites of ectopic vein material.

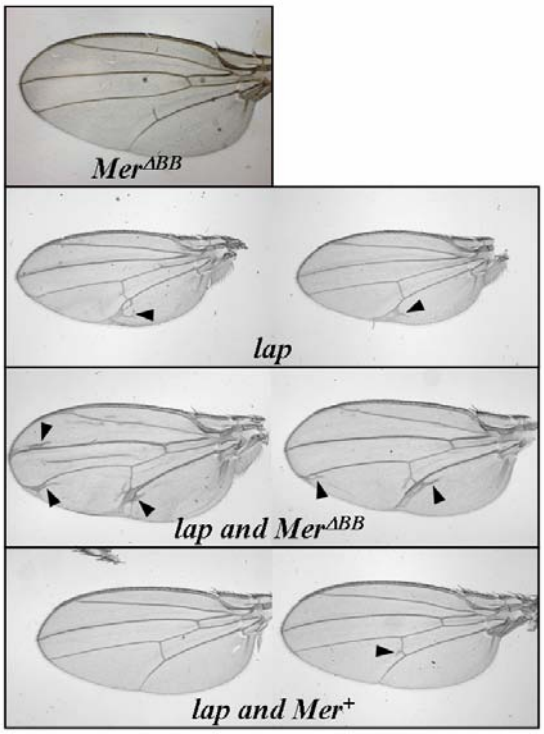


Figure 11. Model for Merlin's function through Lap to regulate the EGF pathway, which is required for vein fate determination.

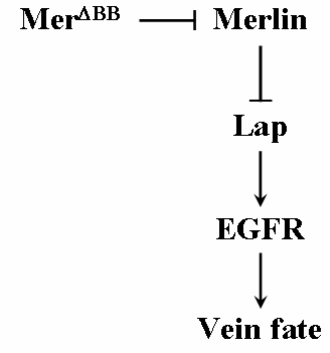
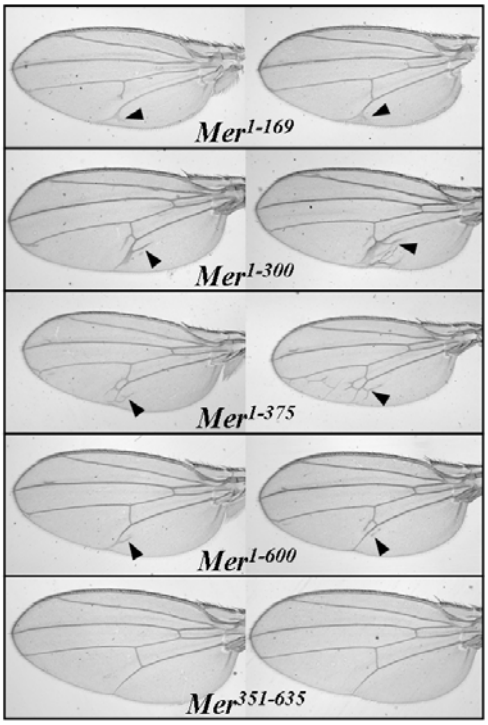


Figure 12. The region of Merlin required for the genetic interaction with Lap. Arrowheads point to ectopic vein materials.



To examine which region of Merlin was required for the genetic interaction with Lap, we simultaneously over-expressed Lap together with various truncated Merlin proteins in the wing pouch using the 1096 driver. When Mer^{1-169} was simultaneously over-expressed together with Lap in the wing disc, ectopic vein formation was still detected in the wing, similar to that seen in the case of Lap over-expression alone (Figure 12). When other N-terminal fragments of Merlin, Mer^{1-330} and Mer^{1-375}, were over-expressed together with Lap, ectopic vein materials continued to be seen. Also, when the first 600 amino acids of Merlin, Mer^{1-600}, was co-expressed with Lap in the wing disc, some ectopic veins were still present in the wing; however, the amount of ectopic vein material appeared to be reduced. Importantly, when the c-terminal fragment of Merlin, $Mer^{351-635}$, was co-expressed with Lap, the normal wing was restored, similar to what was observed when Mer^{+} was simultaneously over-expressed together with Lap. These results indicate that the C-terminal region of the Merlin protein is important for the genetic interaction with Lap. Further investigation

of this interaction will be needed to better understand the role of Merlin in vesicular trafficking.

Because the *Mer*^{3} allele was viable but sterile and because many proteins involved in exocytosis/endocytosis are also important for spermatogenesis, we previously examined the *Mer*^{3} mutant for any defects in this process. We found that the *Mer*^{3} mutant displayed abnormalities in cyst polarization during spermatogenesis. The Merlin protein normally localized in the acrosome of mature sperm, but this localization was altered in the *Mer*^{3} mutant. We now extended our finding to the *Mer*^{4} mutant. We found that the adult male pharates carrying the *Mer*^{4} allele displayed a more severe defect in cyst polarization.

We presented our findings on the Role of Drosophila Merlin in Spermatogenesis and Wg Morphogen Trafficking in the Imaginal Disc to the 2006 CTF International Consortium for the Molecular Biology of NF1, NF2, and Schwannomatosis (Omelyanchuk et al., 2006).

Task 10: To study the distribution of the Merlin protein in the wing imaginal disc, we are in the process of preparing additional antibody. A peptide with the amino acid sequence covering the putative phosphorylation site of *Drosophila* Merlin was synthesized, conjugated to a carrier protein, and used in immunizing rabbits. Antibody response against the Merlin peptide will be tested in the near future.

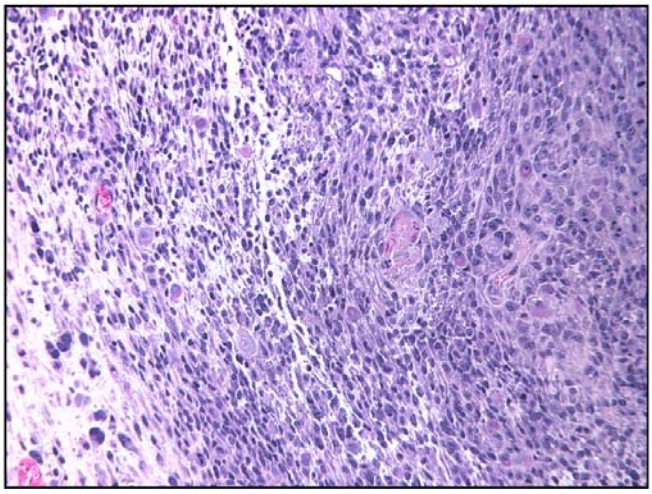
Aim 4: To investigate whether *NF2*$^{-/-}$ mouse schwannoma cells also show cytological abnormalities in mitosis similar to those seen in the *Drosophila* imaginal discs.

Task 11: This task was accomplished and we have obtained the *Nf2*$^{flox2/flox2}$ and P0Cre mice for the experiments described below.

Task 12: To generate a conditional *Nf2* knockout in Schwann cells, the *Nf2*$^{flox2/flox2}$ mice were mated with the P0Cre mice. Expression of the Cre recombinase in Schwann cells (specified by the myelin protein P0 promoter) would result in deletion of exon 2 of the *Nf2* gene (P0Cre;*Nf2*$^{\Delta E2/\Delta E2}$). Giovannini et al. (2000) previously showed that conditional inactivation of the *Nf2* gene in Schwann cells in mice could lead to Schwann cell hyperplasia and schwannoma formation, and the tumors usually appeared 10 months or later after *Nf2* inactivation. We have observed that the P0Cre;*Nf2*$^{\Delta E2/\Delta E2}$ mice developed schwannomas at about one year of age. Histopathological examination revealed that the tumor was a grayish, creamy globoid mass. Histologically consisted of spindle-shaped cells with mitotic figures (Figure 13). We are in the process of harvesting schwannoma tumors for cytological analysis.

Task 13: We have also bred a colony of *Nf2*$^{flox2/flox2}$ (phenotypically normal) and *Nf2*$^{flox2/\Delta E2}$ (heterozygous *Nf2*) mice. We are planning to isolate dorsal root ganglion from these mice for the isolation of Schwann cells. Alternatively we will harvest sciatic nerve for Schwann cell preparation.

Figure 13. Histological analysis of a schwannoma produced in the P0Cre;*Nf2*$^{\Delta E2/\Delta E2}$ *mouse with a deletion of Exon 2 of the Nf2 gene in Schwann cell. Note that the tumor consisted of actively growing, spindle-shaped cells.*



Task 14: We will conduct the cytological analysis of schwannoma cells and Schwann cells in year 3.

Task 15: During the past year, one research abstract was presented to the 2006 CTF International Consortium for the Molecular Biology of NF1, NF2, and Schwannomatosis. To extend our knowledge about potential merlin functions in various species, we analyzed the evolution, diversity, and overall distribution of merlin among different taxa. Our results demonstrate a monophyletic origin of the merlin proteins with their root in the early metazoa. The overall similarity among the primary and secondary structures of all merlin proteins and the conservation of several functionally important residues suggest a universal role for merlin in a wide range of metazoa. These findings were published recently (Golovnina et al., 2005).

KEY RESEARCH ACCOMPLISHMENTS:

(1) Two types of mitosis exit asynchrony, the asynchronous anaphase-telophase figures and the asynchronous telophase-interphase figures, were frequently observed in various *Merlin* mutants, *Mer*^{1} (Gln$^{324\rightarrow stop}$), *Mer*^{2} (Gln$^{318\rightarrow stop}$), *Mer*^{3} (Met$^{177\rightarrow Ile}$), and *Mer*^{4} (Gln$^{170\rightarrow stop}$). The asynchronous mitosis exit phenotype in the *Mer*^{4} genetic background could be corrected by the introduction of a wild-type *Merlin* transgene was substantially diminished. These results suggest that Merlin is important for the control of mitosis exit

(2) Examination of the effect of *Merlin* mutation on cell proliferation in different regions of the wing imaginal disc reveals that cells lacking Merlin possess greater ability to overcome vein restriction. The results further support the role of Merlin in cell motility, cell adhesion and cell proliferation.

(3) We showed that the Merlin and Wg proteins were colocalized in the cells at the D/V compartment border of the wing imaginal disc. Merlin inactivation may lead to an alteration on the determination/maintenance of Wg stripe expression.

(4) We have extended our analysis on the potential genetic interactions between Merlin and the proteins involved in vesicular trafficking. In addition to the potential genetic interaction between the *Merlin* and *porcupine* genes and between the *Merlin* and *shibire* genes, an interaction between the *Merlin* and *lap* (*like-AP180*), which encodes a clathrin adaptor protein important for clathrin-mediated endocytosis of synaptic vesicles, was identified. Our results indicate that Merlin counteracts with Lap and suggest that through Lap, Merlin may regulate the EGFR pathway required for vein fate determination in the wing.

(5) To better understand merlin function across different species, we conducted a study on the evolution, diversity, and overall distribution of merlin among different taxa (Golovnina et al., 2005). We found a monophyletic origin of the merlin proteins with their root in the early metazoa. Also, we compared the overall similarity among the primary and secondary structures of all merlin proteins and identified several conserved and functionally important residues. Our results suggest a universal role for merlin in a wide range of metazoa.

REPORTABLE OUTCOMES:

One research abstract was presented to the 2006 CTF International Consortium for the Molecular Biology of NF1, NF2, and Schwannomatosis. Also, one research paper was published in BMC *Evolutionary Biology*.

Abstract

(1) Omelyanchuk, L.V., Dorogova, N.V., Kopyl, S., Akhmameteva, E.M., Fehon, R.G., and Chang, L.-S. 2006. The Role of *Drosophila* Merlin in Spermatogenesis and Wg Morphogen Trafficking in the Imaginal Disc. Abstract presented to the 2006 CTF International Consortium for the Molecular Biology of NF1, NF2, and Schwannomatosis.

Previously we reported that the viable, but sterile *Merlin* mutant Mer^3 ($Met^{177}\rightarrow Ile$) displayed abnormalities in cyst polarization during spermatogenesis. The Merlin protein normally localized in the acrosome of mature sperm, but this localization was altered in the Mer^3 mutant. We now showed that a more severe defect in cyst polarization could be seen in the adult male pharates carrying the Mer^4 allele ($Gln^{170}\rightarrow stop$). We also found that both the clathrin mutant *Chc4*, known to have abnormal cyst individualization, and the meiotic mutant *ff16* displayed defective nuclei polarization and nuclear shaping. Because Merlin has been shown to associate with endocytic compartments and because mutations in the genes, such as *clathrin* and *ff16*, that are known to be important for vesicle formation and cytokinesis, also affect

nuclei polarization, we examined whether Merlin is involved in the vesicular traffic in somatic tissues. We first examined potential interaction between *Merlin* and *shibere*, a dynamin participating in various microtubule-mediated processes such as cytokinesis and endocytosis. Ectopic expression of a dominant-negative mutant of *shibere* (*shi*^{DN}), *Shi*^{K44A}, by the 1096 wing pouch driver led to a disrupted wing morphology including the loss of MTR staut bristles. The *shiDN* wing margin phenotype was rescued by simultaneous introduction of a UAS-*Mer* construct carrying the *Mer*$^{+}$ or *Mer*$^{1-600}$ transgene. Partial restoration of the phenotype was also seen when truncated *Merlin* expression constructs *Mer*$^{1-330}$ and *Mer*$^{1-375}$ were used, while no restoration was found with *Mer*$^{1-169}$ and *Mer*$^{\Delta BB}$. These results suggest that Merlin plays a role in the vesicular traffic and the FERM domain including the Blue Box is required for the interaction with Shibere. The MTR staut bristles are derived from the cells normally expressing wingless (Wg) morphogen whose movement through the tissue is related to the vesicular traffic. We found that the expression pattern of *patched (ptc),* which marks the A/P compartment border in the wing imaginal disc, was not changed in the *Mer*^{4} mutant, suggesting that the Decapentaplegic (Dpp) morphogen trafficking is likely not affected by *Merlin* mutation. In contrast, the stripe expression pattern of *Wg* at the D/V compartment border was altered in the *Mer*^{4} mutant. Intriguingly, *wg-lacZ* insertion in the *Mer*^{4} background revealed no change in the *wg* regulatory zone as compared with that in the wild-type control. While the expression of *neuralized*, which participates in the determination of dTR and vTR bristles, was controlled by Wg, its expression pattern was significantly deviated in the *Mer*^{4} mutant. In addition, the Wg-regulated *cycE* expression pattern at the D/V border was also affected by the *Mer*^{4} mutation. These results suggest that Merlin plays an important role in Wg trafficking. It has been shown that Porcupine (Porc) facilitates Wg glycosylation in the endoplasmic reticulum and the wing margin is subjected to Wg regulation. We found that both *Mer*^{3} and *Mer*^{4} mutations did not significantly affect the wing margin morphology and only caused additional sensory bristles within the row of staut bristles. While overexpression of *Mer*$^{+}$ or *Mer*$^{\Delta BB}$ did not change the wing margin pattern, overexpression of Porc led to complete disappearance of Staut bristles and irregularities of sensory and mechano-sensory bristles. In contrast, simultaneous overexpression of *Mer* and *porc* restored the staut bristles and normalized the arrangement of mechano- and chemo-sensory bristles. In addition, while overexpression of *Mer*$^{\Delta BB}$ did not affect *Wg* transcription and protein expression in the wing imaginal disc, *porc* overexpression results in complete absence of the Wg protein at the D/V border. Importantly, overexpression of both *porc* and *Mer*$^{+}$ or *Mer*$^{\Delta BB}$ restored the Wg stripe at the D/V border. These results suggest that overexpression of *Merlin* may facilitates Wg secretion. Together, our data support the notion that Merlin participates in the vesicular traffic.

Publication

(1) Golovnina, K., Blinov, A., Akhmametyeva, E.M., Omelyanchuk, L.V.,

and Chang, L.-S. 2005. Evolution and Origin of Merlin, the Product of the Neurofibromatosis Type 2 Tumor-Suppressor Gene. BMC Evolutionary Biology 5:69-86.

In this paper, we examined the evolution, diversity, and overall distribution of merlin among different taxa. By combining bioinformatic and phylogenetic approaches, we demonstrate that merlin homologs are present across a wide range of metazoan lineages. While the phylogenetic tree shows a monophyletic origin of the ERM family, the origin of the merlin proteins is robustly separated from that of the ERM proteins. The derivation of merlin is thought to be in early metazoa. We have also observed the expansion of the ERM-like proteins within the vertebrate clade, which occurred after its separation from Urochordata (Ciona intestinalis). Amino acid sequence alignment reveals the absence of an actin-binding site in the C-terminal region of all merlin proteins from various species but the presence of a conserved internal binding site in the N-terminal domain of the merlin and ERM proteins. In addition, a more conserved pattern of amino acid residues is found in the region containing the so-called "Blue Box," although some amino acid substitutions in this region exist in the merlin sequences of worms, fish, and Ciona. Examination of sequence variability at functionally significant sites, including the serine-518 residue, the phosphorylation of which modulates merlin's intra-molecular association and function as a tumor suppressor, identifies several potentially important sites that are conserved among all merlin proteins but divergent in the ERM proteins. Secondary structure prediction reveals the presence of a conserved α–helical domain in the central to C-terminal region of the merlin proteins of various species. The conserved residues and structures identified correspond to the important sites highlighted by the available crystal structures of the merlin and ERM proteins. Furthermore, analysis of the merlin gene structures from various organisms reveals the increase of gene length during evolution due to the expansion of introns; however, a reduction of intron number and length appears to occur in the merlin gene of the insect group. Our results demonstrate a monophyletic origin of the merlin proteins with their root in the early metazoa. The overall similarity among the primary and secondary structures of all merlin proteins and the conservation of several functionally important residues suggest a universal role for merlin in a wide range of metazoa.

CONCLUSIONS:

The role of Merlin in the control of mitosis exit has been confirmed by the use of various *Merlin* mutants. Merlin mutations lead to two types of mitosis exit asynchrony, the asynchronous anaphase-telophase figures and the asynchronous telophase-interphase figures. Cells lacking Merlin also possess greater ability to overcome vein restriction. The Merlin protein is colocalized with the Wingless morphogen in the cells at the dorsal/ventral compartment border of the wing imaginal disc. Merlin inactivation may lead to an

alteration on the determination/maintenance of Wg stripe expression. In addition to the potential genetic interaction between the *Merlin* and *porcupine* genes and between the *Merlin* and *shibire* genes, an interaction between the *Merlin* and *lap* (*like-AP180*), which encodes a clathrin adaptor protein important for clathrin-mediated endocytosis of synaptic vesicles, was identified. Our results indicate that Merlin counteracts with Lap and through Lap, Merlin may regulate the EGFR pathway required for vein fate determination in the wing. By analyzing the evolution, diversity, and overall distribution of merlin among different taxa, we demonstrate a monophyletic origin of the merlin proteins with their root in the early metazoa. The overall similarity among the primary and secondary structures of all merlin proteins and the conservation of several functionally important residues suggest a universal role for merlin in a wide range of metazoa.

REFERENCES:

Ashburner, M.A. 1989. *Drosophila*, A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Chang, L.-S., Akhmametyeva, E.M., Mihaylova, M., Luo, H., Tae, S., Neff, B., Jacob, A., and Welling, D.B. 2005. Dissecting the molecular pathways in vestibular schwannoma tumorigenesis. Recent Res. Devel. Genes & Genomes 1:1-33.

Entchev, E.V. and Gonzalez-Gaitan, M.A. 2002. Morphogen gradient formation and vesicular trafficking. Traffic 3:98-109.

Fehon, R.G., Oren, T., LaJeunesse, D.R., Melby, T.E., McCartney. B.M. 1997. Isolation of mutations in the Drosophila homologues of the human Neurofibromatosis 2 and yeast CDC42 genes using a simple and efficient reverse-genetic method. Genetics 146:245-252.

Garcia-Bellido, A. and Merriam, J.R. 1971a. Parameters of the wing imaginal disc development of *Drosophila melanogaster*. Dev. Biol. 24:61-87.

Garcia-Bellido, A. and Merriam, J.R. 1971b. Genetic analysis of cell heredity in imaginal discs of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 68:2222-2226.

Golovnina, K., Blinov, A., Akhmametyeva, E.M., Omelyanchuk, L.V., and Chang, L.-S. 2005. Evolution and Origin of Merlin, the Product of the *Neurofibromatosis Type 2* Tumor-Suppressor Gene. BMC Evol. Biol. 5:69-86.

Gonzalez-Gaitan, M., Capdevila, M.P., and Garcia-Bellido, A. 1994. Cell proliferation pattern in the wing imaginal disc of *Drosophila*. Mech. Dev. 40:183-200.

Giovannini, M., Robanus-Maandag, E., van der Valk, M., Niwa-Kawakita, M., Abramowski, V., Goutebroze, L., Woodruff, J.M., Berns, A., and Thomas, G. 2000. Conditional biallelic *Nf2* mutation in the mouse promotes manifestations of human neurofibromatosis type 2. Genes Dev. 14:1617-1630.

Hofmann, K. 2000. A superfamily of membrane-bound O-acyltransferases with implications for Wnt signaling. *Trends Biochem. Sci.* 25:111–112.

Johnston L.A. and Edgar B.A. 1998 Wingless and Notch regulate cell-cycle arrest in the developing Drosophila wing. Nature 394:82-84.

Kitamoto, T. 2002. Targeted expression of temperature-sensitive dynamin to study neural mechanisms of complex behavior in Drosophila. J. Neurogenet. 16:205-228.
LaJeunesse, D.R., McCartney, B.M., Fehon, R.G. 1998. Structural analysis of Drosophila merlin reveals functional domains important for growth control and subcellular localization. J. Cell Biol. 141:1589-1599.
Maitra, S., Kulikauskas, R.M., Gavilan, H., and Fehon, R.G. 2006. The tumor suppressors Merlin and expanded function cooperatively to modulate receptor endocytosis and signaling. Curr. Biol. 16:702-709.
McCartney, B.M., Fehon, R.G. 1996. Distinct cellular and subcellular patterns of expression imply distinct functions for the Drosophila homologues of moesin and the neurofibromatosis 2 tumor suppressor, merlin. J. Cell Biol. 133:843-852.
McCartney, B.M., Kulikauskas, R.M., LaJeunesse, D.R., and Fehon, R.G. (2000). The Neurofibromatosis-2 homologue, Merlin, and the tumor suppressor expanded function together in Drosophila to regulate cell proliferation and differentiation. Development 127:1315–1324.
Neff, B.A., D.B. Welling, E.M. Akhmametyeva, and L.-S. Chang. 2006. The Molecular Biology of Vestibular Schwannomas: Dissecting the Pathogenic Process at the Molecular Level. Otol. Neurotol. 27:197-208
Omelyanchuk, L.V., Dorogova, N.V., Kopyl, S., Akhmameteva, E.M., Fehon, R.G., and Chang, L.-S. 2006. The Role of *Drosophila* Merlin in Spermatogenesis and Wg Morphogen Trafficking in the Imaginal Disc. Abstract presented to the 2006 CTF International Consortium for the Molecular Biology of NF1, NF2, and Schwannomatosis.
Seto, E.S. and Bellen, H.J. 2004. The ins and outs of Wingless signaling. Trends Cell Biol. 14:45-53.
Trofatter JA, MacCollin MM, Rutter JL, Murrell JR, Duyao MP, Parry DM, Eldridge R, Kley N, Menon AG, Pulaski K, Haase VH, Ambrose CM, Munroe D, Bove C, Haines JL, Martuza RL, MacDonald ME, Seizinger BR, Short MP, Buckler AL, Gusella JF. 1993. A novel Moesin-, Exrin-, Radixin-like gene is a candidate for the neurofibromatosis 2 tumor-suppressor. Cell 72:791-800.
van der Bliek, A.M. and Meyerowitz, E.M. 1991. Dynamin-like protein encoded by the Drosophila shibire gene associated with vesicular traffic. Nature 351:411-414.
Zhang, B., Koh, Y.H., Beckstead, R.B., Budnik, V., Ganetzky, B., and Bellen, H.J. 1998. Synaptic vesicle size and number are regulated by a clathrin adaptor protein required for endocytosis. Neuron 21:1465-1475.

ABSTRACT

Presently, the mechanism by which Merlin functions as a tumor suppressor is not understood. By utilizing *Drosophila* genetics, we have found a role of Merlin in the control of mitosis exit. Merlin mutations lead to two types of mitosis exit asynchrony, the asynchronous anaphase-telophase figures and the asynchronous telophase-interphase figures. Also, we show that cells lacking Merlin possess greater ability to overcome vein restriction in the wing. The Merlin protein is colocalized with the Wingless morphogen in the cells at the dorsal/ventral compartment border of the wing imaginal disc. Merlin inactivation may lead to an alteration on the determination/maintenance of Wg stripe expression. We have found potential genetic interactions between the *Merlin* and *porcupine* genes and between the *Merlin* and *shibire* genes. We also discover an interaction between the *Merlin* and *lap* (*like-AP180*), which is important for clathrin-mediated endocytosis of synaptic vesicles, was identified. Our results suggest that Merlin counteracts with Lap and through Lap, Merlin may regulate the EGFR pathway required for vein fate determination in the wing. In addition, by analyzing the evolution, diversity, and overall distribution of merlin among different taxa, we demonstrate a monophyletic origin of the merlin proteins with their root in the early metazoa. The overall similarity among the primary and secondary structures of all merlin proteins and the conservation of several functionally important residues suggest a universal role for merlin in a wide range of metazoa.

APPENDICES:

One Publication

Golovnina, K., Blinov, A., Akhmametyeva, E.M., Omelyanchuk, L.V., and Chang, L.-S. 2005. Evolution and Origin of Merlin, the Product of the Neurofibromatosis Type 2 Tumor-Suppressor Gene. BMC Evolutionary Biology.

BMC Evolutionary Biology

BioMed Central

Research article

Evolution and origin of merlin, the product of the *Neurofibromatosis type 2* (*NF2*) tumor-suppressor gene

Kseniya Golovnina[1], Alexander Blinov[1], Elena M Akhmametyeva[2], Leonid V Omelyanchuk[1] and Long-Sheng Chang*[2]

Address: [1]Institute of Cytology and Genetics, Russian Academy of Sciences, 10 Lavrent'ev Ave., 630090, Novosibirsk, Russia and [2]Center for Childhood Cancer, Children's Research Institute, Children's Hospital and Department of Pediatrics, The Ohio State University, 700 Children's Drive, Columbus, OH 43205-2696, USA

Email: Kseniya Golovnina - ksu@bionet.nsc.ru; Alexander Blinov - blinov@bionet.nsc.ru; Elena M Akhmametyeva - akhmamee@pediatrics.ohio-state.edu; Leonid V Omelyanchuk - ome@bionet.nsc.ru; Long-Sheng Chang* - lchang@chi.osu.edu

* Corresponding author

Published: 02 December 2005

BMC Evolutionary Biology 2005, **5**:69 doi:10.1186/1471-2148-5-69

Received: 18 March 2005
Accepted: 02 December 2005

This article is available from: http://www.biomedcentral.com/1471-2148/5/69

Abstract

Background: Merlin, the product of the *Neurofibromatosis type 2* (*NF2*) tumor suppressor gene, belongs to the ezrin-radixin-moesin (ERM) subgroup of the protein 4.1 superfamily, which links cell surface glycoproteins to the actin cytoskeleton. While merlin's functional activity has been examined in mammalian and *Drosophila* models, little is understood about its evolution, diversity, and overall distribution among different taxa.

Results: By combining bioinformatic and phylogenetic approaches, we demonstrate that merlin homologs are present across a wide range of metazoan lineages. While the phylogenetic tree shows a monophyletic origin of the ERM family, the origin of the merlin proteins is robustly separated from that of the ERM proteins. The derivation of merlin is thought to be in early metazoa. We have also observed the expansion of the ERM-like proteins within the vertebrate clade, which occurred after its separation from Urochordata (*Ciona intestinalis*). Amino acid sequence alignment reveals the absence of an actin-binding site in the C-terminal region of all merlin proteins from various species but the presence of a conserved internal binding site in the N-terminal domain of the merlin and ERM proteins. In addition, a more conserved pattern of amino acid residues is found in the region containing the so-called "Blue Box," although some amino acid substitutions in this region exist in the merlin sequences of worms, fish, and *Ciona*. Examination of sequence variability at functionally significant sites, including the serine-518 residue, the phosphorylation of which modulates merlin's intra-molecular association and function as a tumor suppressor, identifies several potentially important sites that are conserved among all merlin proteins but divergent in the ERM proteins. Secondary structure prediction reveals the presence of a conserved α-helical domain in the central to C-terminal region of the merlin proteins of various species. The conserved residues and structures identified correspond to the important sites highlighted by the available crystal structures of the merlin and ERM proteins. Furthermore, analysis of the merlin gene structures from various organisms reveals the increase of gene length during evolution due to the expansion of introns; however, a reduction of intron number and length appears to occur in the merlin gene of the insect group.

Conclusion: Our results demonstrate a monophyletic origin of the merlin proteins with their root in the early metazoa. The overall similarity among the primary and secondary structures of all merlin proteins and the conservation of several functionally important residues suggest a universal role for merlin in a wide range of metazoa.

Background

The advancement in genome sequencing projects, the accumulation of knowledge in bioinformatics, and the molecular genetic analysis of genes and their functions in a variety of model organisms provides us with an unprecedented opportunity to identify novel genes based on sequences related to characterized genes [1]. This process is conducted using pairwise sequence comparison with the understanding that genes form families wherein related sequences likely share similar functions. Although initial identification of new genes may not yield a clear indication of their respective functions, studies on their evolution may allow validation of their sequence identity and provide information on their putative functional characteristics. For genes evolved from duplication and/or adapted to different evolutionary niches during speciation, detailed sequence comparison can provide additional information regarding their biological and biochemical characteristics [2].

Neurofibromatosis type 2 (NF2) is a highly penetrant, autosomal dominant disorder, whose hallmark is the development of bilateral vestibular schwannomas [3,4]. The tumor suppressor gene associated with NF2 has been identified and termed the *neurofibromatosis type 2* gene (*NF2*) [5,6]. The *NF2* gene encodes a protein named merlin, for moesin-ezrin-radixin like protein, or schwannomin, a word derived from schwannoma, the most prevalent tumor seen in NF2. For simplicity, we refer to the *NF2* gene product as merlin hereafter.

Merlin shares sequence similarity with the ezrin, radixin, and moesin (ERM) proteins, which belong to the protein 4.1 superfamily of cytoskeleton-associated proteins that link cell surface glycoproteins to the actin cytoskeleton [7,8]. Like ERM proteins, merlin consists of three predicted structural domains [5,6,9]. The N-terminal domain, termed the FERM (F for 4.1) domain, is highly conserved among all members of the ERM family and is important for interactions with cell surface glycoproteins, including CD44 and intercellular adhesion molecules [10-13]. Crystal structure analysis shows that the tertiary structure of the FERM domain of merlin closely resembles those of the FERM domain of moesin and radixin [14-18]. The FERM domain of merlin exists as a clover-shaped molecule consisting of three structural subdomains A, B, and C, which are homologous to lobes F1, F2, and F3 in moesin and radixin. Subdomain A, composed of residues 20–100, possesses a ubiquitin-like fold. Subdomain B, consisting of residues 101–215, folds itself into a topology like that of the acetyl-CoA-binding protein. Subdomain C, containing residues 216–313, adopts the pleckstrin homology/phosphotyrosine-binding fold found in a broad range of signaling molecules [14-16]. The second half of merlin contains a predicted α-helix domain, which is also present in the ERM proteins [19]. Although the unique C-terminus of merlin lacks the conventional actin-binding domain found in the ERM proteins [20,21], merlin can directly bind actin using the residues at the N-terminal domain and indirectly through its association with βII-spectrin or fodrin [22-24].

The merlin and ERM proteins are thought to be key regulators of interactions between the actin cytoskeleton and the plasma membrane in polarized cells. They act as important members of signal transduction pathways that control cell growth and participate in the sorting of membrane proteins during exocytic traffic [25,26]. However, unlike the ERM proteins, merlin has a distinct function as a tumor suppressor [27]. Growth suppression by merlin is dependent on its ability to form intramolecular associations [28,29]. In this regard, merlin exists in an "open" (inactive form) or "closed" (active growth-suppressive form) conformation that is regulated by phosphorylation [30-35].

While previous studies have focused primarily on the functional analysis of merlin, limited information is available about its overall distribution across eukaryotes and its evolution. A phylogenetic study indicates that the FERM domains of ERM homologs from sea urchins, *Caenorhabditis elegans*, *Drosophila melanogaster*, and vertebrates share 74–82% amino acid identity and have about 60% identity with those of merlin [25,36-42]. These levels of identity are exceptionally high, suggesting that the protein structures of the merlin and ERM proteins from these species may be well-conserved. The most divergent ERM proteins are found in tapeworms and schistosomes [36-39]. The FERM domains of these parasite proteins share only 44–58% similarity to their vertebrate homologs. The high degree of structural conservation among these proteins points to possible similarities or functional redundancies. Intriguingly, no FERM domain-encoding genes have been identified in the genome of the yeast *Saccharomyces cerevisiae*, implying that FERM domains evolved in response to multicellularity, rather than as a cytoskeletal component [25].

The goal of the present study was to expand our understanding of the taxonomic diversity of merlin and the phylogenetic relationships using experimentally annotated and predicted sequences. By the integration of the BLAST-based analysis using the available partial and whole genome sequences with phylogeny reconstruction, we have generated an evolutionary tree for the entire ERM-family members from various taxa and identified some interesting details about their phylogenetic origin. In addition, we compared sequence variability at functionally significant sites, including the major phosphorylation site of merlin, predicted the secondary structure of the

Table 1: The list of the predicted and experimentally annotated merlin and ERM proteins included in this study.

Species	Proteins	UniProtKB/Swiss-Prot Identifiers	GenBank Accession No.	Entries from Genome Sequencing Projects	Related Resources
Homo sapiens	merlin (NF2)	P35240	AAA36212		http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gmd=Retrieve&dopt=Overview&list_uids=9558
	ezrin	P15311	CAA35893		
	radixin	P35241	AAA36541		
	moesin	P26038	AAA36322		
Pan troglodytes	similar to NF2		XP_515061		http://www.hgsc.bcm.tmc.edu/projects/chimpanzee/
Papio anubis	merlin	P59750	AAO23133		http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gmd=Retrieve&dopt=Overview&list_uids=12965
Bos taurus	ezrin	P31976	AAA30510		http://www.hgsc.bcm.tmc.edu/projects/bovine/
Sus scrofa	radixin	P26044	AAB02865		http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
	moesin	P26042	AAB02864		
Canis familiaris	similar to NF2		XP_534729		http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
Oryctolagus cuniculus	ezrin	Q8HZQ5	AAN06818		http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gmd=Retrieve&dopt=Overview&list_uids=12818
Mus musculus	ezrin	P26040	CAA43086		http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
	radixin	P26043	CAA43087		
	merlin	P46662	CAA52737		
Rattus norvegicus	ezrin	P31977	AAR91694		http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
	NF2		XP_341249		
Gallus gallus	ezrin	Q9YGW6	BAA75497		http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
	radixin	Q9PU45	CAB59977		
	merlin		NP_989828		
Xenopus laevis	unknown protein		AAH77822		http://www.xenbase.org/
Danio rerio	nf2a	Q6Q413	AAS66973		http://www.ensembl.org/Danio_rerio/
	moesin	Q503E6	AAH95359		
Fugu rubripes	radixin			FRUP00000132603	http://genome.jgi-psf.org/
	moesin			FRUP00000156313	
	merlin			FRUP00000136298	
Tetraodon nigroviridis	unnamed protein 1		CAG08868		http://www.ensembl.org/Tetraodon_nigroviridis/
	unnamed		CAG08250		

Table 1: The list of the predicted and experimentally annotated merlin and ERM proteins included in this study. *(Continued)*

	protein 2				
Ciona intestinalis	erm-like			ci0100149701	http://genome.jgi-psf.org/
	merlin-like			ci0100130636	
Ciona savignyi	merlin-like			paired_scaffold_109	http://www.broad.mit.edu/ftp/
Biomphalaria glabrata	erm-like		AAK61353		http://biology.unm.edu/biomphalaria-genome/
Lytechinus variegates	moesin	P52962	AAC46514		http://www.hgsc.bcm.tmc.edu/projects/seaurchin/
Apis mellifera	similar to schwannomin		XP_392673		http://racerx00.tamu.edu/PHP/bee_search.php
Drosophila melanogaster	merlin	Q24564	AAB08449		http://fbserver.gen.cam.ac.uk:7081/
	moesin	P46150	AAB48934		
Drosophila yakuba	merlin-like			predicted in this work	http://genome.wustl.edu/blast/client.pl
Anopheles gambiae	merlin-like fragment		EAA07087		http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
Caenorhabditis elegans	erm1a	P91015	AAB37643		http://www.wormbase.org/
	erm1b	P91016	AAB37642		
	nfm 1a	Q20307	AAA19073		
	nfm 1b	Q95QG5	AAK68385		
Caenorhabditis briggsae	erm-like			BP:CBP03133	http://www.wormbase.org/
	nfm1			BP:CBP05025	
Caenorhabditis remanie	merlin-like			predicted in this work	http://genome.wustl.edu/blast/client.pl
	erm-like				
Brugia malayi	merlin-like			316.m00022	http://www.tigr.org/tdb/e2k1/bma1/
Schistosoma japonicum	JF2		AAB49033		http://www.nhm.ac.uk/hosted_sites/schisto/
Taenia saginata	myosin-like	Q94815	CAA65728		http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=protein
Echinococcus multilocularis	EM10		A45620		http://www.sanger.ac.uk/Projects/Echinococcus/
Echinococcus granulosus	EG10	Q24796	CAA82625		
Phanerochaete chrysosporium	---				http://genome.jgi-psf.org/whiterot1/whiterot1.home
Aspergillus flavus	---				http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
Arabidopsis thaliana	---				http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
Oryza sativa	---				http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
Trypanosoma brucei	---				http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?spec
Cryptosporidium parvum	---				http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=c_parvum

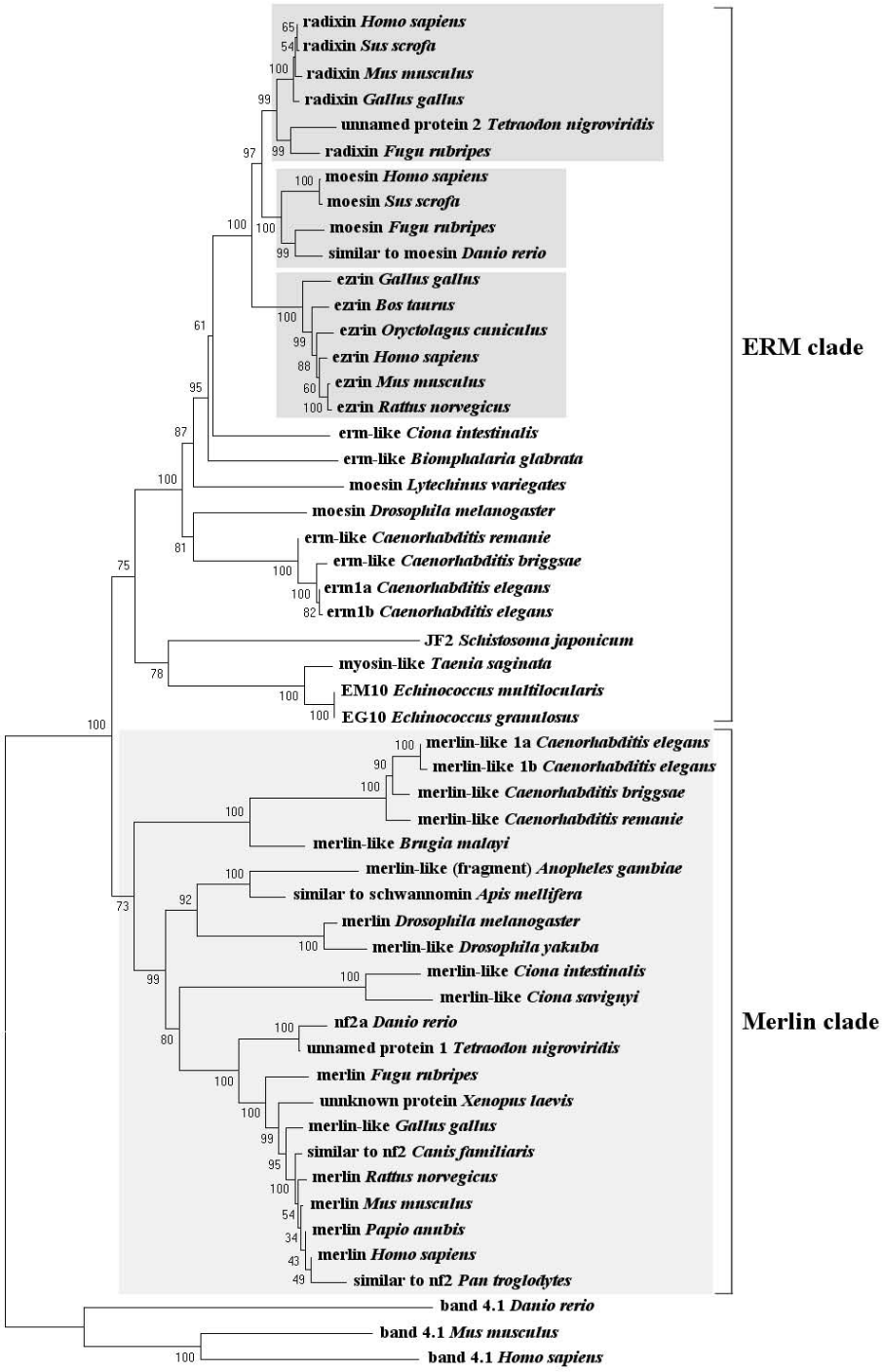


Figure 1

The neighbor-joining tree of the ERM family. The diagram illustrates the basic resolution of the ERM-family members into two major clades, merlin and ERM. Bootstrap support values are shown above each node. Shaded boxes denote different subgroups of the ERM clade in vertebrates, which appeared after the expansion of the ERM-like ancestor. The *Tetraodon nigroviridis* "unnamed protein 1 and 2" sequences (GenBank Accession No. CAG08868 and CAG08250, respectively) and the *Xenopus laevis* "unknown protein" sequence (GenBank Accession No. AAH77822) were grouped based on their similarity to the merlin or ERM sequences.

merlin proteins of various species, and examined the exon-intron structural evolution of the *NF2* gene.

Results and Discussion

BLAST identification of merlin sequences

To identify putative merlin and ERM sequences in a wide range of eukaryotes, we performed BLAST analysis of 15 available genome databases. By searching through all annotated proteins and genome sequences, we identified 50 sequences from 30 species. Table 1 summarizes the full list of the predicted and annotated merlin and ERM proteins identified, and their GenBank and available UniProtKB/Swiss-Prot accession numbers and related resources. No merlin-like sequences were found in the genomes of fungi, plants, and protozoa. While the sequencing projects of the hard ticks are still ongoing at The Institute for Genomic Research (TIGR), amino acid sequences deduced from partial cDNAs of salivary glands, which share a similarity with the FERM domain of merlin, have been noted from *Rhipicephalus appendiculatus* [43], *Amblyomma variegatum* [44], and *Boophilus microplus* [45].

Assembly of predicted merlin sequences from whole genome shotgun

To date, the genomes of *Caenorhabditis remanei and Drosophila yakuba* are represented by a set of contigs [46]. When contigs are ordered, oriented, and positioned with respect to each other by mate-pair reads, they are described as a scaffold. Scaffolds are the main product of the Whole Genome Shotgun strategy and can be assigned to chromosomes using chromosome-specific markers. Although the extensive scaffolds for the genomes of *Caenorhabditis remanei* and *Drosophila yakuba* are not currently available, we were able to assemble predictive protein sequences, which most resemble the merlin sequence of the closely-related organism, *Caenorhabditis elegans* or *Drosophila melanogaster*, respectively, using TBLASTN search across the available sets of contigs. In the *Drosophila yakuba* contig 49.37, we identified a predicted merlin sequence, which is nearly identical to that of the *Drosophila melanogaster* protein with the exception of three positions at the C-terminus, two substitutions at Glu^{468}→Asp and Asn^{579}→Ser and an insertion of Lys at position 575. Also, we found three *Caenorhabditis remanei* contigs, 564.6, 2151.1, and 2151.2, which contained merlin-like sequences with similarity, ranging from 81% to 100%, to its *Caenorhabditis elegans* counterpart. It should be noted that the deduced amino acid sequences were assembled manually, and in some cases, only partial or approximate amino acid sequences could be obtained. Nevertheless, they were useful for the identification of the definite gene in the respective genome and were valuable for the following phylogenetic reconstruction in order to validate the functional relationship and evolution of the definite gene.

Construction of a phylogenetic tree for the ERM family of proteins

To understand the origin and evolution of merlin, we conducted a phylogenetic analysis of the 50 proteins of the ERM family, which were identified from 30 different taxa (Table 1) using the neighbor-joining method [47,48] combined with the molecular evolutionary genetics analysis program MEGA2 [49]. Three protein 4.1 sequences from humans, mice, and zebrafish, respectively, were used as an outgroup. By comparing the bootstrap support values, which denote the number of times a grouping occurs out of 1,000 random samples from the alignment, we constructed a phylogenetic tree for the ERM family of proteins (Figure 1). Based on this phylogenetic analysis, the entire ERM family can be subdivided into the ERM clade and the merlin clade. While both clades show a strongly supported monophyletic origin, the merlin clade can be robustly defined and separated from the ERM clade (the bootstrap support value = 100). We identified a total of 22 sequences for the merlin clade and 28 sequences for the ERM clade. The topology of the phylogenetic tree within the merlin clade appears to agree with the general concept of evolutionary history of speciation.

The merlin clade can be further divided into three groups according to the order of derivation: worms, insects, and Chordata, with the earliest separated genus, *Ciona*, in the last taxonomic unit (Figure 1). The predicted merlin-like sequence from *Caenorhabditis remanei* branched from that of *Caenorhabditis elegans*, and similarly, *Drosophila yakuba* diverged from its *Drosophila melanogaster* counterpart. Both the "unnamed protein 1" of *Tetraodon nigroviridis* and the "unknown protein" of *Xenopus laevis* from the GenBank database are clustered in the Chordata merlin-like group with high bootstrap probabilities (Figure 1), which confirms their identity as merlin homologs. The protein fragment from *Anopheles gambiae*, which bears a sequence similarity to merlin, is grouped together with the *Apis mellifera* merlin-like protein by a bootstrap support value of 100.

Although the ERM-like proteins have been identified in *Taenia saginata*, *Schistosoma japonicum*, *Echinococcus granulosus*, and *Echinococcus multilocularis* [36-39], we did not find any merlin-like sequences in the genomes of these species. The lack of merlin-like sequences in these parasite genomes may be due to incomplete genome sequences in the database; however, this explanation is unlikely because the merlin-like sequence was also not observed in the genome of *Schistosoma mansoni*, which has been rigorously studied [50]. Another possibility is that the absence of merlin-like sequences in these organisms may reflect their adaptation to a parasitic lifestyle and the reduction of various organ systems. Alternatively, the merlin protein may emerge later during evolution. Similarly, no merlin-

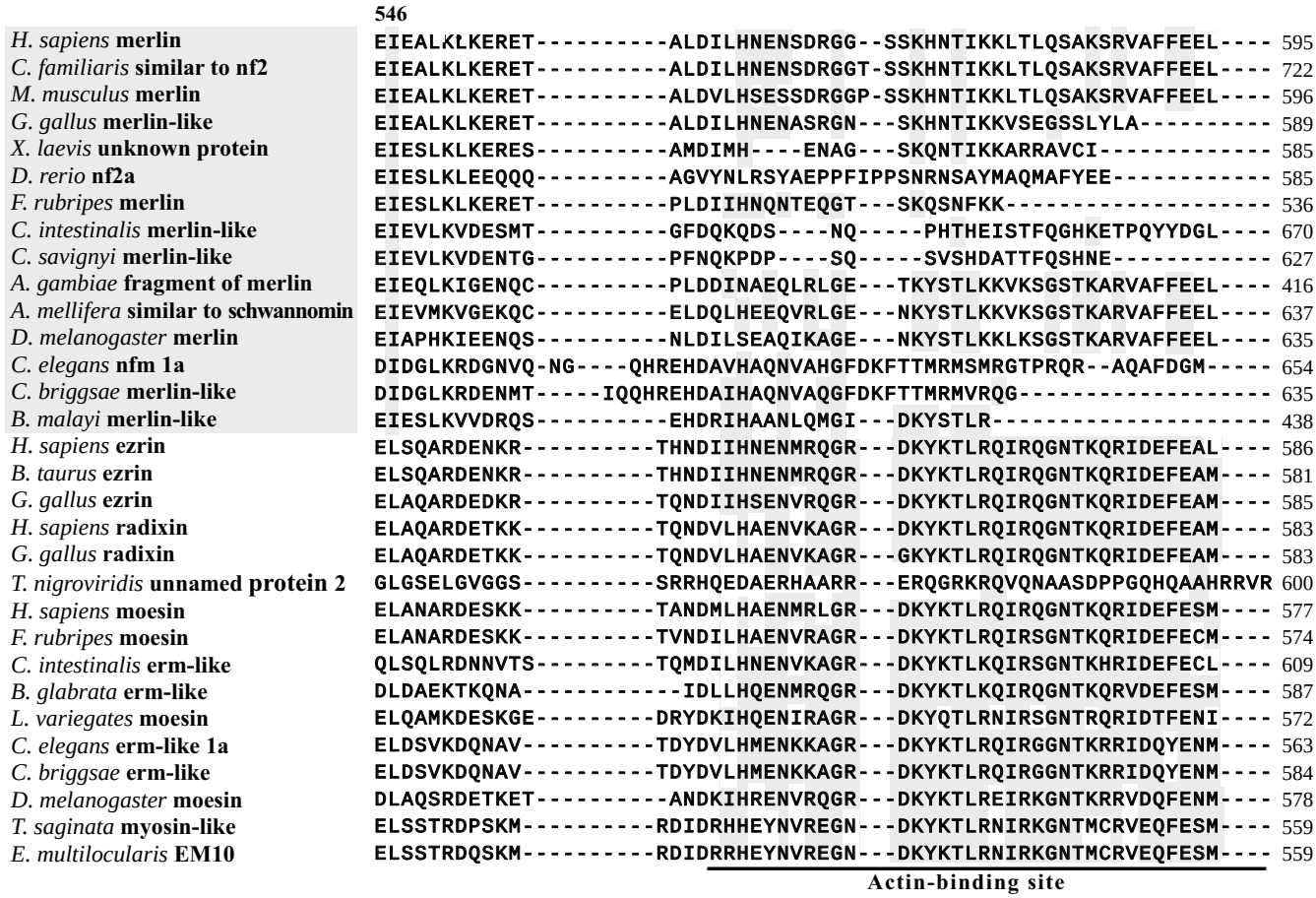


Figure 2

Sequence alignments of functionally important sites in the merlin and ERM proteins of various species. Comparison of the C-terminal region including the actin-binding site and two other predicted significant residues. Databank resources for the ERM-family proteins listed in Table 1 were used in the analysis, and only typical representatives from each group are displayed.

like sequence was found in the complete genomes of protozoa, fungi, and plants. Based on these results, we suppose that the derivation of merlin occurred in the early metazoa after its separation from flatworms.

As illustrated in the ERM clade in Figure 1, the ERM-like proteins found in parasites can be grouped together but form a separate branch from the rest of ERM proteins. Based on the phylogenetic analysis, the clustering of the "unnamed protein 2" of *Tetraodon nigroviridis* with the *Fugu rubripes* radixin protein defines it as a radixin-like protein. It should be noted that the two predicted ERM proteins, erm1a and erm1b of *Caenorhabditis elegans* [51], may represent different transcript variants of the same gene (also see below).

Furthermore, we have observed the evident expansion of the ERM-like ancestor in vertebrates (Figure 1). Since the ERM homolog of *Ciona* emerged prior to the vertebrate clade, it appears that the first duplication of the vertebrate ERM sequence occurred after its divergence from *Ciona*. Subsequent expansion within this sub-family has led to the present existence of three related groups of proteins, ezrin, radixin and moesin; among which, the ezrin group is the most ancient. Such an expanded complement may only be common to the ERM proteins of vertebrates because other metazoa have only one predicted ERM-like homolog [52-56]. Curiously, the increasing number of ERM members that occurred within the vertebrate clade paralleled the evolutionary complexity of the organism. It will be important to understand how these proteins

	Blue Box	204	
H. sapiens **merlin**	INLYQ-MTPEMWEERITAWYAEHRGRARDEAEMEYLK		209
C. familiaris **similar to nf2**	INLYQ-MTPEMWEERITAWYAEHRGRARDEAEMEYLK		318
M. musculus **merlin**	INLYQ-MTPEMWEERITAWYAEHRGRARDEAEMEYLK		209
G. gallus **merlin-like**	INLYQ-MTPEMWEERITAWYAEHRGRARDEAEMEYLK		209
X. laevis **unknown protein**	INLYQ-MTPEMWEERITAWYAEHRGRTRDEAEMEYLK		209
D. rerio **nf2a**	LMQYQ-MTPDMWEEKITAWYAEHRNITRDEAEMEYLK		201
F. rubripes **merlin**	INLYQ-MTAEMWEERITACYAEHRGRTRDEAEMEYLK		170
C. intestinalis **merlin-like**	RDQFQSVTGEMWETQITSWYAQHHGLTRDEAELEYLK		207
C. savignyi **merlin-like**	IDQYQSVTGQMWEAQITPWYAGHHGLTRDEAELEYLK		179
A. gambiae **fragment of merlin**	--QYQ-MTPQMWEERIKTWYADHRGMSRDEAEMEYLK		34
A. mellifera **similar to schwannomin**	IDQYQ-MTPEMWEDRIKIWYADHRGMSRDEAEMEYLK		201
D. melanogaster **merlin**	TDQYQ-MTPEMWEERIKTWYMDHEPMTRDEVEMEYLK		203
C. elegans **nfm-1a**	IDQYD-MSADMWRDRIKRWWSRNAGQSREEAELEYLR		203
C. briggsae **merlin-like**	IDQYD-MSADMWRDRIKRWWSRNAGQSREEAELEYLR		203
B. malayi **merlin-like**	IKQYD-MTPOMWEERIKRWWINNSGQSREDAEMEYLR		196
H. sapiens **ezrin**	MDQHK-LTRDQWEDRIQVWHAEHRGMLKDNAMLEYLK		193
B. taurus **ezrin**	MDQHK-LTRDQWEDRIQVWHAEHRGMLKDSAMLEYLK		193
G. gallus **ezrin**	MDQHK-LSRDQWEERIQVWHAEHSGMLKENAMLEYLK		193
H. sapiens **radixin**	LEQHK-LTKEQWEERIQNWHEEHRGMLREDSMMEYLK		193
G. gallus **radixin**	LEQHK-LTKEQWEERIQNWHEEHRGMLREDSMMEYLK		193
T. nigroviridis **unnamed protein 2**	LEQHK-LTKEQWEERIQTWHEEHRSMLREDAMMEYLK		189
H. sapiens **moesin**	LEQHK-LNKDQWEERIQVWHEEHRGMLREDAVLEYLK		193
F. rubripes **moesin**	LDQHK-LNKDQWEERIQVWHEEHKGMMREESMMEYLK		189
C. intestinalis **erm-like**	YEQHK-MTKEQWEERIQTWHCEHGSMTREDAMIEYLK		217
B. glabrata **erm-like**	YDQHK-LTKEQWEERIKSWYAEHKAMLREDAMIEYLK		194
L. variegates **moesin**	IEQHK-MTKEQWYERVSNWHQEHLSLSKEDAITEYMK		192
C. elegans **erm-like 1a**	LGQFK-LNSEEWERRIMTWWADHRATTREQAMLEYLK		194
C. briggsae **erm-like**	LGQFK-LNSEEWERRIMTWWADHRATTREQAMLEYLK		189
D. melanogaster **moesin**	IDQHK-MSKDEWEQSIMTWWQEHRSMLREDAMMEYLK		194
T. saginata **myosin-like**	KDQYD-QTDEQWFDRIVTYYKDHHDMSREDAMVQYLQ		195
E. multilocularis **EM10**	KEQYD-QTDEQWYERIIAYYKDHHDMSREDAMVQYLQ		195

Figure 3
Alignment of the N-terminal domain, containing the Blue Box and the amino acid residue 204, conserved among the merlin proteins but divergent in the ERM proteins.

evolved and how their functions coordinated because of the important and diverse functions of ERM proteins [8,25,26].

Evolution of the functionally important residues in merlin

Although initial identification of proteins via sequence similarities does not yield a clear indication of their respective functions, analysis of specific conserved regions and residues may provide important information regarding their putative functional characteristics. We conducted pairwise sequence comparison among all obtained merlin and ERM sequences, and identified several regions of interest. The results of the entire sequence alignment are provided in the Additional File 1 and are summarized in

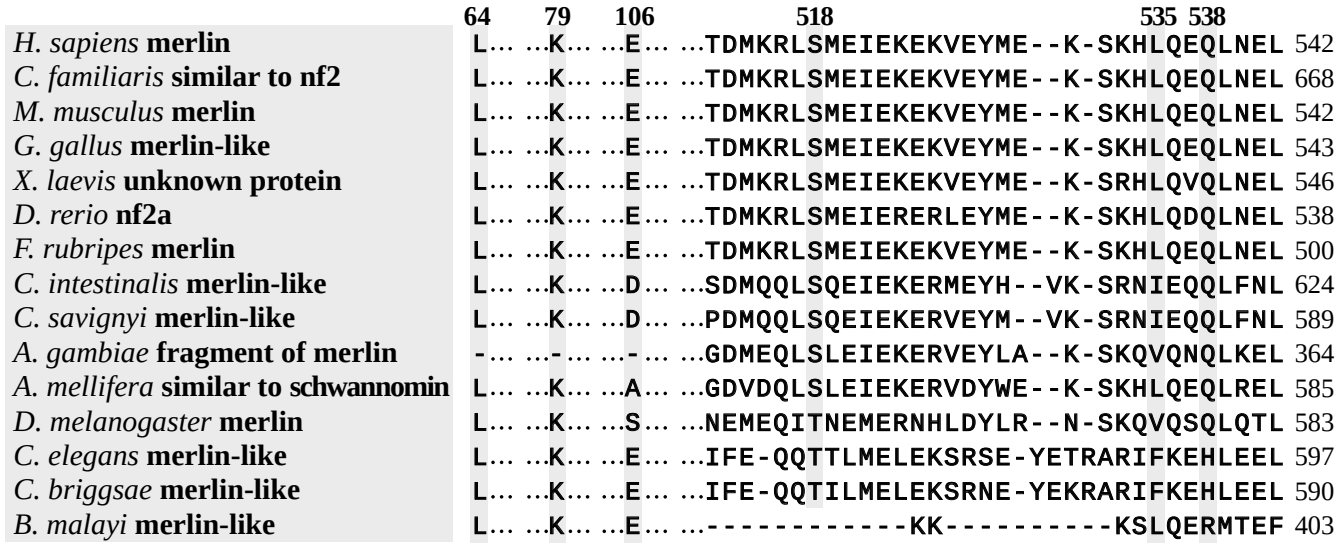


Figure 4
Sequence alignments reveal conservation of several functionally important residues, including the major phosphorylation site of the merlin group.

Figures 2, 3, and 4. Previously, the conservation of the N-terminal FERM domain among human ERM proteins and their functional importance were described [10-13]. In our alignment, we showed that this conservation extended to the merlin and ERM proteins of various species for which sequences were available to date. These data suggest a universal role for the presence of the FERM domain during evolution and further imply an existence of certain evolutionary constraints on the changes of their amino acid residues.

Although merlin lacks the C-terminal actin-binding site found in ERM proteins [7,20,21,57], it can directly interact with the actin cytoskeleton [22,58] or indirectly bind via the actin-binding protein βII spectrin/fodrin [23,24]. Sequence alignment showed extensive amino acid variability in the C-terminal region of the merlin proteins of various species, while a noncontiguous stretch of 25 amino acid residues, including the well-defined actin-binding site, was reliably aligned among all predicted ERM proteins with the exception of the "unnamed protein 2" of *Tetraodon nigroviridis* (Figure 2). According to the phylogenetic tree, the "unnamed protein 2" of *Tetraodon nigroviridis* is classified in the radixin group (Figure 1), and its sequence visibly differs from other radixin proteins only at the C-terminus. The reason for this sequence variability is presently unknown. It may be due to an inaccuracy in sequence assembly from the scaffold. Alternatively, the "unnamed 2 protein" may possess a unique characteristic and will be of considerable interest for functional comparison with other radixin proteins.

Sequence variability at the C-terminal domain of the merlin proteins of various species appears to be high, while some conservation can be found within separate taxonomic groups such as vertebrates, insects, and worms (Figure 2 and Additional File 1). A part of the C-terminal region is absent in *Fugu rubripes*, *Danio rerio*, *C. briggsae*, and *Brugia malayi*. This may be due to partial assembly of the protein sequences, as all of them were predicted by bioinformatics using the available genomes and cDNA sequences. Alternatively, the lack of conservation in the C-terminal region of merlin in these species may imply that this region does not share the same function. In the remaining organisms, the C-terminal amino acid residues have a specific charge distribution, in spite of decreased hydrophilicity, when compared with the C-terminal part of moesin [15]; however, they likely form structures similar to the B, C, and D helices found in moesin.

Unlike ERM proteins, two regions (residues 1–27 and 280–323) in the N-terminal half of merlin have been mapped that are sufficient for binding to F-actin [59,60]. The first 17 amino acids in the N-terminus of human merlin are present in the merlin proteins of various species but not in any ERM proteins (see Additional File 1). The merlin proteins of higher vertebrates contain these residues, eight of which are absent in the merlin proteins of other organisms. Crystal structure analysis suggests that the structure of these extreme N-terminal residues of merlin is disordered in solution but likely becomes ordered as merlin binds to some effector targets [17]. Our sequence alignment indicates that the conservation in the extreme

N-terminus of merlin extends to the first 27 residues. The distribution of specific positively-charged residues also appears to be conserved in this N-terminal portion of the merlin proteins of lower vertebrates and insects. These results suggest that the first 27 amino acids of merlin serve as a common protein-binding motif. It is noteworthy that a similar sequence can be found in the ERM-like protein of *Ciona*; however, the N-terminal region of the *Ciona* protein contains ten positively-charged, basic amino acids, which may affect the binding to actin and/or other proteins (Additional File 1).

The internal actin-binding site, containing residues 280–323 in the N-terminal half of merlin, was found to be highly conserved among all merlin and ERM proteins analyzed, particularly the last 30 amino acid residues (Additional File 1). This region contains an extended helix at the beginning of the α-helical domain and its importance is supported by the identification of several disease-causing mutations (S315F, L316F, L316W, Q324L), which were predicted to destabilize the α-helical segment and disrupt its hydrogen bonding with subdomain A [16-18]. In addition, these residues have been shown to associate with F-actin in moesin [61,62] and to contribute to the ICAM-2-binding site in radixin [14].

Previously, LaJeunesse et al. [63] identified seven functionally important amino acid residues (^{170}YQMTPEM^{177}) in the N-terminal domain of *Drosophila* merlin, called the "Blue Box." These seven amino acids are identical between the human and *Drosophila* merlin proteins but differ from the ERM proteins. Sequence comparison revealed a more conserved pattern of the Blue Box; all seven amino acid residues of the Blue Box were found to be identical in the merlin sequences from vertebrates, fruit flies, and honeybees (Figure 3); however, several amino acid substitutions were found in the Blue Box of worms, fish, and *Ciona*. The most interesting substitutions were found in the merlin-like protein of *Caenorhabditis* from ^{174}ThrProGlu^{176} to ^{174}SerAlaAsp^{176}. It is noteworthy that the methionine residue at position 177 in the Blue Box is conserved among all merlin proteins but not in the ERM proteins. These results further corroborate the functional importance of the seven amino acids in the Blue Box [63].

According to the crystal structure of the FERM domain in human merlin, the Blue Box residues are located in helix α3B of subdomain B [18] and form a defined area that is located on the surface of the protein [17]. Intriguingly, the three-dimensional conformation of merlin's Blue Box region is similar to that of the equivalent region in radixin [18], suggesting that regions in addition to the Blue Box are required for merlin to function as a tumor suppressor. Note that regions closely adjacent to the Blue Box-equivalent residues in human ERM proteins have been shown to participate in the N-terminal to C-terminal intramolecular interaction and ligand-binding, enabling increased mobility and structural changes in the activated FERM domain [14-16,64]. In light of the functional importance of the Blue Box in *Drosophila* merlin, its sequence conservation during evolution, and its location on the surface of merlin, the Blue Box probably participates in specific protein-protein interactions and contributes to other activities of merlin.

As in ERM proteins, phosphorylation affects the subcellular localization and intra- and inter-molecular associations of merlin [13,30-32]. In addition, it modulates the ability of merlin to suppress cell growth [34,35]. Two phosphorylation sites have been mapped to the Ser^{518} and Thr^{576} residues in the merlin protein. Phosphorylation on Ser^{518} has been shown to modulate the ability of merlin to form intramolecular associations and to bind to critical effectors important for growth suppression [34,35]. In contrast, phosphorylation on the Thr^{576} residue has no effect on merlin's functional activity, while phosphorylation on this residue is important for the function of ERM proteins [57,65-67]. Sequence alignment shows that the Ser^{518} residue is conserved across all merlin proteins of various species with the exception of the fruit fly and the worm, which contain a related threonine residue at the corresponding position (Figure 4). Since both the serine and threonine residues can be phosphorylated, we suggest that the corresponding threonine residue in merlin proteins of the fly and the worm may act as a phosphorylation site.

Gutmann et al. showed that mutations within the predicted α-helical region of the human merlin protein had little effect on its function, whereas those in its N- or C-terminus rendered the protein inactive as a negative growth regulator [28,29]. Specifically, five naturally occurring missense mutations, L64P, K79E, E106G, L535P and Q538P, were found to inactivate merlin function. Interestingly, we found that the Leu^{64} and Lys^{79} residues were conserved among the merlin and ERM proteins of various species (Figure 4). According to the crystal structure of the FERM domain of merlin, the L64P substitution would create a cavity in the hydrophobic core of subdomain A and affect its β-sheet structure [17,18]. The significance of this structural information was further supported by the finding that the L64P mutation impaired the ability of merlin to form an intramolecular complex between its two N-terminal interaction sites [28]. Moreover, the L64P mutant lost its ability to bind the cytoplasmic tail of CD44; this interaction correlates with the ability of merlin to function as a growth suppressor [29].

The Lys[79] residue is situated at the end of helix α4A, and mutation at this residue (K79E) may cause the formation of a salt bridge with its neighboring Lys[76] residue, which is normally hydrogen bonded to Tyr[66] in helix α3A [17]. Two equivalent lysine residues, Lys[60] and Lys[63], were found in module F1 of moesin and were predicted to be involved in specific protein interactions, consequently changing the structure of an activated molecule [15,16]. Together with a group of positively-charged amino acids at the beginning of the helix module F3 (R275, K278, R279), these lysine residues interact with the negatively-charged residues (342-REKEE-346) in the C-terminal region [16]. Importantly, most of the homologous positively-charged residues located between lobes F1 and F3 in the radixin protein have been shown to bind to inositol 1,4,5-trisphosphate (IP3) [15]. In addition to Lys[79], the Lys[76] residue was also found to be highly conserved among various merlin and ERM proteins with the exception of the worm protein, which has a Gln[76] instead of Lys[76] (Figure 4 and Additional File 1). Also, the ERM-like proteins of parasites *Taenia saginata*, *Echinococcus granulosus*, and *Echinococcus multilocularis* contain an Arg[76] residue, which is also a basic amino acid residue and may be capable of participating in interactions similar to those of the corresponding lysine residue. On the contrary, in the JF2 protein of *Shistosoma mansoni*, the position equivalent to Lys[60] of moesin is occupied by a glutamic acid residue, and no conservation of residues 275, 278, and 279 in the JF2 protein was found, suggesting a unique structural feature for this S*histosoma* protein.

Several other naturally-occurring missense mutations on human merlin, including E106G, L535P and Q538P, have also been found to affect its functional activity [29,68]. Our sequence alignment revealed that the Glu[106], Leu[535], and Gln[538] residues were conserved among the merlin proteins of the Chordata group (Figure 4), highlighting the general importance of these residues for merlin function. Similar to the L64P mutation described above, the E106G mutation resulted in impaired intramolecular associations of merlin [29]. However, the Leu[64] residue is highly conserved among all merlin and ERM proteins of various species, while Glu[106] is conserved only in the merlin proteins of Chordata and worms. In the crystal structure of the FERM domain of merlin, the Glu[106] residue is located at the linker A-B (α1'B) and participates in the inter-subdomain interaction by forming a hydrogen bond with the Lys[289] residue [17,18]. This interaction may enable subdomain B to rotate closer to subdomain C. Intriguingly, Lys[289] is conserved only among the merlin proteins of mammals, chickens, frogs, honeybees, and mosquitoes (Additional File 1). In the merlin proteins of other species, a negatively-charged aspartic or glutamic acid occupies this position, except in fish. Instead of lysine, an arginine residue was found in the homologous position of all ERM proteins (e.g., Arg[273] for moesin), except for the ERM-like proteins in parasites (Additional File 1). This Arg[273] residue, located at the beginning of the helix of subdomain F3, lies in a specific cleft between subdomains F1 and F3 with the positively-charged R275, K278, and R279 residues. According to the structure of radixin, the IP_3-binding site is located at this basic cleft [14]. This region in the moesin protein has also been shown to interact with its C-terminal domain [16].

It should be mentioned that residues that are conserved in the merlin proteins, but not in the ERM proteins, of various species may be important for elucidating the functional difference between the merlin and ERM proteins. We found that the glutamic acid residue at position 204 of human merlin was conserved among all merlin proteins (Figure 3). In contrast, variable and uncharged amino acids were found at the corresponding position of the ERM proteins. Crystal structure of the FERM domain of human merlin shows that the Glu[204] residue lies in the beginning of helix α4B and is included in a specific stretch of amino acids in subdomain B [17]. By sequence alignment of human merlin and ERM proteins, about 70 amino acids, including this specific stretch of residues, which are unique to merlin but different in ERM proteins, were identified (see Additional File 1). The majority of these amino acids can be subdivided into three clusters; each of them is specific to its corresponding subdomain and is located on the surface of the protein. These results suggest that these 70 amino acids likely take part in protein-protein interactions. Note that the conserved stretch of amino acids in subdomain B also includes the functionally important "Blue Box" discussed above.

Similarly, the isoleucine residue at position 546 was found to be conserved among the merlin proteins of various species, while a leucine residue was present in the corresponding position in all ERM proteins (Figure 2). The residue homologous to Leu[529] in the C-terminal domain of moesin is located at the end of helix A with other hydrophobic residues, tightly contacting the hydrophobic faces of helices B and D of subdomain F2 [15]. Although the information about such an interaction in merlin is presently unavailable, additional crystal structure analysis should allow us to better understand the importance of this amino acid residue. In addition, it will be interesting to test whether mutations in the conserved amino acid residues identified in this study could affect protein function.

Predicted secondary structure of merlin and comparative analysis of the predicted α-helical region

Turunen et al. previously reported that the central region of ERM proteins contained approximately 200 residues that were predicted to be mostly α-helical [19]. To exam-

Figure 5

The predicted secondary structures for the merlin proteins of various species. The region with a predicted β-sheet structure is shaded in grey, while the region with an α-helix structure is shaded in black. These predicted secondary structures correspond to the crystal structural data [18], which are shown above the alignment with the α-helix region indicated with a thick black bar and the β-sheet region with a thin black bar. The predicted α-helical domain in the central-to-C-terminal region of merlin is marked with an open bar. Asterisks denote known domains of the merlin protein with numbers pointing to the end of truncated *Drosophila* merlin protein discussed in the text. "+" denotes the beginning and the end of the predicted α-helical domain. The positions of specific residues in the FERM domain discussed in the text are denoted by black dots below the aligned sequences.

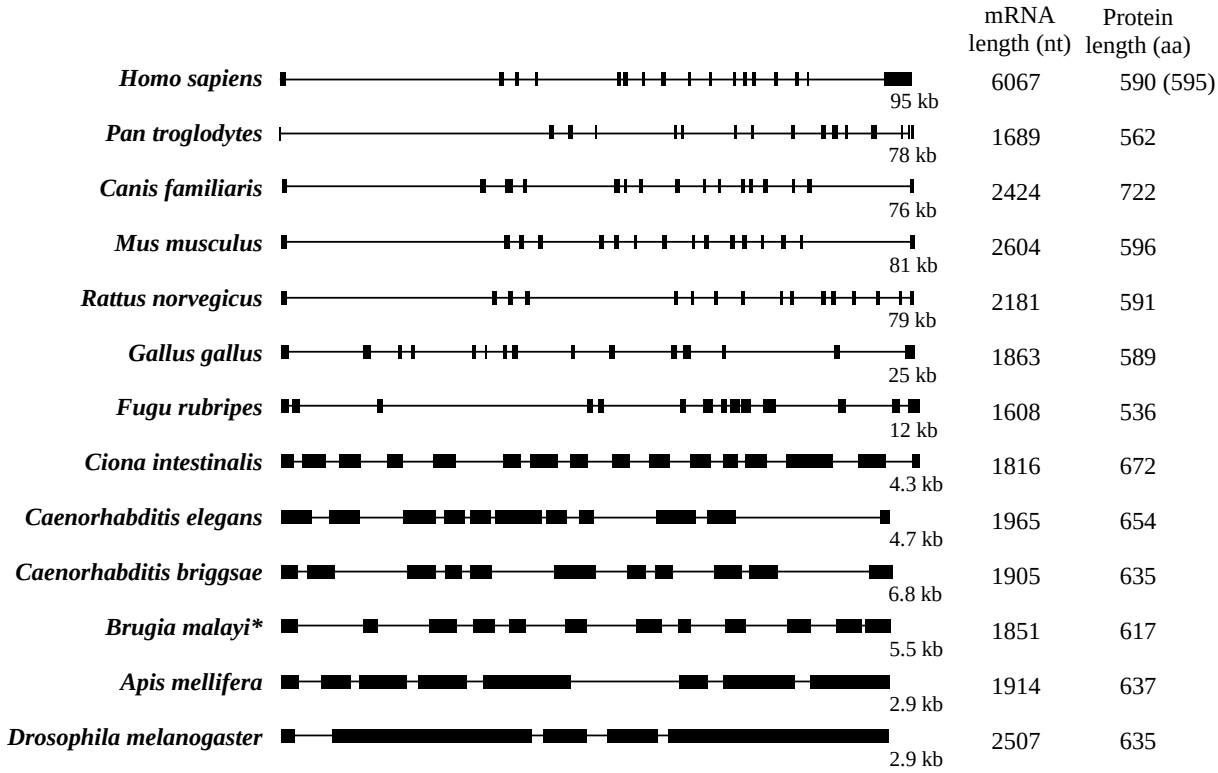


Figure 6

Schematic diagram of the exon and intron structures for the merlin genes of various species. The horizontal line depicts the merlin gene with its size indicated in base pairs (bp) on the right. The upright boxes represent exons. The lengths of the merlin mRNA sequences available in the database are shown in nucleotides (nt) and the lengths of the predicted merlin proteins are also indicated in amino acids (aa). The indicated human *NF2* mRNA refers to the longest, full-length transcript identified, which contains a long 3' untranslated region [72]. Two major *NF2* isoforms I and II are produced via alternative splicing and the lengths of their protein products are shown with that of isoform I indicated in the parentheses. It should be noted that Northern blot analysis detected the rat and mouse *NF2* mRNAs of about 4.5 kb, indicating that the sizes of the rodent *NF2* mRNAs shown here are not full-length. The asterisk (*) denotes the exon-intron structure of *Brugia malayi* predicted from this study.

ine whether there was a similar structure present in all merlin proteins, we analyzed 21 merlin sequences from various organisms and predicted their secondary structures using the JPRED program [69]. The results of such an analysis for six representative species are presented in Figure 5. The predicted locations of α-helices and β-sheets in the N-terminal domain support the experimental findings from the structural analysis of the FERM domain of human merlin protein [18]. In addition, a predicted α-helical structure in the central-to-C-terminal region was found to be conserved among the merlin proteins of various species analyzed. Previously, it was shown that a truncated merlin protein of *Drosophila*, containing residues 1–600, lost the ability to localize to the cytoplasm and was concentrated at the plasma membrane [63]. However, two smaller truncated proteins, consisting of residues 1–350 or 1–375, were only loosely associated with the plasma membrane. These results suggest that the predicted α-helical region of merlin is important for its intracellular localization. Since almost the entire α-helical domain was absent in these truncated proteins, we suggest that it may contain a determinant for membrane association. This notion is further supported by the observation that additional truncated proteins, containing residues 1–300 or missing almost the entire α-helical domain, were diffusely localized to the cytoplasm.

In human merlin, the predicted α-helical structure is situated between residues K312 and K478 (Figure 5 and Additional File 1). The N-terminal border of this structure was clearly recognized for 21 merlin sequences analyzed, whereas the C-terminal boundary could be traced only up to Urochordata (*Ciona*) and Nematoda (*Caenorhabditis*). This α-helical domain, predicted from all 21 merlin

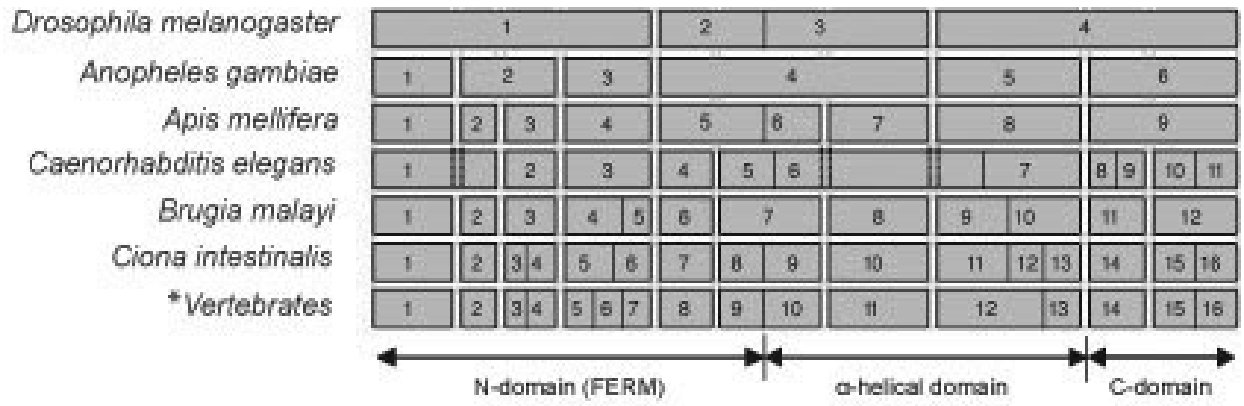


Figure 7

The alignment of exons with specific domains of merlin reveals the presence of homologous introns. Boxes represent the coding exons with numbers indicated accordingly. The locations of the three commonly discussed domains are marked with horizontal arrows under exons. The boundaries among these domains were defined according to the human merlin protein. The asterisk indicates that the exon structure shown is common to the merlin genes of all vertebrate species studied, including *Homo sapiens*, *Pan troglodytes*, *Canis familaris*, *Mus musculus*, *Rattus norvegicus*, *Danio rerio*, *Fugu rubripes*, *and Xenopus laevis*. The overall merlin gene structure of *Gallus gallus* is similar, except that exons 15 and 16 are fused together in this species.

sequences, contains a high density of charged amino acids (from about 25% in *Ciona* to greater than 40% in vertebrates). Sequence alignment reveals 19 conserved amino acid residues in this predicted α-helical domain (Additional File 1). The amino acid identity for the predicted α-helical domain within each phylogenetic group is as follows: 1) Genus *Drosophila* (*D. melanogaster* and *D. yakuba*) – 99% (one amino acid substitution), 2) Genus *Caenorhabditis* (*C. elegans*, *C. briggsae*, and *C. remanie*) – 85.7%, 3) Genus *Ciona* (*C. intestinalis* and *C. savignyi*) – 71.7%, 4) vertebrates – 63.5%, and 5) mammals – 90%. Taken together, these results indicate that the merlin proteins of various species contain a conserved α-helical domain in the central to C-terminal region.

Exon-intron structural evolution of the merlin gene

Recent progress in automated computational analysis of partially and completely sequenced genomes using a gene prediction method and the analysis of expressed sequence tag (EST) has provided considerable opportunities not only to describe novel genes but also their exon-intron structures. Such an approach also allows the examination of the presence of different splicing variants/isoforms. To understand the evolution of the exon-intron structure of the merlin gene, we assembled all available *NF2* gene-related sequences from various taxa. Using the sequences of proteins, mRNAs, and combined contigs [70], we established the structure of the merlin-like gene for *Brugia malayi*, which consists of 12 exons and 11 introns (Figure 6). Analogously, the homolog of the *NF2* gene in *Caenorhabditis elegans* was found to contain 11 exons and 10 introns. It should be mentioned that the two *NF2*-like sequences, nfm-1a and nfm-1b of *Caenorhabditis elegans*, differ from each other only by the sequence of the last exon (Additional File 1), suggesting that they represent cDNA isoforms.

As shown in Figure 6, the general arrangement of the merlin gene structure is conserved among mammalian species, especially at the region that encodes the N-terminal domain, albeit the length of the genes may differ slightly. The human *NF2* gene consists of 17 exons and spans about 95 kb of DNA [5,6,71,72]. *NF2* transcripts undergo alternative splicing, generating multiple isoforms [72-79]. Isoform I, missing exon 16, and isoform II, containing all 17 exons, are the two predominant species. As the result of alternative splicing, isoform 1 encodes a 595 amino acid protein. Isoform II differs from isoform I only at the C-terminus. Insertion of exon 16 into the mRNA provides a new stop codon, resulting in a 590 amino acid protein that is identical to isoform I over the first 579 residues. Because of the presence of a long 3' untranslated region, the longest *NF2* isoform I RNA, containing the sequence from all 17 exons, is about 6.1 kb [72]. The merlin genes of other mammalian species, R*attus norvegicus*, *Canis familiaris*, *Mus musculus* and *Pan troglodytes*, contain 16 exons (Figures 6 and 7). In addition, alternatively spliced merlin isoforms have been found in rodent species [80-82]. On the contrary, the structure of the merlin genes of *Gallus gallus* and *Fugu rubripes* are arranged differently from those of mammalian species, with 15 and 14 exons spanning much shorter DNAs of only about 25 kb and 12.3 kb, respectively (Figure 6). Although the *NF2* gene of *Fugu rubripes* has not yet been completely assembled, we

predict that it lacks the sequences of the first and the last exons of the mammalian *NF2* gene based on our sequence alignment (see Additional File 1).

In spite of the presence of 16 exons and the size of transcript similar to those found in some vertebrates, the gene for the merlin-like protein of *Ciona intestinalis* is relatively small with only about 4.3 kb (Figure 6). This tendency towards reduction of intron length and number continues to be seen in the merlin gene of worms and insects (Figures 6 and 7). The gene for the merlin-like protein of *Caenorhabditis elegans*, consisting of 11 exons, spans about 4.7-kb DNA, and that of *Brugia malayi*, containing 12 exons, is about 5.5 kb in length. The *NF2* homolog of *Drosophila melanogaster* and the gene for the merlin-like protein of *Apis mellifera* are only about 2.9 kb, the smallest in the merlin clade, and consist of 5 and 8 exons, respectively (Figure 6). Interestingly, some conservation of the positions of homologous introns can be found in the *NF2* gene from various species (Figure 7); however, the sizes of their introns are variable. Such an evolutionary conservation of homologous introns implies that the presence of regulatory sequences in the introns to regulate the transcriptional event.

Unlike the sizes and structures of the merlin or merlin-like genes in various organisms, the lengths of the merlin proteins and transcripts have not changed very much during evolution (Figure 6). Moreover, several functionally important domains of merlin also remain conserved. Since the merlin protein of the insect emerged after deviation from that of the worm, which was anciently derived from the common ancestor (Figure 1), it appears that the decrease in gene size and exon number occurred specifically within the insect group. This branch of merlin evolution is likely to develop independently and in the opposite direction from those more recently developed merlin proteins of Chordata. Parallel evolution towards the increase in merlin gene size and exon number between the worm and Chordata appears to be less likely. Thus, it is possible that the common ancestor for the merlin genes of the worm, the insect, and Chordata may contain a few more exons. During evolution, reduction of introns and/or fusion of exons occur within the insect group.

It is evident that the genome of the insect is more complicated than that of the worm. The simplification of the merlin gene structure in the insect is unique and may have a functional significance. This observation may explain the lack of splicing variants of the *NF2* homolog in *Drosophila*, in contrast to those merlin isoforms found in mammals [72-82] and in *Caenorhabditis elegans* as we predicted in this study.

Conclusion

We have conducted the phylogenetic analysis of merlin diversity across metazoan genomes using the experimentally annotated and predicted sequences in conjunction with bioinformatic tools. We show that the merlin proteins have a monophyletic origin with a root in the early metazoa. We have also established the expansion of the ERM-like ancestors within the vertebrate clade that occurred after its separation from Urochordata. Several potentially important sites that are conserved among all merlin proteins but are divergent in the ERM members have been identified. As supported by the crystal structure data, these conserved residues are likely to be important for merlin function. Analysis of the evolution of the merlin gene structure reveals the existence of common splicing variants in human and *Caenorhabditis elegans*. While a trend toward the increase of gene length during evolution was observed, a reduction of intron number and length appears to occur in the merlin gene of the insect group. Taken together, our results have important implications for the evolution of the merlin proteins and their possible functional variability in different taxa.

Methods

BLAST search

Initial sequences of genes and proteins of interest from various organisms were identified from the National Center for Biotechnology Information (NCBI) database [83] using the BLAST algorithm [84]. We then searched the desirable sequences across genomic databases of completely or partially sequenced genomes available at The Sanger Institute [85] and The Institute for Genomic Research (TIGR; [86]). Also, we investigated other available sequence databases that contain information for specific organisms. The sources of sequences of the predicted or experimentally annotated merlin and ERM proteins are summarized in Table 1.

To obtain the entire amino acid sequence of an annotated protein, we used UniProt from Universal Protein Resource [87]. The erythrocyte membrane proteins 4.1 sequences of *Homo sapiens* (GenBank: CAI21970), *Mus musculus* (GenBank: NP_001006665), and *Danio rerio* (GenBank: AAQ97985) were also included in the analysis as an outgroup. Because of the presence of many non-conserved and large introns in the genes of interest, we conducted BLAST search using TBLASTN alignment algorithm in the cases where no protein sequences were available.

Alignments and phylogeny

The CLUSTAL_X program [88] was used to align the characterized or predicted protein sequences from different species. Phylogenetic analysis was carried out using the MEGA2.1 program [49].

Secondary structure prediction

The secondary structure prediction program JPRED [69] was used to predict the secondary structure of each merlin protein from various species. This program uses physical-chemical properties of the amino acid sequence and neural network for recognition of α-helices and β-sheets.

List of Abbreviations

NF2 – Neurofibromatosis type 2

NF2 – the *Neurofibromatosis type 2* gene

ERM – ezrin, radixin, and moesin

FERM – 4.1, ezrin, radixin, and moesin

IP3 – inositol 1,4,5-trisphosphate

TIGR – The Institute for Genomic Research

EST, expressed sequence tag

NCBI – National Center for Biotechnology Information

Authors' contributions

KG and AB carried out the phylogenetic analysis of merlin diversity across metazoan genomes and drafted the manuscript. EMA and LVO helped with the design of the study and preparation of data for the figures. LSC was the principal investigator of the project and participated in the design, coordination, and writing of the manuscript. All authors read and approved the final manuscript

Additional material

Additional File 1

Complete amino acid sequence alignment of the merlin and ERM proteins. *Letters shaded in grey illustrate the conservation pattern of aligned sequences. The names of the merlin proteins from various species are shown in yellow. The conserved residues of the 'Blue Box' are also shaded in yellow. The positions of residues discussed in the text are colored in red. Blue letters denote the conserved residues within the predicted α-helical domain. Numbers indicated at the end of each sequence refer to the positions of the last residue within each protein analyzed.*

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2148-5-69-S1.pdf]

Acknowledgements

We sincerely thank Sarah S. Burns, Maria Mihaylova, and D. Bradley Welling for critical reading of the manuscript, and Hui-Chun Hsiao and Kiselev Arkadyi for the help with the diagrams. This study was supported by grants from the US Department of Defense Neurofibromatosis Research Program and Russian Fund of Fundamental Investigations.

References

1. Dacks JB, Doolittle WF: **Novel syntaxin gene sequences from *Giardia*, *Trypanosoma* and algae: implications for the ancient evolution of the eukaryotic endomembrane system.** *J Cell Sci* 2002, **115:**1635-1642.
2. Hsu S: **Bioinformatics in reproductive biology – functional annotation based on comparative sequence analysis.** *J Rep Immunol* 2004, **63:**75-83.
3. NIH Consensus Statement on Acoustic Neuroma: *Neurofibromatosis Res. Newsletter* 1992, **8:**1-7.
4. Bull. World Health Org: **Prevention and control of neurofibromatosis: Memorandum from a joint WHO/NNFF meeting.** 1992, **70:**173-182.
5. Rouleau GA, Merel P, Luchtman M, Sanson M, Zucman J, Marineau C, Hoang-Xuan K, Demczuk S, Desmaze C, Plougastel B, Pulst SM, Le noir G, Bijlsma E, Fashold R, Dumanski J, de Jong P, Parry D, Eldridge R, Aurias A, Delattre O, Thomas G: **Alteration in a new gene encoding a putative membrane-organizing protein causes neuro-fibromatosis type 2.** *Nature* 1993, **363:**515-521.
6. Trofatter JA, MacCollin MM, Rutter JL, Murrell JR, Duyai MP, Parry DM, Eldridge RE, Kley N, Menon AG, Pulaski K, Haase VH, Ambrose CM, Munroe D, Bove C, Haines JL, Martuza RL, MacDonald ME, Seizinger BR, Short MP, Buckner AJ, Gusella JF: **A novel moezin-, ezrin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor.** *Cell* 1993, **72:**791-800.
7. Algrain M, Arpin M, Louvard D: **Wizardry at the cell cortex.** *Current Biol* 1993, **3:**451-454.
8. Tsukita S, Oishi K, Sato N, Sagara J, Kawai A, Tsukita S: **ERM family members as molecular linkers between the cell surface glycoprotein CD44 and actin-based cytoskeleton.** *J Cell Biol* 1994, **126:**391-401.
9. Takeuchi K, Kawashima A, Nagafuchi A, Tsukita S: **Structural diversity of band 4.1 superfamily members.** *J Cell Sci* 1994, **107:**1921-1928.
10. Chishti AH, Kim AC, Marfatia SM, Lutchman M, Hanspal M, Jindal H, Liu SC, Low PS, Rouleau GA, Mohandas N, Chasis JA, Conboy JG, Gaskard P, Takakuwa Y, Huang SC, Benz EJ Jr, Bretcher A, Fenon RG, Gusella JF, Ramesh V, Solomon F, Marchesi VT, Tsukita S, Hoover KB: **The FERM domain: a unique module involved in the linkage of the cytoplasmic proteins to the membrane.** *Trends Biochem Sci* 1998, **23:**281-282.
11. Hamada K, Shimizu T, Matsui T, Tsukita S, Hakoshima T: **Structural basis of the membrane-targeting and unmasking mechanisms of the radixin FERM domain.** *EMBO J* 2000, **19:**4449-4462.
12. Herrlich P, Morrison H, Sleeman J, Orian-Rousseau V, Konig H, Weg-Remers S, Ponta H: **CD44 acts both as a growth- and invasiveness-promoting molecule and as a tumor-suppressing cofactor.** *Ann NY Acad Sci* 2000, **910:**106-118.
13. Morrison H, Sherman LS, Legg J, Banine F, Isacke C, Haipek CA, Gutmann DH, Ponta H, Herrlich P: **The NF2 tumor suppressor gene product, merlin, mediates contact inhibition of growth through interactions with CD44.** *Genes Dev* 2001, **15:**968-980.
14. Hamada K, Shimizu T, Yonemura S, Tsukita Sh, Tsukita S, Hakoshima T: **Structural basis of adhesion – molecule recognition by ERM proteins revealed by the crystal structure of the radixin – ICAM-2 complex.** *EMBO J* 2003, **22:**502-514.
15. Pearson MA, Reczek D, Bretscher A, Karplus PA: **Structure of the ERM protein moesin reveals the FERM domain fold masked by an extended actin binding tail domain.** *Cell* 2000, **101:**259-270.
16. Edwards SD, Keep NH: **The 2.7 Å crystal structure of the activated FERM domain of moesin: an analysis of structural changes on activation.** *Biochemistry* 2001, **40:**7061-7068.
17. Kang BS, Cooper DR, Devedjiev Y, Derewenda U, Derewenda ZS: **The structure of the FERM domain of merlin, the neurofibromatosis type 2 gene product.** *Acta Crystollogr Sec D* 2002, **58:**381-391.
18. Shimizu T, Seto A, Maita N, Hamada K, Tsukita Sh, Tsukita S, Hakoshima T: **Structural basis for neurofibromatosis type2.** *J Biol Chem* 2002, **277:**10332-10336.
19. Turunen O, Sainio M, Jaaskelainen J, Carpen O, Vaheri A: **Structure-function relationships in the ezrin family and the effect of tumor-associated point mutations in neurofibromatosis 2 protein.** *Biochim Biophys Acta* 1998, **1387:**1-16.

20. Turunen O, Wahlström T, Vaheri A: **Ezrin has a COOH-terminal acting-binding site that is conserved in the ezrin protein family.** *J Cell Biol* 1994, **126:**1445-1453.
21. Gary R, Bretscher A: **Ezrin self-association involves binding of an N-terminal domain to a normally masked C-terminal domain that includes the F-actin binding site.** *Mol Biol Cell* 1995, **6:**1061-1075.
22. Sainio M, Zhao F, Heiska L, Turunen O, den Bakker M, Zwarthoff E, Lutchman M, Rouleau GA, Jaaskelainen J, Vaheri A, Carpen O: **Neurofibromatosis 2 tumor suppressor protein colocalizes with ezrin and CD44 and associates with actin-containing cytoskeleton.** *J Cell Sci* 1997, **110:**2249-2260.
23. Scoles DR, Huynh DP, Morcos PA, Coulsell ER, Robinson NG, Tamanoi F, Pulst SM: **Neurofibromatosis 2 tumor suppressor schwannomin interacts with βII-spectrin.** *Nature Genet* 1998, **18:**354-359.
24. Neill GW, Crompton MR: **Binding of the merlin-I product of the neurofibromatosis type 2 tumour suppressor gene to a novel site in beta-fodrin is regulated by association between merlin domains.** *Biochem J* 2001, **358:**727-735.
25. Bretscher A, Edwards K, Fehon RG: **ERM proteins and merlin: integrators at the cell cortex.** *Nature Rev Mol Cell Biol* 2002, **3:**586-599.
26. Ramesh V: **Merlin and ERM proteins in Schwann cells, neurons and growth cones.** *Nature Rev Neurosci* 2004, **5:**462-470.
27. McClatchey AI: **Merlin and ERM proteins: unappreciated roles in cancer development?** *Nature Rev Cancer* 2003, **3:**877-883.
28. Gutmann DH, Haipek CA, Hoang Lu K: **Neurofibromatosis 2 tumor suppressor protein, merlin, forms two functionally important intramolecular associations.** *J Neurosci Res* 1999, **58:**706-716.
29. Gutmann DH, Hirbe AC, Haipek CA: **Functional analysis of neurofibromatosis 2 (NF2) missense mutations.** *Hum Mol Genet* 2001, **10:**1519-1529.
30. Shaw RJ, Paez JG, Curto M, Yaktine A, Pruitt WM, Saotome I, O'Bryan JP, Gupta V, Ratner N, Der CJ, Jacks T, McClatchey AI: **The Nf2 tumor suppressor, merlin, functions in Rac-dependent signaling.** *Dev Cell* 2001, **1:**63-72.
31. Kissil JL, Johnson KC, Eckman MS, Jacks T: **Merlin phosphorylation by p21-activated kinase 2 and effects of phosphorylation on merlin localization.** *J Biol Chem* 2002, **277:**10394-10399.
32. Xiao GH, Beeser A, Chernoff J, Testa JR: **p21-activated kinase links Rac/Cdc42 signaling to merlin.** *J Biol Chem* 2002, **277:**883-886.
33. Alfthan K, Heiska L, Gronholm M, Renkema GH, Carpen O: **Cyclic AMP-dependent protein kinase phosphorylates merlin at serine 518 independently of p21-activated kinase and promotes merlin-ezrin heterodimerization.** *J Biol Chem* 2004, **279:**18559-18566.
34. Surace EI, Haipek CA, Gutmann DH: **Effect of merlin phosphorylation on neurofibromatosis 2 (NF2) gene function.** *Oncogene* 2004, **23:**580-587.
35. Rong R, Surace EI, Haipek CA, Gutmann DH, Ye K: **Serine 518 phosphorylation modulates merlin intramolecular association and binding to critical effectors important for NF2 growth suppression.** *Oncogene* 2004, **23:**8447-8454.
36. Frosch PM, Frosch M, Pfister T, Schaad V, Bitter-Suermann D: **Cloning and characterisation of an immunodominant major surface antigen of *Echinococcus multilocularis*.** *Mol Biochem Parasitol* 1991, **48:**121-130.
37. Frosch PM, Muhlschlegel F, Sygulla L, Hartmann M, Frosch M: **Identification of a cDNA clone from the larval stage of *Echinococcus granulosus* with homologies to the *E. multilocularis* antigen EM10-expressing cDNA clone.** *Parasitol Res* 1994, **80:**703-705.
38. Kurtis JD, Ramirez BL, Wiest PM, Dong KL, El-Meanawy A, Petzke MM, Johnson JH, Edmison J, Maier RA Jr, Olds GR: **Identification and molecular cloning of a 67-kilodalton protein in *Schistosoma japonicum* homologous to a family of actin-binding proteins.** *Infect Immun* 1997, **65:**344-347.
39. Hubert K, Cordero E, Frosch M, Solomon F: **Activities of the EM10 protein from *Echinococcus multilocularis* in cultured mammalian cells demonstrate functional relationships to ERM family members.** *Cell Motil Cytoskeleton* 1999, **42:**178-188.
40. McCartney BM, Fehon RG: **Distinct cellular and subcellular patterns of expression imply distinct functions for the *Drosophila* homologues of moesin and the neurofibromatosis 2 tumor suppressor, merlin.** *J Cell Biol* 1996, **133:**843-852.
41. Hansson CM, Ali H, Bruder CE, Fransson I, Kluge S, Andersson B, Roe BA, Menzel U, Dumanski JP: **Strong conservation of the human NF2 locus based on sequence comparison in five species.** *Mamm Genome* 2003, **14:**526-536.
42. Chen Y-X, Gutmann DH, Haipek CA, Martinsen BJ, Bronner-Fraser M, Krull CE: **Characterization of chicken Nf2/merlin indicates regulatory roles in cell proliferation and migration.** *Dev Dyn* 2004, **229:**541-54.
43. [http://www.tigr.org/tigr-scripts/tgi/est_report.pl?GB=CD797075&species=r_appendiculatus].
44. [http://www.tigr.org/tigr-scripts/tgi/est_report.pl?GB=BM291669&species=a_variegatum].
45. [http://www.tigr.org/tigr-scripts/tgi/est_report.pl?GB=CK190110&species=B.microplus].
46. [http://genome.wustl.edu/blast/client.pl].
47. Saitou N, Nei M: **The neighbor-joining method: A new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987, **4:**406-425.
48. Kumar S, Gadagkar SR: **Efficiency of the neighbor-joining method in reconstructing deep and shallow evolutionary relationships in large phylogenies.** *J Mol Evol* 2000, **51:**544-53.
49. Kumar S, Tamura K, Jakobsen IB, Nei M: **MEGA2: molecular evolutionary genetics analysis software.** *Bioinformatics* 2001, **17:**1244-1245.
50. [http://www.tigr.org/tdb/e2k1/sma1/].
51. [http://www.wormbase.org/db/gene/gene?name=F42A10.2a;class=Transcript].
52. The C. elegans Sequencing Consortium: **Genome sequence of the nematode C. elegans: a platform for investigating biology.** *Science* 1998, **282:**2012-2018.
53. Adams MD, Celniker SE, Holt RA, Evans CA, *et al.*: **The genome sequence of Drosophila melanogaster.** *Science* 2000, **287:**2185-2195.
54. Dehal P, Satou Y, Campbell RK, Chapman J, *et al.*: **The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins.** *Science* 2002, **298:**2157-2167.
55. Stein LD, Bao Z, Blasiar D, Blumenthal T, *et al.*: **The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics.** *PLoS Biol* 2003, **1:**E45.
56. Foster JM, Kumar S, Ganatra MB, Kamal IH, Ware J, Ingram J, Pope-Chappell J, Guiliano D, Whitton C, Daub J, Blaxter ML, Slatko BE: **Construction of bacterial artificial chromosome libraries from the parasitic nematode *Brugia malayi* and physical mapping of the genome of its Wolbachia endosymbiont.** *Int J Parasitol* 2004, **34:**733-746.
57. Matsui T, Maeda M, Doi Y, Yonemura S, Amano M, Kaibuchi K, Tsukita S: **Rho-kinase phosphorylates COOH-terminal threonines of ezrin/radixin/moezin (ERM) proteins and regulates their head-to-tail association.** *J Cell Biol* 1998, **140:**647-657.
58. Deguen B, Merel P, Goutebroze L, Giovannini M, Reggio H, Arpin M, Thomas G: **Impaired interaction of naturally occurring mutant NF2 protein with actin-based cytoskeleton and membrane.** *Hum Mol Genet* 1998, **7:**217-226.
59. Xu H, Gutmann DH: **Merlin differentially associates with the microtubule and acting cytoskeleton.** *J Neurosci Res* 1998, **51:**403-415.
60. Brault E, Gautreau A, Lamarine M, Callebaut I, Gilles T: **Normal membrane localization and actin association of the NF2 tumor suppressor protein are dependent on folding of its N-terminal domain.** *J Cell Sci* 2001, **114:**1901-1912.
61. Roy C, Martin M, Mangeat P: **A dual involvement of the amino terminal domain of ezrin in F- and G-actin binding.** *J Biol Chem* 1997, **272:**20088-20095.
62. Martin M, Roy C, Montcourrier P, Sahuquet A, Mangeat P: **Three determinants in ezrin are responsible for cell extension activity.** *Mol Cell Biol* 1997, **8:**1543-1557.
63. LaJeunesse DR, McCartney BM, Fehon RG: **Structural analysis of *Drosophila* merlin reveals functional domains important for growth control and subcellular localization.** *J Cell Biol* 1998, **141:**1589-1599.
64. Smith WJ, Nassar N, Bretscher A, Cerione RA, Karplus PA: **Structure of the active N-terminal domain of ezrin.** *J Biol Chem* 2003, **278:**4949-4956.

65. Nakamura F, Amieva MR, Furthmayr H: **Phosphorylation of threonine 558 in the carboxyl-terminal actin-binding domain of moesin by thrombin activation of human platelets.** *J Biol Chem* 1995, **270:**31377-31385.
66. Oshiro N, Fukata Y, Kaibuchi K: **Phosphorylation of moesin by rho-associated kinase (Rho-kinase) plays a crucial role in the formation of microvilli-like structures.** *J Biol Chem* 1998, **273:**34663-34666.
67. Dard N, Louvet-Vallee S, Santa-Maria A, Maro B: **Phosphorylation of ezrin on threonine T567 plays a crucial role during compaction in the mouse early embryo.** *Dev Biol* 2004, **271:**87-97.
68. Stokowski RP, Cox DR: **Functional analysis of the neurofibromatosis type 2 protein by means of disease – causing point mutations.** *Am J Hum Genet* 2000, **66:**873-891.
69. [http://www.compbio.dundee.ac.uk/~www-jpred/].
70. [http://www.tigr.org/tdb/e2k1/bma1/].
71. Zucman-Rossi J, Legoix P, Der Sarkissian H, Cheret G, Sor F, Bernardi A, Cazes L, Giraud S, Ollagnon E, Lenoir G, Thomas G: **NF2 gene in neurofibromatosis type 2 patients.** *Hum Mol Genet* 1998, **7:**2095-2101.
72. Chang L-S, Akhmametyeva EM, Wu Y, Zhu L, Welling DB: **Multiple transcription initiation sites, alternative splicing, and differential polyadenylation contribute to the complexity of human neurofibromatosis 2 transcripts.** *Genomics* 2002, **79:**63-76.
73. Arakawa H, Hayashi N, Nagase H, Ogawa M, Nakamura Y: **Alternative splicing of the NF2 gene and its mutation analysis of breast and colorectal cancers.** *Hum Mol Genet* 1994, **3:**565-568.
74. Bianchi AB, Hara T, Ramesh V, Gao J, *et al.*: **Mutations in transcript isoforms of the neurofibromatosis 2 gene in multiple human tumour types.** *Nature Genet* 1994, **6:**185-192.
75. Hara T, Bianchi AB, Seizinger BR, Kley N: **Molecular cloning and characterization of alternatively spliced transcripts of the mouse neurofibromatosis 2 gene.** *Cancer Res* 1994, **54:**330-335.
76. Hitotsumatsu T, Kitamoto T, Iwaki T, Fukui M, Tateishi J: **An exon 8-spliced out transcript of neurofibromatosis 2 gene is constitutively expressed in various human tissues.** *J Biochem (Tokyo)* 1994, **116:**1205-1207.
77. Pykett MJ, Murphy M, Harnish PR, George DL: **The neurofibromatosis 2 (*NF2*) tumor suppressor gene encodes multiple alternatively spliced transcripts.** *Hum Mol Genet* 1994, **3:**559-564.
78. Nishi T, *et al.*: **Neurofibromatosis 2 gene has novel alternative splicing which control intracellular protein binding.** *Int J Oncol* 1997, **10:**1025-1029.
79. Schmucker B, Tang Y, Kressel M: **Novel alternatively spliced isoforms of the neurofibromatosis type 2 tumor suppressor are targeted to the nucleus and cytoplasmic granules.** *Hum Mol Genet* 1999, **8:**1561-1570.
80. Haase VH, Trofatter JA, MacCollin M, Tarttelin E, Gusella JF, Ramesh V: **The murine NF2 homologue encodes a highly conserved merlin protein with alternative forms.** *Hum Mol Genet* 1994, **3:**407-411.
81. Huynh DP, Nechiporuk T, Pulst SM: **Alternative transcripts in the mouse neurofibromatosis type 2 (NF2) gene are conserved and code for schwannomins with distinct C-terminal domains.** *Hum Mol Genet* 1994, **3:**1075-1079.
82. Gutmann DH, Wright DE, Geist RT, Snider WD: **Expression of the neurofibromatosis 2 (NF2) gene isoforms during rat embryonic development.** *Hum Mol Genet* 1995, **4:**471-478.
83. [http://www.ncbi.nlm.nih.gov/BLAST].
84. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25:**3389-3402.
85. [http://www.sanger.ac.uk/DataSearch].
86. [http://tigrblast.tigr.org/tgi/].
87. [http://www.ebi.uniprot.org/index.shtml].
88. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, **15:**4876-4882.