Network of coregulated spliceosome components revealed by zebrafish mutant in recycling factor p110

Nikolaus S. Trede*, Jan Medenbach†, Andrey Damianov†, Lee-Hsueh Hung†, Gerhard J. Weber‡, Barry H. Paw§, Yi Zhou‡, Candace Hersey‡, Agustin Zapata¶, Matthew Keefe‡, Bruce A. Barut‡, Andrew B. Stuart‖, Tammisty Katz*, Chris T. Amemiya‖, Leonard I. Zon‡**, and Albrecht Bindereif†,††

†Institute of Biochemistry, Justus-Liebig-University, D-35392 Giessen, Germany; *Department of Pediatrics, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112; ‡Howard Hughes Medical Institute, Division of Hematology/Oncology, Children's Hospital Boston, Harvard Medical School, Boston, MA 02115; §Division of Hematology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; ¶Department of Cell Biology, Faculty of Biology, Complutense University, 28040 Madrid, Spain; and ‖Benaroya Research Institute at Virginia Mason, Department of Biology, University of Washington, Seattle, WA 98101

Communicated by Christine Guthrie, University of California, San Francisco, CA, March 2, 2007 (received for review December 20, 2006)

The spliceosome cycle consists of assembly, catalysis, and recycling phases. Recycling of postspliceosomal U4 and U6 small nuclear ribonucleoproteins (snRNPs) requires p110/SART3, a general splicing factor. In this article, we report that the zebrafish *earl grey* (*egy*) mutation maps in the *p110* gene and results in a phenotype characterized by thymus hypoplasia, other organ-specific defects, and death by 7 to 8 days postfertilization. U4/U6 snRNPs were disrupted in *egy* mutant embryos, demonstrating the importance of p110 for U4/U6 snRNP recycling *in vivo*. Surprisingly, expression profiling of the *egy* mutant revealed an extensive network of coordinately up-regulated components of the spliceosome cycle, providing a mechanism compensating for the recycling defect. Together, our data demonstrate that a mutation in a general splicing factor can lead to distinct defects in organ development and cause disease.

small nuclear RNA | small nuclear ribonucleoprotein | splicing | genetic screen | thymus

Messenger RNA splicing requires the ordered assembly of the spliceosome from >100 protein components and five small nuclear RNAs (snRNAs): U1, U2, U4, U5, and U6 (reviewed in refs. 1–3). After splicing catalysis and mRNA release, the spliceosome disassembles, and its components undergo a recycling phase, which still is poorly understood. In humans, recycling of postspliceosomal U4 and U6 small nuclear ribonucleoproteins (snRNPs) to functional U4/U6 snRNPs requires *in vitro* p110/SART3, a general splicing factor referred to as p110 in the present article (4, 5). In addition, p110 functions in recycling of the U4atac/U6atac snRNP (6). Characteristically, p110 associates only transiently with the U6 and U4/U6 snRNPs but is absent from the U4/U6.U5 tri-snRNP and spliceosomes.

The domain structure of the human p110 protein is composed of at least seven tetratricopeptide repeats (TPR) in the N-terminal half, followed by two RNA recognition motifs (RRMs) in the C-terminal half, as well as a stretch of 10 highly conserved amino acids at the C terminus (C10 domain). The N-terminal TPR domain functions in interaction with the U4/U6 snRNP-specific 90K protein, the RRMs are important for U6 snRNA binding, and the conserved C10 domain is critical for interacting with the U6-specific LSm proteins (5, 7, 8). Thus, multiple contacts mediate the interaction between p110 and the U4 and U6 components.

This p110 domain organization is conserved in many other eukaryotes, including *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Schizosaccharomyces pombe*, and *Drosophila melanogaster* (5). The *Saccharomyces cerevisiae* Prp24 protein, although functionally related to human p110, is an exception in that it lacks the entire N-terminal half with the TPR domain (9).

Here we use the zebrafish system to study the system-wide role and *in vivo* function of p110. We describe the phenotype of a zebrafish mutant, called *earl grey* (*egy*), that originated from a genetic screen for mutants of T cell and thymus development. Surprisingly, the embryonically lethal mutation was mapped in the p110 gene. Biochemical characterization of *egy* mutant embryos demonstrated the role of p110 in U4/U6 snRNP recycling *in vivo*. Through microarray expression profiling of the *egy* mutant, we discovered an extensive network of coregulated components of the spliceosome cycle, which would provide a mechanism compensating for the recycling defect. In sum, these data illustrate the usefulness of zebrafish as a vertebrate model system to investigate the role of splicing factors in organ development and human disease.

Results and Discussion

***egy* Phenotype and Locus.** Here we report that the zebrafish *egy* mutation maps in the *p110* gene. *egy* was identified in a genetic screen for mutants of T cell and thymus development by using *N*-ethyl-*N*-nitrosourea as a mutagen. Phenotypically, the *egy* mutant is characterized by microcephaly, microphthalmia (Fig. 1*A*), and death by 7 to 8 days postfertilization (dpf). Identification of the *egy* phenotype was based on the absence of T cells in the bilateral thymic organ [by using *rag1* whole-mount *in situ* hybridization (WISH); Fig. 1*B*]. Disruption in a number of pathways can lead to this phenotype (reviewed in refs. 10 and 11). Disturbance of other hematopoietic lineages was excluded with the demonstration of normal erythropoiesis in *egy* mutants (Fig. 1*C*). Similarly, neural crest cells and their migration to the pharyngeal area (Fig. 1*D*), as well as patterning of the endoderm (Fig. 1*E*), were intact. However, pharyngeal arch formation was defective (Fig. 1*F*), resulting in a thymus devoid of lymphocytes (Fig. 1*G*). Characterization of the effect of the *egy* mutation on other organs revealed that, surprisingly, although insulin expression indicated normal development of the endocrine pancreas in

Author contributions: N.S.T., J.M., and A.D. contributed equally to this work; N.S.T., J.M., A.D., L.-H.H., C.T.A., L.I.Z., and A.B. designed research; N.S.T., J.M., A.D., L.-H.H., G.J.W., B.H.P., Y.Z., C.H., A.Z., M.K., B.A.B., A.B.S., T.K., and C.T.A. performed research; N.S.T., J.M., A.D., L.-H.H., G.J.W., B.H.P., A.Z., A.B.S., C.T.A., L.I.Z., and A.B. analyzed data; and N.S.T., L.I.Z., and A.B. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: snRNP, small nuclear ribonucleoprotein; dpf, days postfertilization; snRNA, small nuclear RNA; TPR, tetratricopeptide repeats; RRM, RNA recognition motif; WISH, whole-mount *in situ* hybridization; m_3G cap, trimethyl cap structure; NCBI, National Center for Biotechnology Information.

**To whom correspondence may be addressed at: Howard Hughes Medical Institute, Department of Hematology/Oncology, Children's Hospital, Harvard Medical School, Karp Family Research Laboratories, 300 Longwood Avenue, Boston, MA 02115. E-mail: zon@enders.tch.harvard.edu.

††To whom correspondence may be addressed at: Institute of Biochemistry, Justus-Liebig-University, Heinrich-Buff-Ring 58, D-35392 Giessen, Germany. E-mail: albrecht.bindereif@chemie.bio.uni-giessen.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0701919104/DC1.

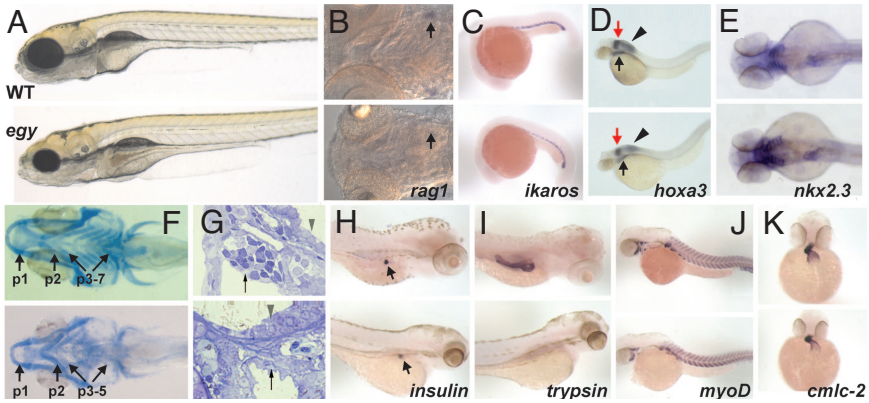

Fig. 1. Phenotype of *egy* mutant zebrafish: organ-selective defects. (*Upper*) Wild-type animals. (*Lower*) Mutant animals. (*A*) Wild-type and *egy* mutant larvae exhibiting microcephaly and microphthalmia at 5 dpf. (*B*) Absence of T cells shown by WISH with a *rag1*-specific probe. Ventral view shows *rag1* expression in 5-dpf wild-type animal, with the arrow indicating the left thymic region. No *rag1* signal was detected in *egy* mutants. (*C*) Normal primitive hematopoiesis in mutants is shown by WISH with an *ikaros* probe at 19 h postfertilization. No difference was detected between wild-type and mutant embryos in the hematopoietic intermediate cell mass. (*D*) *Hoxa3* expression at 2 dpf is equivalent between wild-type and mutant animals, indicating normal neural crest cell development. Red arrows point to rhombomeres 5 and 6, the black arrowhead shows hindbrain expression, and the black arrow indicates pharyngeal area. (*E*) *Nkx2.3* expression by WISH is normal in *egy* mutants, indicating normal patterning of endoderm. (*F*) Alcian blue staining of wild-type and mutant larvae at 7 dpf indicates a defect in pharyngeal arch formation. Gill arches are marked and numbered from anterior to posterior. (*G*) Thin section of thymus (arrow) from wild-type and mutant larvae at 7 dpf shows only a small rim of epithelium devoid of T cells in *egy* mutants. The arrow points to the thymus, and the arrowhead points to the otic vesicle epithelium. (*H*) No difference in insulin expression was detected between wild-type and mutant larvae at 5 dpf. The arrow points to the pancreatic endocrine β-islet cells. For better visualization of the pancreas, larvae are shown with anterior to the right. (*I*) Trypsin, a marker of pancreatic exocrine glands, was strongly expressed in wild-type but not in *egy* mutant animals at 5 dpf. (*J*) *MyoD* expression suggests that formation of somites and skeletal muscle were not affected in *egy* mutants at 2 dpf. (*K*) Heart looping to right and expression of cardiac myosin light chain-2 (*cmlc-2*) was equivalent in wild-type and *egy* mutant animals at 2 dpf.

egy mutants (Fig. 1*H*), markers for exocrine pancreas (represented by trypsin in Fig. 1*I*) were strongly reduced or absent. Besides normal endocrine pancreatic development, somites were normally formed (Fig. 1*J*), and heart looping and cardiac muscle development (Fig. 1*K*) were not affected in *egy* mutants. In summary, the *egy* phenotype is characterized by organ-selective defects and not by a global defect in tissue development.

To identify the mutation leading to the *egy* phenotype, a positional cloning approach was undertaken [see supporting information (SI) Fig. 5*A*]. The *egy* locus mapped to zebrafish LG5, and the critical interval was defined by recombinants on BACzC191D15, located in a region syntenic to human chromosome 12q24. This BAC contained four genes, *HYPE*, *ISCU*, *decorin*, and the zebrafish *p110* ortholog. No gross abnormalities were detected in candidate cDNA sequences of *HYPE*, *ISCU*, and *decorin* from *egy* mutants (data not shown), whereas full-length *p110* cDNA could not be amplified by RT-PCR (SI Fig. 5 *B* and *C* and data not shown). Sequencing of the *p110* gene from mutant-derived BAC clones revealed a large insertion in *p110* intron 15 (Fig. 2*A*). This large genomic aberration likely preceded the *N*-ethyl-*N*-nitrosourea mutagenesis but was uncovered when homozygous gynogenetically diploid animals were created. The fact that three of eight mutants recovered from the T cell/thymus pilot screen fell into the *egy* complementation group (N.S.T., B.H.P., Y.Z., A.Z., C.T.A., and L.I.Z., unpublished data) further supported this hypothesis. The RNA derived from the mutant locus splices from the 5′ splice site of exon 15 onto three exons present in the inserted sequence (Fig. 2*A*, blue boxes with roman numerals I to III), resulting in an aberrant mutant *p110* mRNA with a premature termination codon between the TPR and RRM domains. This finding predicted a functional null mutation because we previously could show that the RRMs are essential for the recycling function (5). Consistent with this finding, no full-length transcript (SI Fig. 5 *B* and *C*) or intact p110 protein (Fig. 2*B*) was detectable in *egy* mutant embryos. Further confirmation that the *egy* mutation lies in the *p110* gene came from *p110* morpholino knockdown (SI Fig. 6), which phenocopied the absence of T cells and the arch defect in *egy* mutants. The analysis of p110 expression in 19-h postfertilization and 2-dpf embryos showed no spatially restricted expression and no appreciable difference between *egy* mutant and wild-type individuals (SI Fig. 7).

***egy* Mutation Leads to Reduced U4/U6 di-snRNP Levels.** The human orthologous *p110* gene maps to the chromosomal location 12q24 and encodes a general splicing factor required for snRNP recycling. Specifically, p110 helps to reanneal U4 and U6 to the functional U4/U6 di-snRNP (4). To obtain functional evidence that the mutated zebrafish gene encodes the *p110* ortholog, we first examined the status of U4 and U6 snRNAs in the *egy* mutant embryos by using coimmunoprecipitations from embryo lysates (Fig. 3). Both the trimethyl cap structure (m_3G cap) and the Sm proteins are present in the U4, but not in the U6, snRNP; therefore, coimmunoprecipitation of U6 snRNA with anti-m_3G cap or anti-Sm antibodies is indicative of the association of U4 and U6 in the U4/U6 di-snRNP (see Fig. 4*C* for a schematic representation of the various snRNPs containing U4 and U6 snRNAs). Quantitation of these Northern blots clearly demonstrated that the U4–U6 interaction was strongly reduced in the mutant embryos (coimmunoprecipitation of U6 by anti-m_3G cap decreased from 12.7% to 6.8%; by anti-Sm, from 4.1% to 1.4%). In contrast, U4 immunoprecipitation levels did not differ significantly between wild-type and mutant embryos (anti-m_3G cap, 15.0% versus 16.2%; anti-Sm, 10.3% versus 9.9%). In addition, specifically in the mutant embryos, the U4 snRNA was partially degraded. The degradation products (marked by asterisks in Fig. 3) were present in the anti-Sm but not in the anti-m_3G cap immunoprecipitate (Fig. 3, compare lanes 7 and 8), demonstrating that these partially degraded U4 snRNAs still contain the 3′ terminal Sm core domain yet lack the characteristic m_3G cap at their 5′ end. These conclusions also were confirmed after glycerol gradient fractionation of embryo lysates, followed by coimmunoprecipitation with anti-m_3G cap antibodies (SI Fig. 8). In addition, we assayed the steady-state levels of U4 and U6 snRNAs in the *egy* mutant embryos (SI Fig. 9). U6 snRNA levels

BIOCHEMISTRY

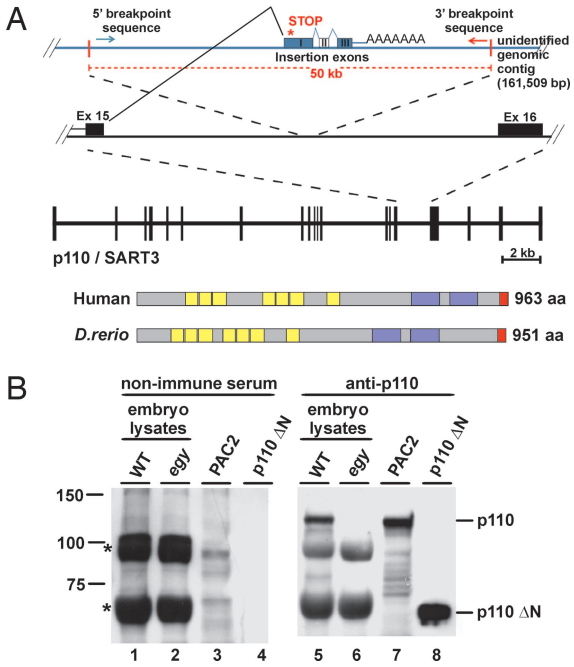

Fig. 2. Identification of the *D. rerio p110/SART3* gene and mapping of the *egy* mutation. (*A*) Molecular analysis of the *egy* mutant locus. The genomic exon 15–16 region is shown in which a 50-kb insertion was identified (*Upper*, depicted in blue) derived from an unidentified genomic contig. Polyadenylation of the mutant transcript is depicted as a stretch of *A* residues. (*Lower*) The complete exon–intron structure of the zebrafish *p110/SART3* gene is given (26 kb; 19 exons shown as boxes, and introns are shown as lines), as well as the schematic domain structures of the human and *D. rerio* p110 proteins (963 and 951 aa, respectively; HAT motifs of the N-terminal TPR-domain shown as yellow boxes, RRMs is in blue, and C-terminal sequence motif is in red). (*B*) p110 protein is absent in *egy* mutant embryos. Equivalent amounts of embryo lysates from wild-type (lanes 1 and 5) or *egy* mutant (lanes 2 and 6) embryos were subjected to Western blotting using nonimmune serum (lanes 1–4) or an anti-zebrafish p110 antibody (lanes 5–8). As positive control, cellular lysate from cultured PAC2 cells (lanes 3 and 7), a zebrafish-derived cell line, as well as recombinant p110 protein lacking the N-terminal TPR domain (p110 ΔN; lanes 4 and 8) were loaded. Asterisks mark nonspecific signals from abundant yolk proteins. Molecular mass markers (in kilodaltons) are shown on the left, and the positions of full-length p110 protein and the ΔN derivative are marked on the right.

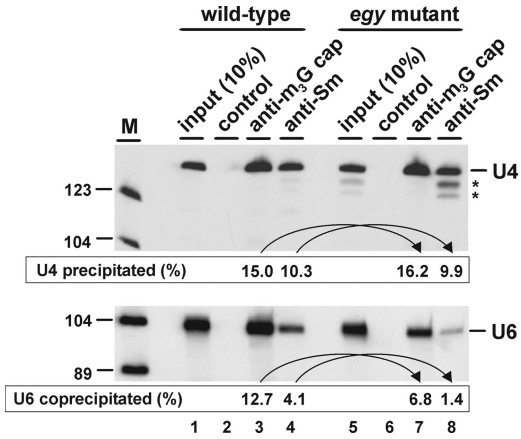

Fig. 3. *egy* mutant zebrafish embryos are defective in U4/U6 di-snRNP formation. snRNPs were immunoprecipitated from wild-type (lanes 1–4) or *egy* mutant (lanes 5–8) zebrafish embryo lysates by using either anti-m_3G cap (lanes 3 and 7) or anti-Sm antibodies (lanes 4 and 8). Precipitated RNAs subsequently were analyzed by Northern blotting against U4 and U6 snRNAs (*Upper* and *Lower*, respectively). Nonimmune serum served as a control (lanes 2 and 6), and 10% of the input was loaded for comparison (lanes 1 and 5). Quantitations of precipitated U4 and coprecipitated U6 snRNAs relative to the input are listed below each lane. Partially degraded U4 snRNA species are marked by asterisks. Molecular weight markers (in nucleotides) are shown on the left.

clearly were reduced, most likely as a result of accumulation of singular U6 snRNPs, which cannot be converted to functional U4/U6 di-snRNPs, followed by U6 snRNA degradation. In summary, the U4–U6 interaction is disrupted in the *egy* embryos to a large degree, demonstrating the *in vivo* function of p110 in U4/U6 snRNP assembly.

Gene Expression Profiling Reveals a Set of Coregulated Splicing Factors. To further characterize the system-wide effects of an inactivating mutation in a specific general splicing factor, microarray analyses from 3- and 5-dpf *egy* embryos were performed. Affymetrix zebrafish GeneChips allowed the expression analysis of ≈14,900 zebrafish transcripts. Initially, we focused on the up-regulated transcripts, for which the orthologous human genes were identified by using the National Center for Biotechnology Information (NCBI) HomoloGene database. Fig. 4*A* lists only genes up-regulated by a fold change ($\log_2$ ratio of mutant to wild-type signals) >1.5 for both 3- and 5-dpf expression profiles. More complete data sets of the up-regulated genes (fold change >1.0) are shown in SI Tables 1 and 2.

Of the 76 up-regulated genes with human homologs (i.e., 66 at 3 dpf and 61 at 5 dpf), 50 genes occur in both 3- and 5-dpf data sets. Surprisingly, 50% of all of the up-regulated genes (38 of 76) encode a subset of snRNP proteins and other splicing-related factors; again, most transcripts of this subset were found in both data sets (31 of 38; see Fig. 4*B* for a graphical representation).

More specifically, first, among the 31 up-regulated genes are six of the seven canonical Sm proteins, which are all part of the common Sm core of the spliceosomal snRNPs (SmB, SmD3, SmE, SmF, SmG, and SmD1), LSm proteins (LSm6, LSm7, and LSm8, which are all part of the heptameric U6 snRNP-specific LSm core), and most U4/U6-, U5-, and U4/U6.U5-tri-snRNP-specific protein components. Second, among the up-regulated factors are the La protein (SSB/La), a known U6 snRNP biogenesis factor, and SNAPC4, an snRNA-specific transcription factor. Third, we found up-regulated factors that are documented as spliceosome-associated, based on recent proteomic analyses in the human system (3). Examples of this group include FRG1, XAB2 (hSyf1), RUVBL1, HTATSF1, and CRNKL1 (hSyf3); however, most of these latter factors have no clearly defined functional role in the spliceosome cycle. Three of the up-regulated factors were confirmed experimentally (LSm8, SmD3, and PRPF31; see SI Fig. 10*A*). In contrast, other spliceosome components that act independently of p110-catalyzed U4/U6 snRNP assembly, such as the U1-specific proteins, the U2-specific core proteins U2A′ and B″, members of the hnRNP family and the classical SR proteins, were not found in the list of up-regulated genes.

These data suggest the existence of an extensive network of coregulated factors of the spliceosome cycle and snRNP biogenesis. Significantly, this network includes not only components physically associated with each other in the same RNA–protein complex but also some factors that are only functionally linked to snRNP biogenesis (La protein) and snRNA transcription (SNAPC4). The specificity of this network is underscored by the fact that all LSm proteins represented by probes on the microarray were up-regulated, with the exception of LSm1, which is not part of the U6 snRNP core but plays an important role in RNA degradation.

Consistency between the 3- and 5-dpf expression profiles is remarkably high. Almost all of the strongly up-regulated factors

A

3 dpf		5 dpf	
FC	Hs_Homolog	FC	Hs_Homolog
3.19	KIAA0947	3.26	KIAA0947
2.67	TP53	2.85	C13orf22
2.54	PRPF31	2.82	SNRPD3
2.49	FRG1	2.80	PRPF31
2.39	SNRPD3	2.75	LSM8
2.26	C13orf22	2.71	TP53
2.23	LSM8	2.70	PRPF3
2.17	PRPF3	2.48	LSM7
2.09	LSM7	2.45	FRG1
2.07	U5-116KD	2.39	U5-116KD
2.03	HTATSF1	2.34	SNRPG
2.00	WDSAM1	2.29	MPP1
1.94	SNRPG	2.19	ODF2
1.91	HSGT1	2.13	SSB
1.90	SF3A2	2.13	HSGT1
1.88	CCNG1	2.05	XAB2
1.86	ODF2	2.04	PRPF4
1.86	PRPF4	2.03	CLU
1.80	POLR2E	2.03	SNRPE
1.73	SART1	1.99	PHLDA3
1.73	SNRPE	1.99	POLR2E
1.72	MDM2	1.98	RUVBL1
1.69	RUVBL1	1.96	HTATSF1
1.69	SSB	1.92	MDM2
1.67	HSF1	1.91	SNRPF
1.65	SNAPC4	1.90	3'HEXO
1.63	LOC440520	1.90	SF3A2
1.62	RY1	1.89	LOC440520
1.60	HIRIP3	1.83	SF3B4
1.56	FLJ12671	1.78	HTATSF1
1.52	AQR	1.77	FLJ12671
		1.76	SART1
		1.74	ATF3
		1.72	RY1
		1.71	HCCS
		1.69	LOC341720
		1.67	FLJ10349
		1.64	WDSAM1
		1.61	C20orf14
		1.59	SOCS3
		1.55	CRNKL1
		1.50	POLR2B

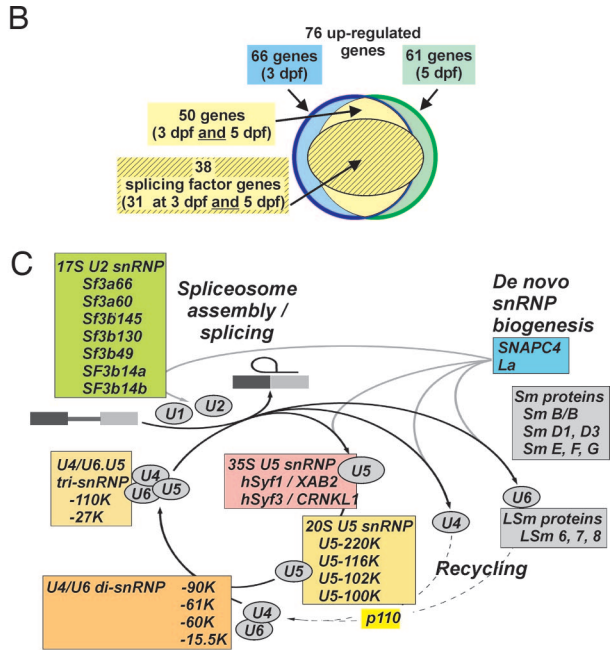

Fig. 4. Expression profiling of *egy* mutant zebrafish: extensive network of coregulated splicing factors and compensation model. (*A*) List of up-regulated genes with a mean fold change >1.5 (FC, $\log_2$ ratios of signals from mutant embryos versus wild-type siblings at 3 and 5 dpf). Only genes in which human homologs could be identified in the NCBI database (human homologs are given in the column *Hs_Homolog*) are listed. Splicing and snRNP biogenesis factors are highlighted in red. (*B*) Graphical representation of the quantitative distribution of up-regulated genes. (*C*) Compensation model and schematic representation of the recycling phase of the spliceosome cycle. The U4/U6.U5 tri-snRNP is disrupted during splicing (shown on the top left corner) and has to be reassembled from its individual components: the U4, U5, and U6 snRNPs. This reassembly proceeds in a stepwise manner: First, the U4/U6 di-snRNP is generated, which then associates with the 20S U5 snRNP that is generated by dismantling of the postspliceosomal 35S U5 snRNP. A deficiency of the p110 protein (highlighted in yellow) results in inefficient recycling of the U4 and U6 snRNPs (dashed lines). Factors up-regulated in the *egy* mutant embryos are listed next to the snRNP complexes for which they are specific (35S and 20S U5 snRNPs, U4/U6 di-snRNP, U4/U6.U5 tri-snRNP, and 17S U2 snRNP), up-regulated common snRNP core components are shown on the right (Sm/LSm proteins), and factors involved in *de novo* snRNP biogenesis in the top right corner.

at 3 dpf remain in the same group at 5 dpf (Fig. 4 *A* and *B*). Note that this list is necessarily incomplete because, first, only a subset of all splicing factors is represented in the Affymetrix zebrafish array, and, second, annotation of the zebrafish genome is unfinished.

What is the biological role of such an extensive network linking >30 splicing-related genes? We propose a model that the absence of p110 in the *egy* mutant elicits a compensatory mechanism, increasing the expression of functionally linked splicing factors and thereby alleviating the specific recycling defect (summarized in Fig. 4*C*). Based on previous *in vitro* studies (4, 12), the biochemical activity of p110 is to promote the U4–U6 interaction. Consistent with this finding, inactivation of the *p110* gene in the zebrafish *egy* mutant results in reduced U4/U6 di-snRNP levels and the accumulation of singular U4 and U6 snRNPs, which apparently cannot be recycled efficiently to functional U4/U6 di-snRNPs. As a consequence of this block in the spliceosome recycling phase, a deficiency of U4/U6.U5 tri-snRNPs is expected to develop in the mutant.

Up-regulation of the subset of splicing-relevant factors identified here would provide a mechanism to partially compensate for this specific recycling defect. For example, increased synthesis of Sm and LSm core proteins as well as of U4/U6 and U5 snRNP-specific components would allow more *de novo* biosynthesis of tri-snRNPs, even under conditions of inefficient U4/U6 di-snRNP formation; higher SNAPC4 levels should stimulate snRNA transcription, and increased La levels may help stabilizing *de novo* transcribed U6 snRNA and U6 snRNP assembly. Consistent with such a compensatory mechanism is the finding that splicing factors were highly enriched in the list of up-regulated genes but were not found among the down-regulated genes (see SI Tables 1–4 for more complete data sets of up- and down-regulated genes). How the lack of functional tri-snRNP is monitored and signaled in the nucleus to up-regulate the subset of splicing-relevant factors remains to be elucidated.

We note that, based on *in vitro* studies, U4/U6 annealing function also has been reported for LSm proteins, suggesting a redundant function of p110 and LSm2–8 in recycling (8, 13, 14). However, the reduced levels of functional di-snRNP, which we observe in *egy* mutant embryos and likely reflect the residual annealing activity of the LSm proteins, clearly are not sufficient for viability. Finally, there is a subset of up-regulated factors that are spliceosome components, but so far without any well defined function, suggesting they may act as novel auxiliary splicing factors during the recycling phase.

Regarding the genes down-regulated in *egy* mutants, we note that most of them are expressed in an organ-specific manner, particularly in the eye and exocrine pancreas, whereas many other tissues are not affected (SI Tables 3 and 4; for validation, see SI Fig. 10*B*); there, a relatively small number of genes are most highly expressed and also require accordingly high mRNA splicing activities. The striking organ-specific distribution of the down-regulated genes appears to reflect the characteristic *egy* phenotype. Most likely, it is based on the fact that in the *egy* mutant the compensatory mechanism (splicing factor up-regulation) does not suffice for rescuing tissues with very high proliferative rates.

Our study establishes zebrafish as a valuable model system to study the system-wide role of splicing factors and their relevance in human disease. Recently, p110 has been implicated as a human disease gene in disseminated superficial actinic porokeratosis (DSAP), an uncommon autosomal dominant chronic skin disorder in a Chinese pedigree (15). Moreover, mutations in several other human splicing factors, all of them components

of the U4/U6.U5 tri-snRNP, could be linked to retinitis pigmentosa: PRPF31 (U4/U6–61K; ref. 16), PRPF3 (U4/U6–90K; ref. 17), and PRPF8 (hprp8; ref. 18). In addition, mutations in the survival of motor neurons (SMN) gene cause spinal muscular atrophy (SMA) (reviewed by refs. 19 and 20). Together, these data suggest that snRNP biogenesis and recycling easily can become limiting factors in the spliceosome cycle, resulting in specific disease phenotypes.

Materials and Methods

Zebrafish Strains. Zebrafish were maintained as described in ref. 21. *N*-ethyl-*N*-nitrosourea mutagenesis was carried out on the AB strain as described in ref. 22. Mutants of T cell and/or thymus development (23) were revealed by using an antisense probe to zebrafish *rag1*. *egy* heterozygous AB individuals were crossed to the WIK strain to generate map-cross generations.

Tissue Culture. Zebrafish PAC2 cells were obtained from N. Foulkes (Max Planck Institute for Developmental Biology, Tübingen, Germany) and cultured in L15 medium (Invitrogen, Carlsbad, CA), supplemented with 15% FBS. Cytoplasmic (S100) and nuclear extracts were prepared as described in ref. 24.

Mutagenesis, Gene Mapping, and Positional Cloning. *N*-ethyl-*N*-nitrosourea mutagenesis was performed as described in ref. 21. Eggs from F_1 females were subjected to early pressure (EP) to generate F_2 gynogenetic diploid offspring (25). F_2 individuals were subjected at 5 dpf to WISH by using the *rag1* probe. F_1 females with mutant offspring were map-crossed, and F_2 individuals were in-crossed to verify the mutant phenotype in the F3 generation. Eggs from heterozygous F_2 females were used for EP mapping (26). For details on gene mapping, positional cloning, and primer sequences, see *SI Materials and Methods*.

WISH. WISH was performed as described in ref. 27.

Zebrafish p110 Expression and Antibody Production. Sequences coding for *Danio rerio* p110 ΔC (amino acids 1–517) and ΔN (amino acids 518–958) were PCR-amplified and cloned into pETM11 by using NcoI and KpnI restriction sites. Recombinant purified proteins were used for rabbit immunization (Biogenes, Berlin, Germany). Affinity-purified antibodies were used for Western blotting.

Embryo Lysates and Western Blotting. *D. rerio* embryo lysates were prepared by homogenizing animals in 10 vol of a buffer containing 20 mM Hepes/KOH (pH 7.5), 150 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, and 1.25% Nonidet P-40. After incubation on ice and pelleting debris by centrifugation, the supernatant was mixed with 1 vol of a buffer containing 20 mM Hepes/KOH (pH 7.5), 50 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, and 40% glycerol. Lysates were separated by 8% SDS/PAGE and analyzed by Coomassie staining or Western blotting (4) by using affinity-purified p110 antibody (1:400). Nonimmune serum (1:1,000) served as a negative control.

snRNP Immunoprecipitation and Northern Blotting. U1, U2, U4, U5, and U6 snRNA as well as 5S rRNA sequences were amplified by RT-PCR from zebrafish embryo total RNA (for oligonucleotides, see *SI Materials and Methods*), cloned into pCR2.1-TOPO vector (Invitrogen), and sequenced. For immunoprecipitation of U4 and U4/U6 snRNPs either anti-cap (H20) antibodies covalently coupled to Sepharose (gift of R. Lührmann, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany) or anti-Sm (Y12) antibodies bound to protein G-Sepharose (Amersham, Piscataway, NJ) were incubated with zebrafish embryo lysate (or glycerol gradient fractions containing 5% glycerol or less) in buffer D (28). After extensive washing, bound RNAs were eluted and analyzed by Northern blotting (5, 29).

Microarray Analysis. ***RNA extraction.*** *egy* mutant embryos and wild-type siblings were collected in duplicates at 3-dpf stage and in triplicates at 5-dpf stage and separately processed. After homogenization (Tekmar Tissumizer, Cincinnati, OH), total RNA was extracted by using TRIzol reagent (Invitrogen) and purified on RNeasy columns (Qiagen, Valencia, CA) according to the manufacturer's recommendations. The quantity and quality of total RNA was assessed by absorbance at 260 and 280 nm and by gel electrophoresis.

Target preparation, hybridization, and signal detection. Seven micrograms of total RNA was converted to cDNA (SuperScript II kit; Invitrogen) by priming with an oligo(dT) primer that included a T7 RNA polymerase promoter site at the 5′ end. cDNA then was used directly in an *in vitro* transcription reaction in the presence of biotinylated nucleotides (BioArray High-Yield RNA transcript labeling kit; Enzo, Farmingdale, NY) to produce biotin-labeled cRNA (antisense RNA). Fifteen micrograms of cRNA subsequently was fragmented (Ambion, Foster City, CA) and hybridized to Affymetrix zebrafish GeneChips, according to the manufacturer's guidelines. After staining with a streptavidin-phycoerythrin conjugate (Molecular Probes, Carlsbad, CA), the fluorescence of bound RNA was quantitated by using a GeneChip scanner (Affymetrix, Santa Clara, CA).

Gene expression profile analysis. The raw expression data were processed and analyzed by using the Bioconductor (www.bioconductor.org) affy package and the Golden Spike R package (www.elwood9.net/spike; refs. 30 and 31). Probe sets with q values <0.001 were defined as differentially expressed between wild-type and mutant embryos. Alternatively, three other Bioconductor packages (vsn, rma, and qcrma) were implemented, and ≈90% of the significantly differentially expressed genes were in agreement between the different methods (data not shown). Probe sets that were differentially expressed are listed with their target-sequence accession nos. (provided by Affymetrix), which were mapped to the NCBI UniGene and EntrezGene databases (August 2005) for zebrafish gene annotation. The human homologs are annotated according to NCBI HomoloGene database (August 2005) (see SI Tables 1–4). Array data have been deposited in NCBI Gene Expression Omnibus (GEO) under the accession no. GSE 5586.

Real-Time Quantitative RT-PCR. Total RNA from 3- and 5-dpf wild-type and mutant embryos was prepared with TRIzol reagent (Invitrogen). Traces of genomic DNA were removed by RQ1 DNase (Promega, Madison, WI). Comparable amounts of total RNA were subjected to reverse transcription with a pool of gene-specific primers, by using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. Control reactions were performed in the absence of reverse transcriptase. Aliquots were analyzed by quantitative real-time PCR (ICycler; Bio-Rad, Hercules, CA) using CYBR Green JumpStart *Taq* ReadyMix (Sigma, St. Louis, MO) and mRNA-specific primer sets (SI Table 5).

We thank Nick Foulkes (Max Planck Institute for Developmental Biology, Tübingen, Germany) for zebrafish PAC2 cells and Reinhard Lührmann (Max Planck Institute for Biophysical Chemistry, Göttingen, Germany) for H20 anti-m_3G cap antibodies. We acknowledge the expert assistance of Tatsuya Otha and Hirohide Kawasaki. This work was supported by grants from the Irvington Institute for Immunological Research (to N.S.T.), National Heart, Lung, and Blood Institute Grant K08 HL04233–05 (to N.S.T.), from the National Institutes of Health (to B.H.P., C.T.A., and L.I.Z.), from the Howard Hughes Medical Institute (to L.I.Z.), Deutsche Forschungsgemeinschaft Grant Bi 316/10 (to A.B.), Federal Ministry for Education and Research BMBF NGNF-2 program (to A.B.), and from the Fonds der Chemischen Industrie (to A.B.).

1. Will CL, Lührmann R (2006) in *The RNA World*, eds Gesteland RF, Cech TR, Atkins JF (Cold Spring Harbor Lab Press, Woodbury, NY), 3rd Ed, pp 369–400.
2. Brow DA (2002) *Annu Rev Genet* 36:333–360.
3. Jurica MS, Moore MJ (2003) *Mol Cell* 12:5–14.
4. Bell M, Schreiner S, Damianov A, Reddy R, Bindereif A (2002) *EMBO J* 21:2724–2735.
5. Medenbach J, Schreiner S, Liu S, Lührmann R, Bindereif A (2004) *Mol Cell Biol* 24:7392–7401.
6. Damianov A, Schreiner S, Bindereif A (2004) *Mol Cell Biol* 24:1700–1708.
7. Kwan SS, Brow DA (2005) *RNA* 11:808–820.
8. Rader SD, Guthrie C (2002) *RNA* 8:1378–1392.
9. Shannon KW, Guthrie C (1991) *Genes Dev* 5:773–785.
10. Traver D, Herbomel P, Patton EE, Murphey RD, Yoder JA, Litman GW, Catic A, Amemiya CT, Zon LI, Trede NS (2003) *Adv Immunol* 81:253–330.
11. Blackburn CC, Manley NR (2004) *Nat Rev Immunol* 4:278–289.
12. Raghunathan PL, Guthrie C (1998) *Science* 279:857–860.
13. Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R (1999) *EMBO J* 18:5789–5802.
14. Karaduman R, Fabrizio P, Hartmuth K, Urlaub H, Lührmann R (2006) *J Mol Biol* 356:1248–1262.
15. Zhang ZH, Niu ZM, Yuan WT, Zhao JJ, Jiang FX, Zhang J, Chai B, Cui F, Chen W, Lian CH, *et al.* (2005) *Br J Dermatol* 152:658–663.
16. Vithana EN, Abu-Safieh L, Allen MJ, Carey A, Papaioannou M, Chakarova C, Al-Maghtheh M, Ebenezer ND, Willis C, Moore AT, *et al.* (2001) *Mol Cell* 8:375–381.
17. Chakarova CF, Hims MM, Bolz H, Abu-Safieh L, Patel RJ, Papaioannou MG, Inglehearn CF, Keen TJ, Willis C, Moore AT, *et al.* (2002) *Hum Mol Genet* 11:87–92.
18. McKie AB, McHale JC, Keen TJ, Tarttelin EE, Goliath R, van Lith-Verhoeven JJ, Greenberg J, Ramesar RS, Hoyng CB, Cremers FP, *et al.* (2001) *Hum Mol Genet* 10:1555–1562.
19. Kennan A, Aherne A, Humphries P (2005) *Trends Genet* 21:103–110.
20. Faustino NA, Cooper TA (2003) *Genes Dev* 17:419–437.
21. Mullins MC, Hammerschmidt M, Haffter P, Nüsslein-Volhard C (1994) *Curr Biol* 4:189–202.
22. Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden FJ, Jiang YJ, Heisenberg CP, *et al.* (1996) *Development (Cambridge, UK)* 123:1–36.
23. Trede NS, Zapata A, Zon LI (2001) *Trends Immunol* 22:302–307.
24. Lee KA, Bindereif A, Green MR (1988) *Gene Anal Tech* 5:22–31.
25. Beattie CE, Raible DW, Henion PD, Eisen JS (1999) *Methods Cell Biol* 60:71–86.
26. Johnson SL, Africa D, Horne S, Postlethwait JH (1995) *Genetics* 39:1727–1735.
27. Thisse B, Heyer V, Lux A, Alunni V, Degrave A, Seiliez I, Kirchner J, Parkhill JP, Thisse C (2004) *Methods Cell Biol* 77:505–519.
28. Dignam JD, Lebovitz RM, Roeder RG (1983) *Nucleic Acids Res* 11:1475–1489.
29. Bell M, Bindereif A (1999) *Nucleic Acids Res* 27:3986–3994.
30. Weber GJ, Choe SE, Dooley KA, Paffett-Lugassy NN, Zhou Y, Zon LI (2005) *Blood* 106:521–530.
31. Choe SE, Boutros M, Michelson AM, Church GM, Halfon MS (2005) *Genome Biol* 6:R16.