1.36 ng mm^{-2} and a k_d of 0.081 s^{-1} (Fig. 4B). When the two proteins were mixed, a Γ_{eq} of 5.95 ng mm^{-2} was reached. An apparent k_d > 0.0008 s^{-1} was determined by fitting a mono-exponential function to the dissociation phase, which is 3 and 2 orders of magnitude lower as compared with the single proteins (fig. S7). Both SOPC and DOPS:SOPC bilayers showed no interaction with the control protein MBP (Fig. 4, A and B). Dissociation of bound CHMP2AΔC and CHMP3ΔC in Hepes buffered saline (HBS) containing 1 M NaCl revealed a $k_d \geq 1$ s^{-1}, which was much faster than dissociation in HBS alone. In contrast, the CHMP2AΔC-CHMP3ΔC polymer did not dissociate with a higher rate ($k_d \leq 0.00046$ s^{-1}) in the presence of 1 M NaCl (fig. S7), indicating resistance to change in ionic strength. Once assembled on membranes, CHMP2AΔC-CHMP3ΔC did not exchange with soluble or membrane-bound CHMP3ΔC (fig. S8). Thus, CHMP2A-3 complexes assembled on membranes in vitro in the absence of CHMP4-6 subcomplexes, even though yeast Snf7-Vps20 (CHMP4-6) complexes may recruit Vps2-Vps24 (CHMP2A-3) complexes to membranes in vivo (*7*).

To assess the influence of CHMP2AΔC-CHMP3ΔC tubes on membrane shapes, we used large unilamellar vesicles (LUVs) composed of DOPS:SOPC. LUV incubation with either CHMP2AΔC or CHMP3ΔC had no effect on their floatation in sucrose gradients (fig. S9, A and B), whereas preformed CHMP2AΔC-CHMP3ΔC tubes restricted LUV floatation to the middle of the gradient (fig. S9, C and D). Negative staining EM confirmed CHMP2AΔC-CHMP3ΔC tube membrane interaction via their outer surfaces (Fig. 4, C and D). However, no systematic remodeling of the LUV membranes was observed. Potential membrane remodeling by the CHMP copolymer or vice versa was further explored by assembling the polymer in the presence of LUVs. Although CHMP2AΔC-CHMP3ΔC assembly in the presence of SOPC LUVs had no effect on tube morphology (Fig. 4E), the presence of DOPS:SOPC LUVs produced shorter tubes (Fig. 4F), displaying loose helical coils (Fig. 4G) and cone-shaped tubes that appeared closed at the narrower end (Fig. 4H). Thus, this suggests a mechanism where lipid interaction affects CHMP polymerization.

Because modified VPS4 and CHMP3 exert dominant negative effects on HIV-1 budding (*8*, *9*, *17*, *25*) and cytokinesis (*6*, *18*), CHMP2A-CHMP3-VPS4 complexes may catalyze a common step such as membrane fission. The CHMP2A-CHMP3 polymer presents a membrane binding topology that is inverse to that of dynamin membrane complexes (*26*), which catalyze endocytotic vesicle abscission (*27*). ESCRT-III coupled to VPS4 may exert a similar role in budding processes directed away from the cytosol. Thus, we propose that a helical CHMP2A-CHMP3 polymer assembles on the inside of a membrane bud, which may induce membrane deformation, leading to constriction and eventually abscission when coupled to VPS4 activity, the only energy-providing candidate in the pathway (*2*, *15*) (fig. S10).

References and Notes

1. R. L. Williams, S. Urbe, *Nat. Rev. Mol. Cell Biol.* **8**, 355 (2007).
2. S. Saksena, J. Sun, T. Chu, S. D. Emr, *Trends Biochem. Sci.* **32**, 561 (2007).
3. J. H. Hurley, *Curr. Opin. Cell Biol.* **20**, 4 (2008).
4. P. D. Bieniasz, *Virology* **344**, 55 (2006).
5. J. G. Carlton, J. Martin-Serrano, *Science* **316**, 1908 (2007); published online 6 June 2007 (10.1126/science.1143422).
6. E. Morita *et al.*, *EMBO J.* **26**, 4215 (2007).
7. M. Babst, D. J. Katzmann, E. J. Estepa-Sabal, T. Meerloo, S. D. Emr, *Dev. Cell* **3**, 271 (2002).
8. U. K. von Schwedler *et al.*, *Cell* **114**, 701 (2003).
9. T. Muziol *et al.*, *Dev. Cell* **10**, 821 (2006).
10. A. Zamborlini *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19140 (2006).
11. S. Lata *et al.*, *J. Mol. Biol.* **378**, 816 (2008).
12. Y. Lin, L. A. Kimpler, T. V. Nalsmith, J. M. Lauer, P. I. Hanson, *J. Biol. Chem.* **280**, 12799 (2005).
13. S. Shim, L. A. Kimpler, P. I. Hanson, *Traffic* **8**, 1068 (2007).
14. P. I. Hanson, R. Roth, Y. Lin, J. E. Heuser, *J. Cell Biol.* **180**, 389 (2008).
15. M. Babst, B. Wendland, E. J. Estepa, S. D. Emr, *EMBO J.* **17**, 2982 (1998).
16. N. Bishop, P. Woodman, *Mol. Biol. Cell* **11**, 227 (2000).
17. B. Strack, A. Calistri, E. Popova, H. Göttlinger, *Cell* **114**, 689 (2003).
18. J. D. Dukes, J. D. Richardson, R. Simmons, P. Whitley, *Biochem. J.* **411**, 233 (2008).
19. E. H. Egelman, *Ultramicroscopy* **85**, 225 (2000).
20. Materials and methods are available as supporting materials on *Science* Online.
21. M. D. Stuchell-Brereton *et al.*, *Nature* **449**, 740 (2007).
22. T. Obita *et al.*, *Nature* **449**, 735 (2007).
23. A. Scott *et al.*, *EMBO J.* **24**, 3658 (2005).
24. P. Whitley *et al.*, *J. Biol. Chem.* **278**, 38786 (2003).
25. J. Martin-Serrano, A. Yarovoy, D. Perez-Caballero, P. D. Bieniasz, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12414 (2003).
26. S. M. Sweitzer, J. E. Hinshaw, *Cell* **93**, 1021 (1998).
27. K. Takei, P. S. McPherson, S. L. Schmid, P. De Camilli, *Nature* **374**, 186 (1995).

28. We thank X. Siebert (Institut de Biologie Structurale) for advice on UROX. This work was supported by Deutsche Forschungsgemeinschaft (SPP 1175) (W.W.), the Agence Nationale de la Recherche sur le SIDA (W.W.), Université Joseph Fourier (W.W.), the Agence Nationale de la Recherche (G.S.), the CNRS (G.S.), the NIH (grant AI29873, H.G.), and by postdoctoral fellowships from the European Molecular Biology Organization (S.L.) and the International Human Frontier Science Program Organization (S.L.). The EM map has been deposited at the European Bioinformatics Institute, accession code EMD-1536.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5894/1354/DC1
Materials and Methods
Figs. S1 to S10
References

28 May 2008; accepted 24 July 2008
Published online 7 August 2008;
10.1126/science.1161070
Include this information when citing this paper.

A Neoplastic Gene Fusion Mimics Trans-Splicing of RNAs in Normal Human Cells

Hui Li,[1] Jinglan Wang,[1] Gil Mor,[2] Jeffrey Sklar[1]*

Chromosomal rearrangements that create gene fusions are common features of human tumors. The prevailing view is that the resultant chimeric transcripts and proteins are abnormal, tumor-specific products that provide tumor cells with a growth and/or survival advantage. We show that normal endometrial stromal cells contain a specific chimeric RNA joining 5′ exons of the *JAZF1* gene on chromosome 7p15 to 3′ exons of the Polycomb group gene *JJAZ1/SUZ12* on chromosome 17q11 and that this RNA is translated into JAZF1-JJAZ1, a protein with anti-apoptotic activity. The *JAZF1-JJAZ1* RNA appears to arise from physiologically regulated trans-splicing between precursor messenger RNAs for *JAZF1* and *JJAZ1*. The chimeric RNA and protein are identical to those produced from a gene fusion found in human endometrial stromal tumors. These observations suggest that certain gene fusions may be pro-neoplastic owing to constitutive expression of chimeric gene products normally generated by trans-splicing of RNAs in developing tissues.

Recurrent, specific gene fusions arising from chromosomal rearrangements are characteristic features of many neoplasms, especially those having hematopoietic and mesenchymal origins (*1*–*6*). In most fusions, recombination occurs within introns that interrupt the coding sequences, giving rise to the expression of chimeric proteins (*2*). The prevailing view is that the chimeric proteins resulting from chromosomal rearrangements are entirely abnormal and have neoplastic effects leading to the growth and/or survival advantage of cells containing them.

An observation that seems at odds with this view is that chimeric mRNAs identical to those derived from fusion genes can often be detected in low abundance by reverse transcription–

[1]Department of Pathology, Yale University School of Medicine, New Haven, CT 06520, USA. [2]Department of Obstetrics and Gynecology, Yale University School of Medicine, New Haven, CT 06520, USA.

*To whom correspondence should be addressed. E-mail: jeffrey.sklar@yale.edu

polymerase chain reaction (RT-PCR) of RNA from healthy tissues (*7*). The explanation generally offered for this finding is that specific chromosomal rearrangements occur within small numbers of cells in healthy tissues but that the chimeric proteins generated by them are alone insufficient to drive substantial clonal expansion.

We have previously described a gene fusion due to a (7;17)(p15;q21) chromosomal translocation found in about 50% of human endometrial stromal sarcomas (ESSs) (*8*, *9*). The fusion joins the first three exons (from a total of five) in the gene *JAZF1* to the last 15 (from 16) in the Polycomb group gene *JJAZ1/SUZ12*. Expression of the chimeric JAZF1-JJAZ1 protein in cultured human embryonic kidney (HEK) 293 cells confers resistance to apoptosis and, when accompanied by suppression of the unrearranged *JJAZ1* allele, increased rates of proliferation (*9*).

We examined normal human endometrial tissues for possible chimeric *JAZF1-JJAZ1* RNA, beginning with endometrial stromal cell lines. RNA extracted from the immortalized, normal human endometrial stromal cell line (HESC) (*10*) was analyzed for the presence of *JAZF1-JJAZ1* chimeric RNA by RT-PCR with primers containing sense and antisense sequence flanking the site of joining between *JAZF1* and *JJAZ1* (*11*). A single amplification product generated by this reaction was identical in size to that amplified from human ESSs carrying a *JAZF1-JJAZ1* gene fusion due to the presence of a t(7;17)(p15;q21) (Fig. 1A). RT-PCR for the *JAZF1-JJAZ1* RNA in two additional, nonimmortalized primary cell lines derived from the normal endometrial stroma of two other patients amplified similarly sized products. Nucleotide sequence analysis of the RT-PCR products from each cell line yielded the same sequence of nucleotides at the *JAZF1-JJAZ1* junction as was found in RNA of tumors with the gene fusion. RT-PCR for *JAZF1-JJAZ1* RNA failed to amplify products from the RNA extracted from a variety of other epithelial and mesenchymal cell lines, all of which contained *JAZF1* and *JJAZ1* RNA (fig. S1).

To investigate the specificity of the junction between *JAZF1* and *JJAZ1* RNA sequences in the HESC cell line, we carried out detailed RT-PCR studies on RNA from the HESC cell line using antisense primers for *JJAZ1* exon 3 sequence paired with six different sense primers for five exons of *JAZF1*. With primers for the first three exons of *JAZF1*, single products were obtained matching the sizes predicted for the joining of *JAZF1* exon 3 to *JJAZ1* exon 2 (Fig. 1B). Similar results were achieved when the sense primer for *JAZF1* exon 3 was paired with six different antisense primers distributed among the 16 exons of *JJAZ1*. These results are consistent with the *JAZF1-JJAZ1* RNA joined at exon 3 and exon 2 of the respective genes being the only abundant *JAZF1-JJAZ1* RNA in HESC cells.

To determine whether the *JAZF1-JJAZ1* RNA is translated into protein, we performed Western blot analysis with JJAZ1-specific antibody on protein extracts prepared from HESC cells. This analysis detected a protein identical in size to JAZF1-JJAZ1 protein detected in ESSs (Fig. 1C and supporting online text).

The detection of *JAZF1-JJAZ1* RNA in endometrial stromal cell lines was duplicated by RT-PCR analysis of RNA extracted from formalin-fixed, paraffin-embedded tissues from normal human uteri. *JAZF1-JJAZ1* RNA was detected primarily in endometrium from late secretory and early proliferative phases of the menstrual cycle (Fig. 2A). No *JAZF1-JJAZ1* RNA sequences were amplified from normal myometrium at any phase of the cycle (fig. S2).

Because of the general association of the *JAZF1-JJAZ1* RNA with endometrium from particular phases of the menstrual cycle, we investigated the effects of steroid hormones on the production of the chimeric transcript in HESC cells. Low concentrations of progesterone seemed to slightly increase amounts of the *JAZF1-JJAZ1* RNA seen in the absence of added hormone, whereas both estrogen and, at higher concentrations, progesterone suppressed detection of the chimeric RNA (Fig. 2B). These findings are consistent with the results of analyses on endometrial tissue, showing that *JAZF1-JJAZ1* RNA is present predominantly at the beginning and end of the menstrual cycle, when hormone concentrations are low.

Because normal endometrium is subjected to hypoxia and undergoes apoptosis during the late secretory phase of the menstrual cycle, we investigated whether hypoxia can induce production of *JAZF1-JJAZ1* RNA in HESC cells by treatment with desferroximine (DFO), which simulates hypoxic conditions. HESC cells treated with 250 μM DFO for 8 hours showed increased amounts of *JAZF1-JJAZ1* RNA (Fig. 2C). Another,

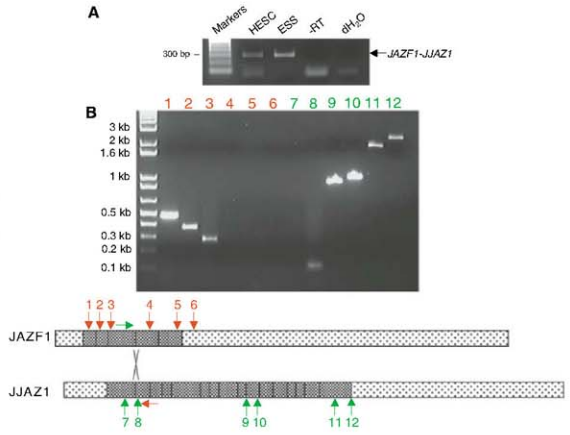

Fig. 1. Detection of chimeric *JAZF1-JJAZ1* RNA and protein in the HESC cell line. **(A)** RT-PCR for the chimeric junction in the *JAZF1-JJAZ1* transcript with primers complementary to antisense and sense sequence in *JAZF1* exon 3 and *JJAZ1* exon 2/3 (*11*). The figure shows the results of agarose gel electrophoresis of amplification products. Results of RT-PCR with RNA from an ESS containing the t(7;17)(p15;q21) are shown in the lane labeled ESS; results of the RT-PCR procedure with RNA from the HESC cell line omitting reverse transcriptase are shown in the lane labeled RT; results without template RNA are shown in the lane labeled dH_2O. **(B)** Analyses by RT-PCR for the specificity of exon joining between *JAZF1* and *JJAZ1* RNAs in HESC cells. Lanes 1 to 6 used six different forward primers at the positions of the downward orange arrows above the diagram of the *JAZF1* transcript, paired with a reverse primer indicated by the orange arrow below the *JJAZ1* transcript. Lanes 7 to 12 used the forward primer indicated in green above the *JAZF1* transcript, paired with six different reverse primers at the positions of the upward green arrows below the *JJAZ1* transcript. **(C)** Western blot of protein extracts from ESS tissue, the HESC cell line, and HEK 293 cells (as a negative control) for]JAZ1 and]AZF-]JAZ1 protein with]JAZ1-specific antibody.

nonimmortalized normal endometrial stromal cell line, HESC-597, also showed up-regulation of *JAZF1-JJAZ1* RNA when treated with DFO. Cultures of cells derived from tissue other than endometrial stroma showed no detectable chimeric RNA with DFO treatment. Analysis of the RNA in DFO-treated HESC cells by a nuclease protection assay (fig. S3) indicated that somewhat less chimeric RNA was produced in these cells than in tumor cells containing the *JAZF1-JJAZ1* fusion, consistent with relative amounts of 10 to 35% detected by quantitative RT-PCR. The nuclease protection assay also confirmed that detection of chimeric RNA was not a methodologic artifact associated with RT-PCR. Quantitative RT-PCR revealed that treatment of HESC cells with DFO did not appreciably change the amount of either *JAZF1* or *JJAZ1* RNA (fig. S4), suggesting that DFO raises the levels of *JAZF1-JJAZ1* RNA by a mechanism independent of increased transcription of the two genes.

To investigate whether *JAZF1-JJAZ1* RNA is produced from a t(7;17)(p15;q21) in HESC cells, we first showed that this cell line, which had not intentionally been cloned, had in fact originated from a single cell immortalized in culture (fig. S5). We then performed Southern blot analyses of HESC DNA by using probes that had previously detected t(7;17)(p15;q21) rearrangements in ESSs. No rearranged bands were detected. Cytogenetic analysis of numerous metaphase spreads from HESC cells revealed no abnormalities in chromosomes 7, 17, or any other chromosome (Fig. 3A). Similarly, analysis by fluorescence in situ hybridization (FISH) with pairs of bacterial artificial chromosome (BAC) probes for DNA flanking on either side the chromosome 7p15 breakpoint and separately the 17q21 breakpoint detected no breakage in these regions of the genome (fig. S5). Additionally, a probe consisting of a yeast artificial chromosome (YAC) that contains DNA spanning the 7p15 breakpoint showed no splitting of the fluorescent signal (Fig. 3B). Finally, no superimposition of signals was observed when probes for chromosomes 7p15 and 17q21 were used together in FISH studies.

To investigate the possibility that (7; 17)(p15;q21) translocations or their equivalents arose in cells in culture at some point after immortalization, we subcloned HESC cells by limiting dilution. Thirty-seven subclones derived on average from half a cell per culture were tested for the production of *JAZF1-JJAZ1* RNA. RT-PCR of RNA from these subclones detected *JAZF1-JJAZ1* RNA in all clones examined, and for most subclones, the amount of *JAZF1-JJAZ1* RNA increased when DFO was added to the culture (Fig. 3C). Furthermore, analyses of all clones examined were negative for rearrangements at the chromosome 7p15 site by FISH with flanking probes (Fig. 3D). Ten of these subclones were also tested for rearrangements at the 17q12 site by FISH, and none of these showed abnormalities.

Given the evidence against DNA recombination in HESC cells and the precise joining of sequences at exon boundaries in *JAZF1-JJAZ1* RNA, we reasoned that the mechanism most likely responsible for production of this RNA is trans-splicing of pre-mRNAs for the *JAZF1* and *JJAZ1* genes. To test this hypothesis, we prepared in vitro splicing extracts from the nuclei of HESC cells. Samples of this extract were mixed with samples of a nuclear extract from a primary rhesus fibroblast cell line RF (*12*). RT-PCR of *JJAZ1* intron 1 RNA sequence revealed that unspliced pre-mRNA was present in the HESC and RF nuclear extracts (fig. S7). Nucleotide sequence analysis of exon 3 in the *JAZF1* gene of RF cells showed two single base-pair sequence differences from the human *JAZF1* gene that permitted both selective RT-PCR of any RNA containing either rhesus or human exon 3 and the ability to distinguish between the products amplified from these RNAs. With selective primers and conditions, amplification of RNA after incubation of mixed extracts yielded products in which RF *JAZF1* exon 3 was joined to exon 2 of *JJAZ1* (Fig. 4B). Sequence analysis of the RT-PCR products confirmed that the amplified *JAZF1-JJAZ1* sequences contained exon 3 of RF *JAZF1* (Fig. 4C). The amount of product generally increased when the HESC extract was prepared from cells cultured with DFO, although the extent of increase varied considerably among experiments. No product was obtained from extracts of HESC cells or RF cells alone, or when adenosine 5′-triphosphate (ATP), an obligate cofactor for splicing, was omitted from the splicing reaction. Similar results were obtained with extracts from the nonimmortalized endometrial stromal cell line, HESC-597 (Fig. 4D).

To study the mechanism further and to rule out the possibility of polymerase switching during transcription, we carried out the in vitro trans-splicing assay with HESC nuclear extract mixed with purified RF RNA. The amount of trans-spliced product detected was similar to that in the

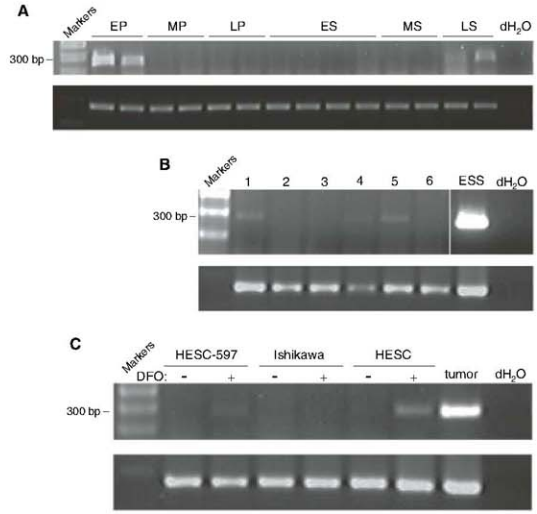

Fig. 2. Detection of *JAZF1-JJAZ1* RNA in endometrial tissues and effects of hormones and hypoxia on amounts of *JAZF1-JJAZ1* RNA in cultured cells. Analyses were performed by RT-PCR, as in Fig. 1A. Unlabeled panels show results of RT-PCR for β-actin RNA as a control for input RNA. **(A)** Detection of *JAZF1-JJAZ1* RNA in total RNA extracted from endometrial samples representing various phases of the menstrual cycle: EP, early proliferative; MP, mid-proliferative; LP, late proliferative; ES, early secretory; MS, mid-secretory; LS, late secretory. At least two separate uteri were tested for each phase. **(B)** Effect of hormone treatment on *JAZF1-JJAZ1* RNA in HESC cells. After 2 days of serum starvation, medium containing serum and no drug (lane 1), 17 β-estradiol at 5×10^{-8} M (lane 2), progesterone at 1×10^{-6} M, 1×10^{-7} M, or 1×10^{-8} M (lanes 3 to 5), or 17 β-estradiol at 5×10^{-8} M plus progesterone at 1×10^{-7} M (lane 6) was added to the cells for 24 hours. **(C)** Effect of DFO treatment on *JAZF1-JJAZ1* RNA in HESC and HESC-597 cells. No chimeric transcript could be detected in Ishikawa cells, an endometrial carcinoma line.

assay performed with a mixture of HESC and RF nuclear extracts (Fig. 4E). Elimination of all PCR-detectable traces of DNA from the RF RNA preparation by treatment with deoxyribonuclease I did not affect the production of the trans-spliced product (figs. S7 and S8).

The data presented here are consistent with trans-splicing of the pre-mRNAs transcribed from the *JAZF1* and *JJAZ1* genes in normal

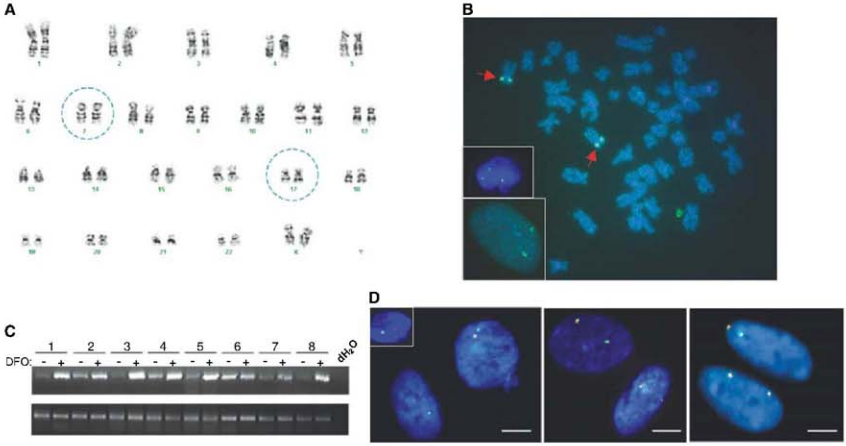

Fig. 3. Absence of the t(7;17)(p15;q21) in HESC cells. **(A)** Cytogenetic analysis of HESC cells. Normal chromosomes 7 and 17 are circled. **(B)** FISH analysis of HESC cells with a YAC probe containing DNA spanning the *JAZF1* locus. Arrows point to the intact YAC signal in metaphase chromosomes. (Lower inset) A representative interphase HESC nucleus; (upper inset) an ESS control showing splitting of the probe signal. **(C)** RT-PCR analysis for *JAZF1-JJAZ1* RNA in representative subclones of the HESC cell line with and without DFO treatment. **(D)** FISH analysis of three HESC subclones with two BAC probes, labeled red or green and each containing DNA sequences that flank the *JAZF1* locus on one side or the other. The juxtaposition of red and green signals indicates no separation of these sequences. (Inset) Separation of signals in an ESS control. Bars, ~10 μm.

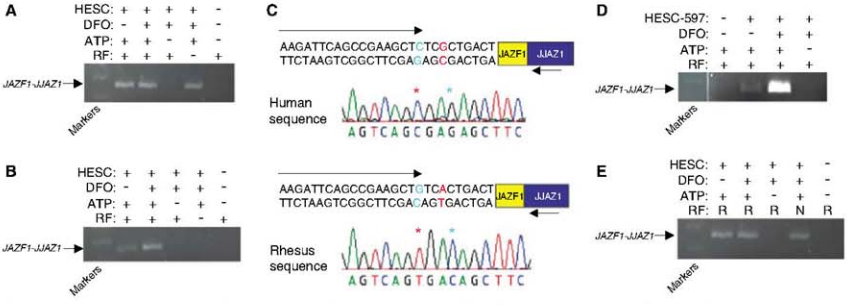

Fig. 4. In vitro trans-splicing reactions. **(A)** RT-PCR analysis for *JAZF1-JJAZ1* RNA with a human-specific primer. RT-PCR products from chimeric RNA were amplified only when ATP was supplied, and the amount of product increased when the HESC nuclear extracts were prepared from cells pretreated with DFO. **(B)** RT-PCR analysis for *JAZF1-JJAZ1* RNA with a rhesus-specific primer. No band was detected in HESC or RF extracts alone, but products were observed when the two kinds of nuclear extracts were mixed. The amount of product increased when the HESC extract was prepared from the cells pretreated with DFO. **(C)** Sequence analysis of the RT-PCR products amplified with antisense human- or rhesus-specific primers. Aqua-colored bases and stars indicate the interspecies sequence differences included in the species-specific primers; red bases and stars indicate the species-specific sequence differences detected in the products. **(D)** RT-PCR analysis for *JAZF1-JJAZ1* RNA in nuclear extracts of the HESC-597 and RF cells with rhesus-specific primers. Mixed extracts produced detectable signal in the presence of ATP. **(E)** RT-PCR analysis for *JAZF1-JJAZ1* RNA in HESC nuclear extracts mixed with purified RF RNA by means of amplification with the rhesus-specific primer. R, RNA; N, nuclear extract.

endometrial stromal cells and tissues to yield chimeric products identical to those produced by a recurrent gene fusion in endometrial stromal tumors. Trans-splicing of noncoding, leader exons to separately transcribed pre-mRNAs is common in certain lower eukaryotes, such as protozoa and nematodes (*13*–*16*). However, in vertebrates, only a few examples of trans-splicing have been described (*17*–*25*), and most of these involve splicing between pre-mRNAs of the same gene to generate mRNAs with duplicated exons (*17*–*20*). For these reasons, trans-splicing in vertebrates has sometimes been regarded as a nonfunctional by-product of a somewhat sloppy splicing system (*26*). This conclusion seems inapplicable to *JAZF1-JJAZ1* RNA because the *JAZF1-JJAZ1* fusion gene is a recurrent finding in a high fraction of endometrial stromal tissues, and fusion genes associated with chromosomal translocations in cancer have repeatedly been shown to contribute to the neoplastic phenotype of the tumors containing them (*27*, *28*). Additionally, the expression of the *JAZF1-JJAZ1* coding sequences in cultured cells has demonstrated effects on cell survival and proliferation (*9*). Whether JAZF1-JJAZ1 protein in tissues provides protection from hypoxia, to which endometrium is subjected during the late secretory phase and possibly the early proliferative phase of the menstrual cycle, remains to be determined.

The mechanisms involved in the trans-splicing of RNAs and the regulation of this process are unclear. Juxtaposition of the loci encoding the RNAs that participate in trans-splicing would not seem essential because in vitro splicing in nuclear extracts of RNAs at physiological concentrations was found to be efficient. These results also indicate that cotranscriptional splicing is not an absolute requirement for trans-splicing. Whether in vivo trans-splicing of RNA transcribed from loci that participate in chromosomal rearrangements predisposes DNA at those sites to recombination with or without prior intranuclear colocalization of the loci will require further investigation.

In view of the regulated trans-splicing between *JAZF1* and *JJAZ1* pre-mRNAs in normal endometrium, the t(7;17)(p15;q21) found in ESSs might be considered a mutation that leads to constitutive production of the *JAZF1-JJAZ1* mRNA and its protein product. This relation is similar to that seen in other oncogenic mutations associated with tumor development, namely, that mutations lead to overproduction or irreversible activation of gene products rather than to creation of "new" genes, as the fusion genes resulting from many chromosomal translocations and other DNA rearrangements have generally been thought to be.

If RNA products of fusion genes other than *JAZF1-JJAZ1* also mimic normal products resulting from trans-splicing of pre-mRNAs, it would explain the ability to frequently amplify from healthy tissues chimeric RNAs associated with chromosomal rearrangements in tumors. Considering the large number of recurrent gene fusions found in tumors, it would further suggest that trans-spliced RNAs may be relatively common in normal cells and tissues (supporting online text). At a minimum, the finding of the trans-spliced *JAZF1-JJAZ1* RNA in normal cells implies a risk to inferring the presence of chromosomal rearrangements in tissue specimens for the diagnosis and detection of cancer, especially in the context of minimal disease. Additionally, it is possible that drugs designed to target chimeric proteins produced by neoplastic gene fusions may have toxicities due to inhibited function of similar proteins in normal cells.

References and Notes

1. M. A. Pierotti *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1616 (1992).
2. F. Mitelman, B. Johansson, F. Mertens, *Nat. Rev. Cancer* **7**, 233 (2007).
3. O. Bernard *et al.*, *Oncogene* **6**, 1477 (1991).
4. P. J. Kourlas *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2145 (2000).
5. A. Pardanani *et al.*, *Blood* **102**, 3093 (2003).
6. S. A. Tomlins *et al.*, *Science* **310**, 644 (2005).
7. S. Janz, M. Potter, C. S. Rabkin, *Genes Chromosomes Cancer* **36**, 211 (2003).
8. J. I. Koontz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6348 (2001).
9. H. Li *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 20001 (2007).
10. G. Krikun *et al.*, *Endocrinology* **145**, 2291 (2004).
11. Materials and methods are available as supporting material on *Science* Online.
12. R. C. Desrosiers *et al.*, *J. Virol.* **71**, 9764 (1997).
13. L. Bonen, *FASEB J.* **7**, 40 (1993).
14. M. Krause, D. Hirsh, *Cell* **49**, 753 (1987).
15. N. Agabian, *Cell* **61**, 1157 (1990).
16. R. E. Sutton, J. C. Boothroyd, *Cell* **47**, 527 (1986).
17. C. Caudevilla *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12185 (1998).
18. A. N. Akopian *et al.*, *FEBS Lett.* **445**, 177 (1999).
19. S. A. Frantz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5400 (1999).
20. T. Takahara, S. I. Kanazu, S. Yanagisawa, H. Akanuma, *J. Biol. Chem.* **275**, 38067 (2000).
21. C. Finta, P. G. Zaphiropoulos, *J. Biol. Chem.* **277**, 5882 (2002).
22. G. Flouriot, H. Brand, B. Seraphin, F. Gannon, *J. Biol. Chem.* **277**, 26244 (2002).
23. C. Fitzgerald *et al.*, *J. Biol. Chem.* **281**, 38172 (2006).
24. Z. Jehan *et al.*, *Genome Res.* **17**, 433 (2007).
25. C. Zhang *et al.*, *DNA Cell Biol.* **22**, 303 (2003).
26. T. Maniatis, B. Tasic, *Nature* **418**, 236 (2002).
27. C. S. Huettner, P. Zhang, R. A. Van Etten, D. G. Tenen, *Nat. Genet.* **24**, 57 (2000).
28. A. T. Look, *Science* **278**, 1059 (1997).

29. Supported by the National Cancer Institute (grant R01 CA85995) and a generous gift from the Burnstein-Turnbull family. We thank M. Martel for assistance in the histologic evaluation of tissue samples; R. Means for providing the rhesus fibroblast cell line RF; P. Li for cytogenetic analysis; V. Patriub for technical assistance; and H. Taylor, J. Steitz, D. Mishler, E. Ullu, and A. Krensky for helpful discussion.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5894/1357/DC1
Materials and Methods
SOM Text
Figs. S1 to S8
References

20 February 2008; accepted 7 July 2008
10.1126/science.1156725

Germline Allele-Specific Expression of *TGFBR1* Confers an Increased Risk of Colorectal Cancer

Laura Valle,[1] Tarsicio Serena-Acedo,[1] Sandya Liyanarachchi,[1] Heather Hampel,[1] Ilene Comeras,[1] Zhongyuan Li,[2] Qinghua Zeng,[2] Hong-Tao Zhang,[2] Michael J. Pennison,[2] Maureen Sadim,[2] Boris Pasche,[2]* Stephan M. Tanner,[1]* Albert de la Chapelle[1]*

Much of the genetic predisposition to colorectal cancer (CRC) in humans is unexplained. Studying a Caucasian-dominated population in the United States, we showed that germline allele-specific expression (ASE) of the gene encoding transforming growth factor–β (TGF-β) type I receptor, *TGFBR1*, is a quantitative trait that occurs in 10 to 20% of CRC patients and 1 to 3% of controls. ASE results in reduced expression of the gene, is dominantly inherited, segregates in families, and occurs in sporadic CRC cases. Although subtle, the reduction in constitutive *TGFBR1* expression alters SMAD-mediated TGF-β signaling. Two major *TGFBR1* haplotypes are predominant among ASE cases, which suggests ancestral mutations, but causative germline changes have not been identified. Conservative estimates suggest that ASE confers a substantially increased risk of CRC (odds ratio, 8.7; 95% confidence interval, 2.6 to 29.1), but these estimates require confirmation and will probably show ethnic differences.

The annual worldwide incidence of colorectal cancer (CRC) exceeds 1 million, being the second to fourth most common cancer in industrialized countries (*1*). Although diet and lifestyle are thought to have a strong impact on CRC risk, genes have a key role in the predisposition to this cancer. A positive family history of CRC occurs in 20 to 30% of all probands. Highly penetrant autosomal dominant and recessive hereditary forms of CRC account for at most 5% of all CRC cases (*2*). Although additional high- and low-penetrance alleles have been proposed, much of the remaining predisposition to CRC remains unexplained (*3*).

Aberrations in the transforming growth factor–β (TGF-β) pathway are heavily involved in CRC carcinogenesis (*4*). Although mutations in the TGF-β type II receptor gene have been explicitly associated with CRC (*5*), the type I receptor gene (*TGFBR1*) has received less attention, although there is evidence that a common variant may be associated with cancer risk (*6, 7*). We hypothesized that *TGFBR1* is a notable candidate for a gene that, when mutated, causes predisposition to CRC or acts as a modifier of other genes, resulting in a predisposition. Our study was undertaken to test this assumption.

Given the previously existing evidence that inherited allele-specific expression of *APC* acts as a mechanism of predisposition to familial adenomatous polyposis (*8*) and of an analogous mechanism involving *DAPK1* in chronic lymphocytic leukemia (*9*), we searched for a similar association of *TGFBR1* with CRC. We hypothesized that the putative change might be subtle; for instance, lowered rather than extinguished expression of one allele referred to here as ASE, for allele-specific expression. To test for ASE in *TGFBR1*, we chose three single-nucleotide polymorphisms (SNPs) (rs334348, rs334349, and rs1590) in the 3′ untranslated region (3′UTR), to which primer extension with fluorescent nucleotides (SNaPshot) (*10*) was applied. These three SNPs are separated by 1916 and 1778 base pairs (bp), respectively, yet they exhibit total linkage disequilibrium.

Among a total of 242 patients with microsatellite instability (MSI)–negative CRC (*10*), 96 (39.7%) were heterozygous for the three 3′UTR SNPs, of whom 12 showed ASE variation ratios higher than 1.5, whereas no patient showed ratios below 0.67. Forty-nine additional cases were heterozygous for one further SNP (rs7871490) located in the 3′UTR that was not in strong linkage disequilibrium with the above three markers, and 17 out of 49 (17/49) had ASE values higher than 1.5. Thus, 29 out of 138 (21%) informative CRC patients showed ASE in the *TGFBR1* gene. Three additional cases had borderline values (fig. S1 and table S1).

DNA samples from the blood of healthy Columbus, Ohio–area controls (195 individuals) (*10*) were genotyped for the four SNPs. One hundred and nine (55.9%) were heterozygous, and ASE analysis in 105 of them revealed ratios ranging between 0.72 and 3.25 (fig. S1). Only three controls showed ratios above 1.5. Our results in both the CRC patients and controls suggest that the degree of ASE is a quantitative trait (Fig. 1). Differences in the degree of ASE between patients and controls showed a *P* value of 0.1208 when a Wilcoxon rank sum test was applied and a *P* value of 0.0207 when a permutation test (100,000 permutations) was applied.

At this stage, it is not possible to determine whether the degree of predisposition to CRC is proportional to the degree of ASE or whether there is a threshold value that separates "abnormal" values that predispose to CRC from "normal" values that do not. A ratio of 1 means that both alleles are equally expressed, whereas a ratio of 1.5 means a 33% difference, as does a ratio of 0.67. To define a cutoff point, we applied receiver operating characteristic (ROC) analysis, which estimates the sensitivity and specificity of cutoff points. As shown in table S2, the value of 1.5 maximizes both characteristics, providing the highest Youden's index. When a cutoff of 1.5 was used, the *P* value comparing cases and controls was 7.655×10^{-5}. Although there is no overall need to define a firm cutoff point, we used the value of 1.5 to categorize CRC cases and controls into ASE and non-ASE. In order to determine whether the observed ratios falling outside this range represent an increase or decrease in the transcript of one allele, a reverse transcription polymerase chain reaction (RT-PCR) experiment was performed, taking advantage of hybrid clones monoallelic for chromosome 9 created from two individuals with ASE (patients 1 and 26, table S1). Each of the four hybrid clones contained either the maternal or paternal copy of chromosome 9, plus the mouse genome (*10*). As shown in Fig. 2A, ASE determination in the diploid samples indicated that the expression of one allele (a) was reduced as compared to that of the other allele (b). In the four monoallelic hybrid clones, the densitometric values of the RT-PCR of human *TGFBR1* were compared with the corresponding values for mouse *Gpi* (*10*). One allele (a) showed reduced expression in both patients. These experiments support the notion of lowered expression of one allele, and in both patients the same allele was affected (Fig. 2, A and B).

To assess the effect of ASE on TGF-β signaling, lymphoblastoid cell lines from four ASE patients and four non-ASE healthy controls were exposed to TGF-β (*10*), which binds TGFBR2 and leads to the formation of the TGFBR2/TGFBR1/TGF-β heteromeric complex. We observed differences in levels of phosphorylated SMAD2 (pSMAD2), an important downstream effector and surrogate marker of TGF-β signaling (*11, 12*). There were constitutive differences in pSMAD2 expression between ASE patients and non-ASE controls in the absence of exogenously added TGF-β (time 0; Fig. 3A). Differences in pSMAD2 levels became more pronounced upon exposure to TGF-β. These differences were observed at low TGF-β concentrations (<5 pM) (Fig. 3B) and occurred in four out of four ASE cases as compared to non-ASE controls.

It has been shown that phosphorylation of SMAD3 is an essential step in signal transduction by TGF-β for the inhibition of cell proliferation (*13*). Furthermore, *Smad3*-deficient mice are prone to developing colon cancer (*14, 15*). To assess the impact of *TGFBR1* ASE on the phosphorylation of SMAD3, we used an antibody targeting the $Ser^{423/425}$ site in SMAD3 (*10, 16*). Constitutive levels of pSMAD3 were detectable in the lymphoblastoid cell lines of three non-ASE controls, whereas pSMAD3 was barely detectable in one ASE case (Fig. 3C). Exposure to TGF-β did not result in any detectable increase in pSMAD3 in the lymphoblastoid cell lines of the ASE patients. The pSMAD2 and pSMAD3 results indicate that patients with ASE exhibit decreased SMAD-mediated signaling when compared with non-ASE controls.

A GCG trinucleotide variable number of tandem repeat polymorphism occurs in exon 1 of *TGFBR1*. The most common allele contains nine repeats leading to a stretch of nine alanines (9A) in the signal peptide of the receptor protein. The second most common allele has six repeats (6A) and occurs in approximately 14% of all individuals in most Caucasian populations (*6*). The 6A allele has been associated with a low-level but statistically significant predisposition to sev-

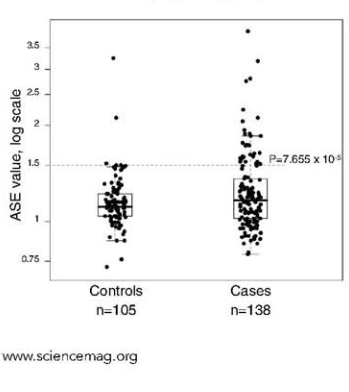

Fig. 1. *TGFBR1* ASE distribution in 138 CRC patients and 105 controls studied by SNaPshot. The ASE cutoff value of 1.5 chosen to categorize the cases is indicated, together with its associated *P* value obtained from comparing the proportions of cases (29/138) and controls (3/105) above the indicated value.

[1]Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, OH 43210, USA. [2]Cancer Genetics Program, Division of Hematology/Oncology, Department of Medicine and Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.

*To whom correspondence should be addressed. E-mail: b-pasche@northwestern.edu (B.P.); Stephan.Tanner@osumc.edu (S.M.T.); Albert.delaChapelle@osumc.edu (A.d.l.C.)

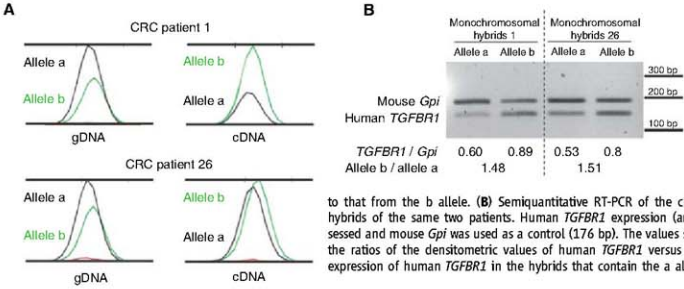

Fig. 2. ASE determination in two ASE CRC probands. **(A)** ASE detection in blood DNA by SNaPshot. The ASE ratio was calculated by normalizing the ratio between the peak areas of the two alleles in cDNA with the same parameters in genomic DNA (gDNA). In both examples, the transcript from the a allele is reduced with respect to that from the b allele. **(B)** Semiquantitative RT-PCR of the cDNA from monochromosomal hybrids of the same two patients. Human *TGFBR1* expression (amplicon size 135 bp) was assessed and mouse *Gpi* was used as a control (176 bp). The values shown below the gel represent the ratios of the densitometric values of human *TGFBR1* versus mouse *Gpi*, showing reduced expression of human *TGFBR1* in the hybrids that contain the a allele.

eral forms of cancer (*17*–*20*). Recent studies suggest that the association of 6A with colon cancer is either weak [odds ratio (OR) 1.2, 95% confidence interval (CI) 1.01 to 1.43] (*17*) or borderline significant (OR 1.13, CI 0.98 to 1.30) (*21*). We typed this polymorphism in all 242 CRC cases studied by us and found 9A/9A in 197, 9A/6A in 40, 6A/6A in 4, and 1 failed (table S3). There were clearly more 9A/6A heterozygotes among the patients with ASE (14/29) than in those without ASE (22/108) ($P = 0.0052$, chi-square test). We tentatively concluded from these data that the 6A allele is probably in linkage disequilibrium with one of the putative mutations that causes ASE, but 6A is not in itself causative of ASE.

All 29 patients showing ASE and three patients with borderline ASE values (1.49, 1.49, and 1.46) ($n = 32$ patients) were studied for genetic changes occurring in the germ line. By sequencing of all nine exons, 2 kb upstream of exon 1, and the entire 3′UTR (*10*), a single sequence change in the coding exons was identified in patient 30, consisting of a coding DNA 1204 T→A (c.1204T>A) missense change in exon 7 that changes a tyrosine to asparagine (p.Tyr401Asn). Its pathogenicity is currently being assessed. Several changes, all previously reported as polymorphic, were identified in the 3′UTR and promoter regions. In three patients, a deletion (del) of two bases (c.1-1782_1783delCA) at 1783 bp upstream of exon 1 was identified in a repetitive sequence of short interspersed nuclear elements. Multiplex ligation-dependent probe amplification (*10*) did not suggest any large rearrangements, deletions, or duplications of exons. In a study of promoter methylation, none of the comparisons of germline methylation status between ASE and non-ASE cases and ASE cases versus controls were significant (supporting online material text and table S4). Thus, germline promoter methylation is unlikely to play a role in ASE.

We hypothesized that changes occurring in noncoding regions of the gene could be responsible for the reduction in expression. To fully study this possibility, overlapping fragments of 1.7 to 10 kb were amplified by long-range PCR, cloned, and sequenced. In all, approximately 96.5 kb covering the whole gene and 3′UTR (49 kb), 35 kb upstream of exon 1 (up to the next gene *COL15A1*), and 12.5 kb downstream of the 3′UTR (Fig. 4) were fully sequenced in the four monochromosomal hybrids (patients

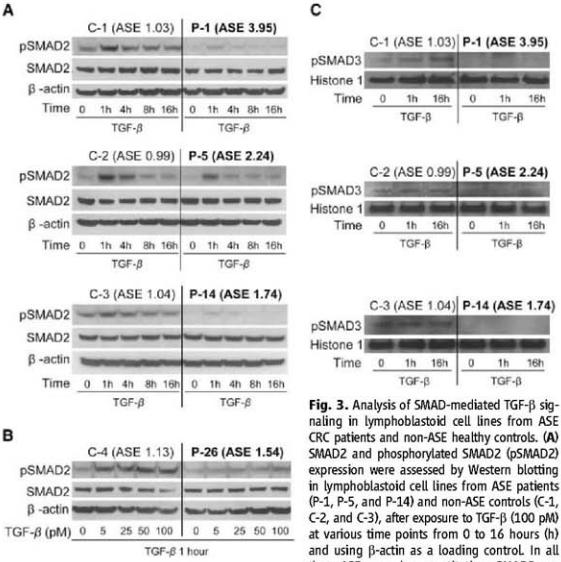

Fig. 3. Analysis of SMAD-mediated TGF-β signaling in lymphoblastoid cell lines from ASE CRC patients and non-ASE healthy controls. **(A)** SMAD2 and phosphorylated SMAD2 (pSMAD2) expression were assessed by Western blotting in lymphoblastoid cell lines from ASE patients (P-1, P-5, and P-14) and non-ASE controls (C-1, C-2, and C-3), after exposure to TGF-β (100 pM) at various time points from 0 to 16 hours (h) and using β-actin as a loading control. In all three ASE cases, less constitutive pSMAD2 was observed than in non-ASE controls. The differences in pSMAD2 expression between ASE and non-ASE cell lines were further enhanced after exposure to TGF-β. **(B)** SMAD2 and p-SMAD2 expression 1 hour after exposure to different TGF-β concentrations. The effect shown in (A) also occurs at low concentrations of TGF-β (5 pM). **(C)** pSMAD3 detection in nuclear extracts from three ASE patients and three non-ASE controls after exposure to TGF-β1. The three non-ASE lymphoblastoid cell lines had pSMAD3 expression in the nucleus, whereas nuclear pSMAD3 expression was undetectable in two ASE cases (P-1 and P-14) and barely detectable in one case (P-5).

1 and 26) and in diploid DNA from four other ASE patients (patients 5, 11, 14, and 21) (*10*). Our sequencing strategy allowed us to determine the phase of every change within each amplicon and over larger regions when at least one change occurred in the overlapping fragments. In all, 25 and 104 changes were identified in the down-regulated alleles of patients 1 and 26, respectively, whereas 31 and 6 changes were detected in their wild-type counterparts. Diploid DNA from the four patients harbored 61, 37, 33, and 135 changes, respectively.

Excluding changes known to be present in the wild-type alleles, 140 changes were identified in the down-regulated alleles. Only the c.1-1782_1783delCA change stood out as a candidate mutation. It occurred in 3/29 (10.3%) ASE patients, in 0/3 ASE controls, in 1/51 (2%) non-ASE CRC patients, and 1/81 (1.2%) non-ASE controls. In summary, these investigations did not uncover the genetic changes causing ASE.

Genotyping of most changes identified by sequencing was carried out in all available ASE CRC patients, including borderline cases (n = 31), and in 55 non-ASE CRC patients. Construction of haplotypes from the available genotype and haplotype data was performed with PHASE v.2.1.1 (*10*). In all, 60 polymorphisms covering 73.5 kb (from 12 kb upstream of exon 1 to 12.5 kb downstream of the 3′UTR) were used for haplotype inference (table S5). For all ASE and non-ASE patients, the program was run with 1000 permutations with overlapping 10-SNP sliding windows. Haplotype frequency distributions in ASE and non-ASE populations showed significant differences in a genomic region covering the area between the 3′ end of intron 3 to ~5 kb downstream of the 3′ end of the UTR (Fig. 4).

The group of patients carrying the minor allele for the three 3′UTR SNPs in linkage disequilibrium (group 1) was very different from the other group derived from the study of SNP rs7871490 (group 2). Haplotype analysis was performed separately in the two groups, using 50 and 21 SNPs, respectively. In group 1 (n = 53), one major haplotype for the affected alleles was present in 11/14 (78.6%) of ASE but also in 22/39 (56.4%) non-ASE patients (Fig. 4). For group 2 (n = 33), another major haplotype for the affected allele was present in 14/17 (82.4%) of ASE and in 1/16 (6.3%) of non-ASE patients (Fig. 4). Fisher's exact test to compare haplotype proportions showed P values of 0.2031 and 1.260×10^{-5} for groups 1 and 2, respectively. The 6A allele of the 9A/6A polymorphism occurred in the ASE haplotype in all 14 cases of group 2, but not in group 1, where all ASE cases except one were homozygous for the 9A allele.

In search of somatic changes in line with Knudson's two-hit hypothesis, loss of heterozygosity (LOH) analyses as well as a search for somatic mutations in the coding sequences of the gene were performed in DNA from the tumors of 26 ASE patients. Using the described threshold (*10*), 6 cases out of 26 showed LOH. In three of these six cases, the wild-type allele, the one with normal expression in blood, was lost or reduced, whereas in the other three cases, the allele showing germline ASE was lost. Exon-by-exon sequencing of the entire gene in tumors from 26 ASE patients revealed somatic changes in three tumors that were not found in blood DNA. The mutations were: c.634G>A (p.Gly212Asp) in one tumor and c.682_685delAAG (p.Glu228del) in two tumors. These mutations occurred in exon 4, which encodes the kinase domain of the protein. LOH analyses and exon 4 sequencing in 49 tumors of CRC patients without ASE showed that none of these tumors had evidence of somatically acquired mutations, and five showed LOH (table S3). Fisher's exact test comparing proportions of LOH and mutations between ASE and non-ASE cases showed P values of 0.1708 and 0.0355, respectively. The occurrence of somatic mutations in ASE cases but not in controls supports the role of *TGFBR1* as a tumor suppressor gene. On the other hand, the fact that

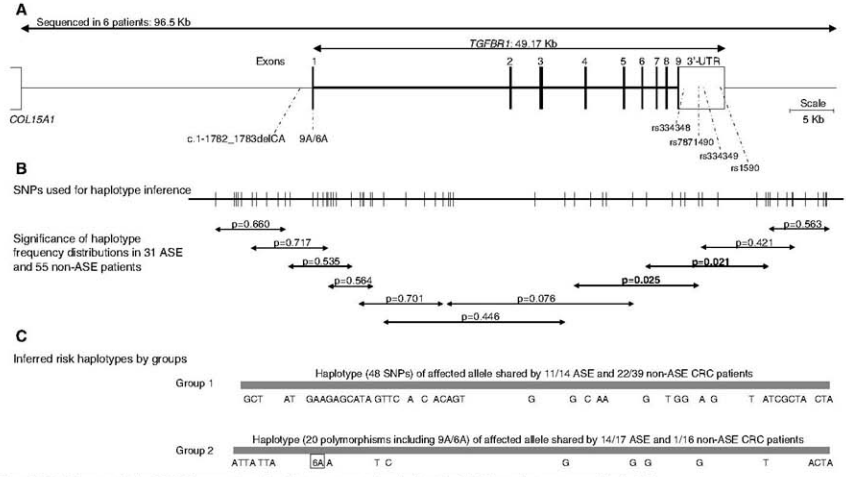

Fig. 4. **(A)** Diagram of the *TGFBR1* genomic region. The uppermost line depicts the 96.5-kb region sequenced in six ASE patients (four monochromosomal hybrids and four diploid DNAs). Shown are the locations of the 2-bp CA deletion upstream of exon 1, the 9A/6A polymorphism in exon 1, and the four SNPs in the 3′UTR used for ASE determinations. **(B)** Locations of the 60 SNPs used for haplotype inference in ASE (n = 31) and non-ASE (n = 55) CRC patients. The arrowed shorter lines each depict a 10-SNP overlapping window. P values indicate the significance of differences in haplotype distribution between ASE and non-ASE individuals. **(C)** Two major haplotypes identified in ASE patients are shown. ☐ 6A allele of the 9A/6A polymorphism

LOH affected the ASE allele as often as the wild-type allele could indicate random losses.

The cohort of MSI-negative CRC patients had been deliberately enriched in familial cases (*10*). In the cohort of 138 patients with available ASE values, 59 out of 136 (43.4%) were familial according to the criteria indicated above, and family information was not available in two cases. Among the cases showing ASE, 53.6% were familial (table S3). The proportion of ASE was higher among familial than nonfamilial cases: 15/59 (25.4%) familial cases versus 13/77 nonfamilial cases (16.9%). A chi-square test to compare proportions showed that this difference was not statistically significant ($P = 0.314$).

The above data suggest that ASE contributes somewhat more to familial than to sporadic CRC but do not allow its inheritance to be assessed. If ASE is regularly inherited as a dominant trait, the expectation is that 50% of first-degree relatives (FDRs) also have ASE. Data from four families that are informative in this regard are shown in fig. S2. In all, among 11 FDRs, ASE was greater than 1.5 in 4, borderline in 2 (ASE values 1.40 and 1.44), and low in 5. There was no instance of ASE being incompatible with Mendelian dominant inheritance. In all four families, co-segregation of ASE with the inferred risk haplotype, representing the down-regulated allele, occurred. The highest Kong and Cox nonparametric LOD score was 1.25, with a P value of 0.008 (nonparametric z score = 4.12; P value = 0.00002). Among the four to six ASE-positive FDRs, two had CRC, one had endometrial cancer and a tubular colonic adenoma, one had prostate cancer, and another had multiple polyps in the colon and rectum (table S6). Although fragmentary, these data suggest dominant inheritance of ASE with incomplete penetrance of CRC in ASE carriers.

There is indirect evidence to support the notion that ASE of *TGFBR1* contributes to CRC development. The TGF-β pathway is strongly involved in the carcinogenesis of colon and other cancers, and its signaling is dependent on the integrity of both of its receptors (TGFBR1 and TGFBR2) (*22*, *23*). In a comprehensive study of CRC tumors, somatic mutations occurred with high frequency in 69/13,023 genes. Among these 69 genes were *TGFBR2*, *SMAD4*, *SMAD2*, and *SMAD3*, attesting to the importance of the TGF-β pathway in CRC (*24*). There is rapidly increasing evidence that subtle variations in gene expression play central roles not only in development in various organisms but also in human disease (*8*, *9*, *25*). Linkage analysis of a cohort of sibling pairs concordant or discordant for colorectal carcinoma or adenoma highlighted a region in chromosome 9q22-31 (*26*). Subsequently, borderline significant linkage to the same region was observed in families segregating colorectal cancer or adenoma without microsatellite instability (*27*, *28*). This evidence is compatible with, but in no way proves, a role for *TGFBR1*.

We were unable to determine what mechanism causes ASE. The haplotype data support the implication of ancestral mutations for most ASE patients. Moreover, the elusive genomic change causing ASE is likely to occur in cis, but the data do not exclude the possibility that ASE arises as a result of trans-acting genes that preferentially affect the risk haplotypes. Such genes could well be RNA genes as predicted earlier (*29*). Very recently, the existence of extensive quantitative trait loci for gene expression was documented in two large studies (*30*, *31*).

How common is ASE of *TGFBR1*? Using our definition, it occurred in 29/138 tested CRC patients (21%) and in 3/105 tested controls (3%). In the extreme, if none of the non-informative CRC cases had ASE, the frequency would be 29/242 (12%), and for the controls, 3/195 (1.5%). Because not all individuals are informative (heterozygous for a transcribed SNP), the true frequency in cases and controls cannot be precisely assessed at present. Using the above alternative numbers, we can calculate the OR of CRC in carriers of ASE. In the first scenario, the OR is 9.0 (CI 2.7 to 30.6), and in the conservative one, OR is 8.7 (CI 2.6 to 29.1).

What proportion of all CRC is attributable to ASE of *TGFBR1*? From the available data of the present case-control study, we estimated the population attributable risk (PAR). If ASE occurs in 21% of cases and 3% of controls, the estimated PAR is 18.7% (CI 10.8 to 25.8). If ASE occurs in 12% of cases and 1.5% of controls, the estimated PAR is 10.6% (CI 6.0 to 14.9). These numbers are estimates, representing the Caucasian-dominated population of central Ohio, and are heavily dependent on the relevant allele frequencies, which may show strong inter-ethnic variation. We nevertheless conclude that ASE of *TGFBR1* is a major contributor to the genetic predisposition to CRC.

References and Notes

1. D. M. Parkin, F. Bray, J. Ferlay, P. Pisani, *CA Cancer J. Clin.* **55**, 74 (2005).
2. A. de la Chapelle, *Nat. Rev. Cancer* **4**, 769 (2004).
3. N. M. Lindor *et al.*, *JAMA* **293**, 1979 (2005).
4. P. M. Siegel, J. Massague, *Nat. Rev. Cancer* **3**, 807 (2003).
5. S. Markowitz *et al.*, *Science* **268**, 1336 (1995).
6. B. Pasche *et al.*, *Cancer Res.* **58**, 2727 (1998).
7. Y. Xu, B. Pasche, *Hum. Mol. Genet.* **16**, R14 (2007).
8. H. Yan *et al.*, *Nat. Genet.* **30**, 25 (2002).
9. A. Raval *et al.*, *Cell* **129**, 879 (2007).
10. Materials and methods are available as supporting material on *Science* Online.
11. J. Massague, R. R. Gomis, *FEBS Lett.* **580**, 2811 (2006).
12. J. Massague, *Mol. Cell* **29**, 149 (2008).
13. X. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10669 (1997).
14. Y. Zhu, J. A. Richardson, L. F. Parada, J. M. Graff, *Cell* **94**, 703 (1998).
15. N. M. Sodir *et al.*, *Cancer Res.* **66**, 8430 (2006).
16. G. Sekimoto *et al.*, *Cancer Res.* **67**, 5090 (2007).
17. B. Pasche *et al.*, *J. Clin. Oncol.* **22**, 756 (2004).
18. B. Pasche *et al.*, *Cancer Res.* **59**, 5678 (1999).
19. B. Pasche *et al.*, *JAMA* **294**, 1634 (2005).
20. Y. Bian *et al.*, *J. Clin. Oncol.* **23**, 3074 (2005).
21. J. Skoglund *et al.*, *Clin. Cancer Res.* **13**, 3748 (2007).
22. B. Bierie, H. L. Moses, *Cytokine Growth Factor Rev.* **17**, 29 (2006).
23. J. Massague, S. W. Blain, R. S. Lo, *Cell* **103**, 295 (2000).
24. T. Sjoblom *et al.*, *Science* **314**, 268 (2006).
25. H. Yan, W. Zhou, *Curr. Opin. Oncol.* **16**, 39 (2004).
26. G. L. Wiesner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12961 (2003).
27. Z. E. Kemp *et al.*, *Cancer Res.* **66**, 5003 (2006).
28. J. Skoglund *et al.*, *J. Med. Genet.* **43**, e7 (2006).
29. M. Morley *et al.*, *Nature* **430**, 743 (2004).
30. H. H. Goring *et al.*, *Nat. Genet.* **39**, 1208 (2007).
31. B. E. Stranger *et al.*, *Nat. Genet.* **39**, 1217 (2007).

32. We thank H. He, K. Sotamaa, and W. Frankel for help. This work was supported by NIH grants CA67941, CA16058, CA112520, and CA108741; grants from the Walter S. Mander Foundation, Chicago, IL; and the Jeannik M. Littlefield–American Association for Cancer Research Grant in Metastatic Colon Cancer Research. L.V. was supported by a fellowship from the Fundación Ramón Areces.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1159397/DC1
Materials and Methods
SOM Text
Figs. S1 and S2
Tables S1 to S6
References

21 April 2008; accepted 5 August 2008
Published online 14 August 2008;
10.1126/science.1159397
Include this information when citing this paper.

New Products Focus: Mass Spectrometry

LC-MS/MS System

The TSQ Vantage is a new liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) system that offers increased sensitivity without increased noise for better reproducibility, accuracy, and precision in quantitative analysis. Technical breakthroughs have improved the TSQ Vantage's signal-to-noise performance, providing better reproducibility and precision in the quantitative analysis of small molecules, biomolecules, and peptides. The instrument offers up to 10 times the signal-to-noise ratio compared with the TSQ Quantum series. The TSQ Vantage's new S-Lens ion optics system features a novel electrostatic field technology to capture virtually every ion and efficiently transfer it into the HyperQuad quadrupole mass analyzer. The S-Lens design is an advance over high-pressure, skimmer-based ion source designs because it eliminates mass discrimination and lowers the gas load on the expensive turbo-molecular pumps.

Thermo Fisher Scientific
For information 800-532-4752
www.thermo.com/tsqvantage

Glycopeptide Analysis Software

The new version of the innovative mass spectrometry (MS) data analysis tool, SimGlycan, can analyze glycopeptides in addition to released glycans. SimGlycan matches MS/MS spectra with its own comprehensive, robust, and annotated database to predict the structure of glycans. It generates a list of all the probable glycans that are close to the experimental data, saving time and laborious work. Each structure is scored to help the user judge which results closely match the experimental data. Alongside the probable glycan structure, the software makes available information such as the glycan class, reaction, pathway, and enzyme. Users can resolve glycan structures in glycopeptide molecules by specifying the sequence or mass of the attached peptide moiety.

Premier Biosoft International
For information 650-856-2703
www.PremierBiosoft.com

Surfactant

ProteasMAX Surfactant is designed to enhance the enzymatic performance of trypsin and other proteases in preparation for analysis by mass spectrometry or liquid chromatography. The surfactant provides a more complete digestion prior to protein analysis, resulting in more accurate data in a shorter time and reduced risk of sample degradation. The new surfactant exposes cleavage sites usually inaccessible because of secondary or tertiary structure and can increase sample recovery from gels. It can reduce standard protein digestion time from overnight to just three hours. It also enhances protein solubilization at room temperature, which means that high temperatures and precipitation can be avoided, even for complex proteins. Because ProteasMAX degrades during the digestion reaction, researchers can proceed directly to analysis of peptides by mass spectrometry without additional deactivation steps.

Promega
For information 608-274-4330
www.promega.com

Cell Lysis Kit

The ProteaPrep Cell Lysis Kit, Mass Spec Grade, is designed for the efficient recovery of purified protein lysates from biological samples. The kit is a proprietary mixture of salts, glycerol, and an acid-labile surfactant optimized for efficient solubilization, extraction, and recovery of proteins during cell lysis. Because the kit contains no lysozyme or harsh detergents, it permits recovery of cell lysates free from artifactual protein contaminants and promotes efficient sample protein solubilization. It is optimal for mass spectrometry, enzymatic digestions, two-dimensional gel electrophoresis, and protein quantitation procedures.

Protea Biosciences
For information 877-776-8321
www.proteabio.com

Tandem Quadrupole MS

The Waters Xevo TQ MS System is an advanced tandem quadrupole mass spectrometer that allows scientists with varying levels of mass spectrometry (MS) expertise to quickly produce high quality data. The system features IntelliStart, a new technology that simplifies instrument setup and troubleshooting. It automatically ensures the system is ready for use by performing mass calibration, setting MS resolution, generating compound-specific MS methods, and optimizing source conditions. The instrument is equipped with a unique collision cell that can be operated in conventional T-Wave-1 enabled mode or in a new ScanWave-enabled mode. ScanWave improves the duty cycle and enhances the full-scan capability to meet the demands of complex analyses, allowing users to more easily confirm the identities and structures of targeted analytes.

Waters
For information 508-482-2614
www.waters.com

High-Resolution Time-of-Flight MS

The maXis electrospray ultrahigh-resolution time-of-flight mass spectrometer offers high resolution of 40,000–60,000 FWHM over a broad mass range and mass accuracy typically between 600 and 800 parts per billion, at speeds of up to 20 full spectra per second. The maXis offers 20 Hz full spectra acquisition at high mass resolution for high-speed liquid chromatography, a high dynamic range of five orders of magnitude for trace detection in complex mixtures, and sub-parts-per-million mass accuracy in both mass spectrometry (MS) and MS/MS mode. The instrument makes high-resolution mass spectrometry over a broad mass range compatible with ultra performance liquid chromatography.

Bruker Daltonics
For information 978-667-9580
www.bdal.com

Electronically submit your new product description or product literature information! Go to **www.sciencemag.org/products/newproducts.dtl** for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.