

Research article

Open Access

## Sequence changes in predicted promoter elements of *STK11/LKB1* are unlikely to contribute to Peutz-Jeghers syndrome

Nicholas CM Hearle\*<sup>1</sup>, Ian Tomlinson<sup>2</sup>, Wendy Lim<sup>3</sup>, Victoria Murday<sup>4</sup>, Edwin Swarbrick<sup>5</sup>, Guan Lim<sup>6</sup>, Robin Phillips<sup>7</sup>, Peter Lee<sup>8</sup>, John O'Donohue<sup>9</sup>, Richard C Trembath<sup>10</sup>, Patrick J Morrison<sup>11</sup>, Andrew Norman<sup>12</sup>, Rohan Taylor<sup>13</sup>, Shirley Hodgson<sup>14</sup>, Anneke Lucassen<sup>15</sup> and Richard S Houlston<sup>16</sup>

Address: <sup>1</sup>Section of Cancer Genetics, Institute of Cancer Research, Sutton, UK, <sup>2</sup>Molecular and Population Genetics Laboratory, Imperial Cancer Research Fund, London, UK, <sup>3</sup>Section of Cancer Genetics, Institute of Cancer Research, Sutton, UK, <sup>4</sup>Dept of Medical Genetics, Yorkhill Hospital, Glasgow, UK, <sup>5</sup>Dept of Gastroenterology, New Cross Hospital, Wolverhampton, UK, <sup>6</sup>Dept of Gastroenterology, Epsom General Hospital, UK, <sup>7</sup>Polyposis Registry, St. Mark's Hospital, Watford Road, Harrow, UK, <sup>8</sup>Dept of Surgery, Hull Royal Infirmary, UK, <sup>9</sup>Dept of Gastroenterology, University Hospital Lewisham, London, UK, <sup>10</sup>Dept of Clinical Genetics, Leicester Royal Infirmary, Leicester, UK, <sup>11</sup>Department of medical genetics, Belfast City Hospital, UK, <sup>12</sup>Clinical Genetics Unit, Birmingham Women's Hospital, Birmingham, UK, <sup>13</sup>Molecular Genetics Unit, St Georges Hospital Medical School, London, UK, <sup>14</sup>Department of Clinical Genetics, St. Georges Hospital, London, UK, <sup>15</sup>Wessex Clinical Genetics Service, The Princess Anne Hospital, Southampton, UK and <sup>16</sup>Section of Cancer Genetics, Institute of Cancer Research, Sutton, UK

Email: Nicholas CM Hearle\* - nick.hearle@icr.ac.uk; Ian Tomlinson - wdlim@yahoo.co.uk; Wendy Lim - wdlim@yahoo.co.uk; Victoria Murday - Victoria.Murday@yorkhill.scot.nhs.uk; Edwin Swarbrick - Edwin.swarbrick@rwh-tr.nhs.uk; Guan Lim - guan.lim@epsom-sthelier.nhs.uk; Robin Phillips - marie.gun@cancer.org.uk; Peter Lee - peterwlee@pmc.edu.my; John O'Donohue - john.o'donohue@uhl.nhs.uk; Richard C Trembath - rtrembath@hgmp.mrc.ac.uk; Patrick J Morrison - patrick.morrison@bch.n-i.nhs.uk; Andrew Norman - andrew.norman@bwhct.nhs.uk; Rohan Taylor - R.TAYLOR@sghms.ac.uk; Shirley Hodgson - shodgson@sghms.ac.uk; Anneke Lucassen - A.M.Lucassen@soton.ac.uk; Richard S Houlston - richard.houlston@icr.ac.uk

\* Corresponding author

Published: 17 March 2005

Received: 15 February 2005

BMC Genomics 2005, 6:38 doi:10.1186/1471-2164-6-38

Accepted: 17 March 2005

This article is available from: <http://www.biomedcentral.com/1471-2164/6/38>

© 2005 Hearle et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Germline mutations or large-scale deletions in the coding region and splice sites of *STK11/LKB1* do not account for all cases of Peutz-Jeghers syndrome (PJS). It is conceivable that, on the basis of data from other diseases, inherited variation in promoter elements of *STK11/LKB1* may cause PJS.

**Results:** Phylogenetic foot printing and transcription factor binding site prediction of sequence 5' to the coding sequence of *STK11/LKB1* was performed to identify non-coding sequences of DNA indicative of regulatory elements. A series of 33 PJS cases in whom no mutation in *STK11/LKB1* could be identified were screened for sequence changes in the putative promoter defined by nucleotides -1090 to -1472. Two novel sequence changes were identified, but were found to be present in healthy individuals.

**Conclusion:** These findings indicate that promoter sequence changes are unlikely to contribute to PJS.

## Background

Peutz-Jeghers syndrome (PJS; MIM 175200) is a rare autosomal dominant disorder typified by hamartomatous polyposis of the gastrointestinal tract and melanin pigmentation of the oro-facial region [1]. Although germline mutations in the coding sequence of the serine-threonine kinase gene *STK11/LKB1* have been found to cause PJS [2,3], such mutations only account for up to 80% of cases [4-12]. In addition to locus heterogeneity [13] mutations in regulatory sequences of *STK11/LKB1* may cause PJS.

Identifying regulatory genomic sequences through functional assays is time consuming and problematic. As natural selection is more likely to tolerate sequence changes in redundant, non-functional sequence than in functionally important sequences, regulatory elements in non-coding sequence will be highly conserved through evolution. Comparison of sequence between both closely related and highly divergent species therefore allows for the identification of non-coding sequence that has a high probability of being important to the regulation of gene transcription. This alternative approach to the identification of promoter elements is termed "phylogenetic foot printing" [14].

Here we describe the phylogenetic foot printing of the 5' upstream region of *STK11/LKB1* and the sequence analysis of this region in a series of PJS cases in whom exonic and splice site mutations in the gene had been excluded.

## Results and discussion

Figure 1 shows the evolutionary conserved region 5' of the coding sequence of *STK11* as identified by the phylogenetic foot-printing programs ECR Browser [15] and CON-SITE [16]. There was a high degree of concordance between the predictions from the programs. ECR Browser predicted three regions of high conservation upstream of *STK11*, encompassing nucleotides -981 to -1668 (nucleotide 0 representing the *STK11* translation initiation signal). Sixteen transcription factor binding sites (TFBS) predicted by rVista [17] resided between nucleotides -1053 and -1472. ConSite predicted 27 TFBSs conserved between human and mouse between nucleotides -1090 and -1605. Through the integration of phylogenetic foot printing and TFBS prediction data from ECR browser/ rVista and ConSite, a consensus region containing predicted TFBS positions was identified between nucleotides -1090 to -1472.

Additional analysis using the TFBS prediction programs Transplorer (BioBase Biological Databases, Wolfenbüttel, Germany) and Proscan v1.7 [18] identified 9 binding sites between nucleotides -1142 and -1724, and 76 binding sites between nucleotides -881 and -1572 respectively, confirming the presence of regulatory elements within the

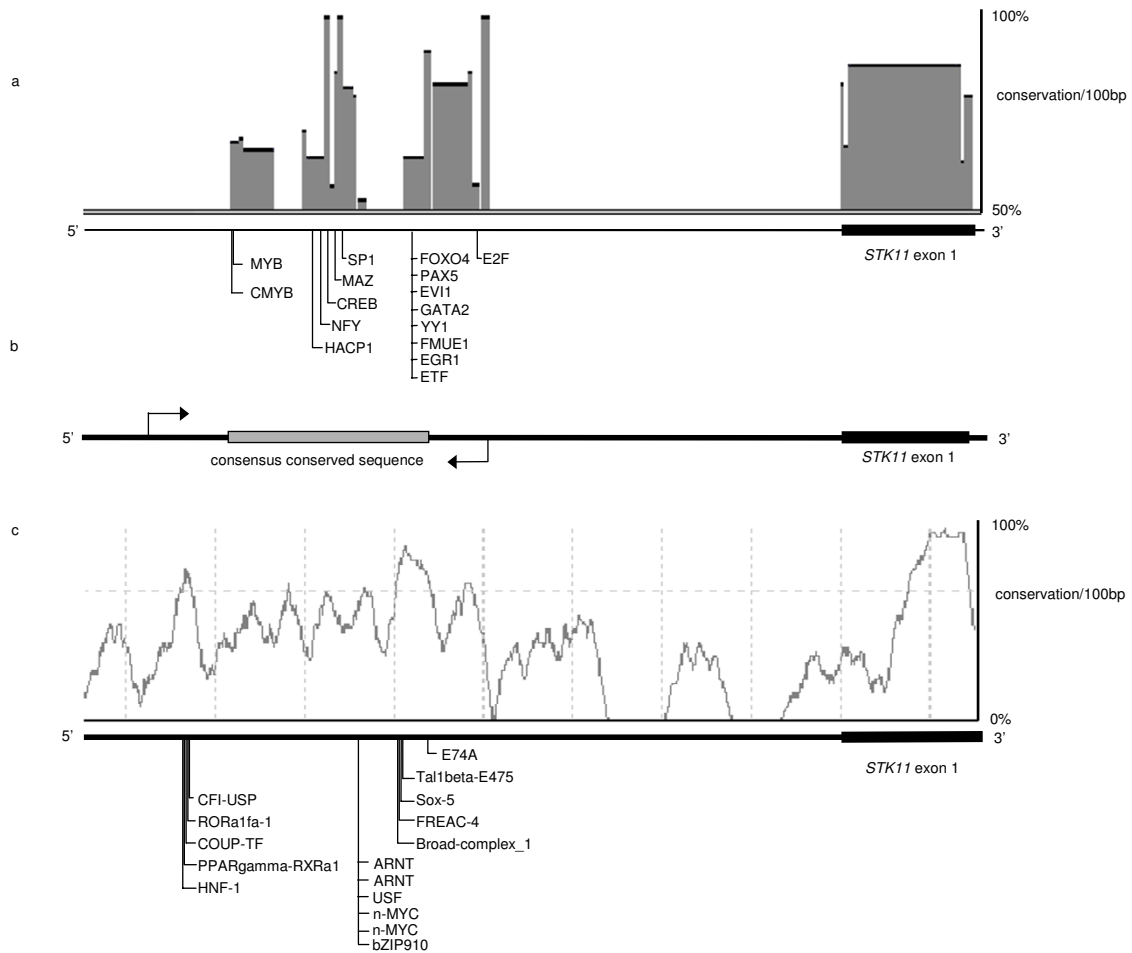
consensus region. ECR Browser/ rVista and ConSite have previously been shown to correctly identify 88% and 66% of TFBS in functionally verified promoter elements respectively [15,16] and it is highly likely, therefore, that the consensus region encompassing nucleotides -1090 to -1472, contains *STK11/LKB1* promoter elements.

DNA from a series of 33 PJS patients that did not harbour germline *STK11/LKB1* mutations (mean age at diagnosis 14 years, range 1-38) was studied for mutations in the region between nucleotides -1001 to -1815, encompassing all TFBSs predicted by rVista and ConSite. Seven of the cases had a documented family history of PJS. The diagnosis of PJS in all cases was based on established criteria [1]. Three sequence changes were identified. The change G → C at position -1566 was found in four cases (three heterozygotes and one homozygote) and represented a previously documented single nucleotide polymorphism (rs3795061). An additional single nucleotide change at position -1268 (G → T) was identified in eight cases (seven heterozygotes and four homozygotes). A single sample of control DNA used as a sequence reference was also found to be homozygous for the change. Finally, a two base pair deletion coupled with a single base pair insertion at position -1709 (n-1709insTdelCC) was identified in one case. A series of healthy population controls (n = 92) was screened by High Performance Liquid Chromatography (HPLC) and one individual (1/92) was found to harbour the sequence change. None of these three sequence changes identified were therefore deemed to be potentially pathogenic.

There is a high degree of redundancy in promoter elements of genes, however point mutations in promoter regions of *PTEN* and *MLH1* have been reported to be disease causing [19,20]. To investigate the possibility of large-scale deletions or insertions undetectable by straightforward PCR primers P1Fwd and P3Rev were used to amplify an 814 bp fragment with products visualised on a 2% agarose gel. No large-scale deletions or insertions were detected in any of the patients.

## Conclusion

As understanding of the contribution of coding sequence changes to disease becomes clearer, attention will focus on regulatory elements of genes. Phylogenetic foot printing using programs such as ECR browser and ConSite present potentially powerful tools in identifying regulatory elements, enabling the analysis of these sequences without time consuming functional studies. Although the efficiency of in-silico delineation of promoter elements has not been rigorously evaluated, ECR Browser/ rVista and ConSite have been shown to correctly identify 88% and 66% of TFBS respectively [15,16]. On the basis of our findings, however, it appears unlikely (upper 95%



**Figure 1**

*in-silico* identification of the putative *STK11/LKB1* promoter. (a) ECR browser output. Regions of high sequence conservation 5' of *STK11* exon I are annotated with transcription factor binding sites as predicted by rVista v2.0 between nucleotides -981 to -1668. (c) ConSite output. Sequence of high conservation is annotated with predicted transcription factor binding sites (TFBS) as predicted by the JASPER TFBS database between nucleotides -1104 to -1613. (b) Ideogram showing consensus conserved region, defined at the 3' end by the E74A TFBS predicted ConSite, and at the 5' end by the CMYB TFBS predicted by rVista. Arrows indicate the genomic sequence analysed.

confidence interval, prevalence; 9%) that mutations in the promoter region of *STK11/LKB1* are responsible for PJS cases not attributable to exonic sequence changes.

**Methods**

**Bioinformatics**

Phylogenetic foot printing of 3 Kb of sequence upstream of human and mouse *STK11/LKB1* (NT\_011255) was performed using the promoter predication programs ECR Browser [15] and CONSITE [16]. For both programs a

conservation cut-off of 70% identity over 100 base pairs was adopted, in accordance with published criteria [21,22]. Regions identified by ECR Browser with greater than 70% conservation were analysed using the TFBS prediction program rVista v2.0 [17]. Conserved regions identified by ConSite were annotated with TFBSs based on the JASPAR database [23] and links provide sequence information. Further analysis was carried out using the TFBS prediction programs Proscan version1.7 [18] and Transplorer (Biobase Biological Databases, Wolfenbüttel, Ger-

many). The promoters and TFBS predicted by each program were aligned to delineate a consensus region indicative of a putative promoter.

### Mutational analysis

The possession of germline mutations in *STK11/LKB1*, including a large-scale deletion of the gene, was excluded using methods previously described [24]. Mutational analysis of the minimal consensus conserved region was carried out by direct DNA sequencing in both directions using the BigDye v3.1 Terminator Cycle Sequencing Ready Reaction Kit in conjunction with an ABI3100 semi-automated genetic analyser (Applied Biosystems, Foster City, USA). Overlapping oligonucleotides amplifying three fragments spanning the region were designed using the Primer3 program [25]: P1Fwd – GCACAGGAGGGT-TCAATATTTTC, P1Rev – TTGCGGACCTGGAAGGAG, P2Fwd – ACTGGAATTGGCCACTTTGT, P2Rev – GATACAGCGCGCTCATTG, P3Fwd – GTCTCCCCATGCCT-GCTTC, P3Rev – GGCCAGCCCATCCAAGG. Predicted PCR products were subjected to searches of the genome using the BLAST program [26] to confirm specificity. Chromatograms were analysed by two independent researchers using the programs Chromas [27] and MutationSurveyor (SoftGenetics, State College, USA). High performance liquid chromatography was performed using the WAVE system (Transgenomics, Omaha, USA).

### Authors' contributions

NCMH performed all bioinformatical and molecular studies and drafted the manuscript. IT, WL, VM, ES, GL, RP, PL, JO, RT, PJM, AN, RCT, SH and AL contributed cases of PJS to the study. RSH conceived the study and drafted the manuscript, and participated in the study's design and coordination. All authors read and approved the final manuscript.

### References

- Tomlinson IP, Houlston RS: **Peutz-Jeghers syndrome.** *J Med Genet* 1997, **34**:1007-1011.
- Hemminki A, Markie D, Tomlinson I, Avizienyte E, Roth S, Loukola A, Bignell G, Warren W, Aminoff M, Hoglund P, Jarvinen H, Kristo P, Pelin K, Ridanpaa M, Salovaara R, Toro T, Bodmer W, Olschwang S, Olsen AS, Stratton MR, de la CA, Aaltonen LA: **A serine/threonine kinase gene defective in Peutz-Jeghers syndrome.** *Nature* 1998, **391**:184-187.
- Jenne DE, Reimann H, Nezu J, Friedel W, Loff S, Jeschke R, Muller O, Back W, Zimmer M: **Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase.** *Nat Genet* 1998, **18**:38-43.
- Boardman LA, Couch FJ, Burgart LJ, Schwartz D, Berry R, McDonnell SK, Schaid DJ, Hartmann LC, Schroeder JJ, Stratakis CA, Thibodeau SN: **Genetic heterogeneity in Peutz-Jeghers syndrome.** *Hum Mutat* 2000, **16**:23-30.
- Jiang CY, Esufali S, Berk T, Gallinger S, Cohen Z, Tobi M, Redston M, Bapat B: **STK11/LKB1 germline mutations are not identified in most Peutz-Jeghers syndrome patients.** *Clin Genet* 1999, **56**:136-141.
- Mehenni H, Gehrig C, Nezu J, Oku A, Shimane M, Rossier C, Guex N, Blouin JL, Scott HS, Antonarakis SE: **Loss of LKB1 kinase activity in Peutz-Jeghers syndrome, and evidence for allelic and locus heterogeneity.** *Am J Hum Genet* 1998, **63**:1641-1650.
- Nakagawa H, Koyama K, Miyoshi Y, Ando H, Baba S, Watatani M, Yasutomi M, Matsuura N, Monden M, Nakamura Y: **Nine novel germline mutations of STK11 in ten families with Peutz-Jeghers syndrome.** *Hum Genet* 1998, **103**:168-172.
- Olschwang S, Markie D, Seal S, Neale K, Phillips R, Cottrell S, Ellis I, Hodgson S, Zuber P, Spigelman A, Iwama T, Loff S, McKeown C, Marchese C, Sampson J, Davies S, Talbot I, Wyke J, Thomas G, Bodmer W, Hemminki A, Avizienyte E, de la CA, Aaltonen L, Tomlinson I: **Peutz-Jeghers disease: most, but not all, families are compatible with linkage to 19p13.3.** *J Med Genet* 1998, **35**:42-44.
- Wang ZJ, Churchman M, Avizienyte E, McKeown C, Davies S, Evans DG, Ferguson A, Ellis I, Xu WH, Yan ZY, Aaltonen LA, Tomlinson IP: **Germline mutations of the LKB1 (STK11) gene in Peutz-Jeghers patients.** *J Med Genet* 1999, **36**:365-368.
- Westerman AM, Entius MM, Boor PP, Koole R, de Baar E, Offerhaus GJ, Lubinski J, Lindhout D, Halley DJ, De Rooij FW, Wilson JH: **Novel mutations in the LKB1/STK11 gene in Dutch Peutz-Jeghers families.** *Hum Mutat* 1999, **13**:476-481.
- Ylikorkkala A, Avizienyte E, Tomlinson IP, Tiainen M, Roth S, Loukola A, Hemminki A, Johansson M, Sistonen P, Markie D, Neale K, Phillips R, Zuber P, Twama T, Sampson J, Jarvinen H, Makela TP, Aaltonen LA: **Mutations and impaired function of LKB1 in familial and non-familial Peutz-Jeghers syndrome and a sporadic testicular cancer.** *Hum Mol Genet* 1999, **8**:45-51.
- Yoon KA, Ku JL, Choi HS, Heo SC, Jeong SY, Park YJ, Kim NK, Kim JC, Jung PM, Park JG: **Germline mutations of the STK11 gene in Korean Peutz-Jeghers syndrome patients.** *Br J Cancer* 2000, **82**:1403-1406.
- Mehenni H, Blouin JL, Radhakrishna U, Bhardwaj SS, Bhardwaj K, Dixit VB, Richards KF, Bermejo-Fenoll A, Leal AS, Raval RC, Antonarakis SE: **Peutz-Jeghers syndrome: confirmation of linkage to chromosome 19p13.3 and identification of a potential second locus, on 19q13.4.** *Am J Hum Genet* 1997, **61**:1327-1334.
- Tagle DA, Koop BF, Goodman M, Slightom JL, Hess DL, Jones RT: **Embryonic epsilon and gamma globin genes of a prosimian primate (*Galago crassicaudatus*). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints.** *J Mol Biol* 1988, **203**:439-455.
- Ovcharenko I, Nobrega MA, Loots GG, Stubbs L: **ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes.** *Nucleic Acids Res* 2004, **32**:W280-W286.
- Lenhard B, Sandelin A, Mendoza L, Engstrom P, Jareborg N, Wasserman WW: **Identification of conserved regulatory elements by comparative genome analysis.** *J Biol* 2003, **2**:13.
- Loots GG, Ovcharenko I, Pachter L, Dubchak I, Rubin EM: **rVista for comparative sequence-based discovery of functional transcription factor binding sites.** *Genome Res* 2002, **12**:832-839.
- Prestidge DS: **Predicting Pol II promoter sequences using transcription factor binding sites.** *J Mol Biol* 1995, **23**:923-932 [<http://bimas.dcrf.nih.gov/molbio/proscan/>].
- Zhou XP, Waite KA, Pilarski R, Hampel H, Fernandez MJ, Bos C, Dasouki M, Feldman GL, Greenberg LA, Ivanovich J, Matloff E, Patterson A, Pierpont ME, Russo D, Nassif NT, Eng C: **Germline PTEN promoter mutations and deletions in Cowden/Bannayan-Riley-Ruvalcaba syndrome result in aberrant PTEN protein and dysregulation of the phosphoinositide-3-kinase/Akt pathway.** *Am J Hum Genet* 2003, **73**:404-411.
- Green RC, Green AG, Simms M, Pater A, Robb JD, Green JS: **Germline hMLH1 promoter mutation in a Newfoundland HNPCC kindred.** *Clin Genet* 2003, **64**:220-227.
- Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM, Frazer KA: **Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons.** *Science* 2000, **288**:136-140.
- Nobrega MA, Ovcharenko I, Afzal V, Rubin EM: **Scanning human gene deserts for long-range enhancers.** *Science* 2003, **302**:413.
- Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B: **JASPAR: an open-access database for eukaryotic transcription factor binding profiles.** *Nucleic Acids Res* 2004, **32 Database issue**:D91-D94.
- Prestidge DS: **Predicting Pol II promoter sequences using transcription factor binding sites.** *J Mol Biol* 1995, **249**:923-932.
- Lim W, Hearle N, Shah B, Munday V, Hodgson SV, Lucassen A, Eccles D, Talbot I, Neale K, Lim AG, O'Donohue J, Donaldson A, Macdonald RC, Young ID, Robinson MH, Lee PW, Stoodley BJ, Tomlinson I,

Alderson D, Holbrook AG, Vyas S, Swarbrick ET, Lewis AA, Phillips RK, Houlston RS: **Further observations on LKB1/STK11 status and cancer risk in Peutz-Jeghers syndrome.** *Br J Cancer* 2003, **89**:308-313.

26. **Primer3** [[http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)]
27. **NCBI BLAST** [<http://www.ncbi.nlm.nih.gov/BLAST/>]
28. **Chromas** [<http://www.technelysium.com.au>]

Publish with **BioMed Central** and every scientist can read your work free of charge

*"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."*

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

