IRS2 is a candidate driver oncogene on 13q34 in colorectal cancer

Citation for published version:

Digital Object Identifier (DOI):
10.1111/iep.12021

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
International Journal of Experimental Pathology

Publisher Rights Statement:
Available under Open Access

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
IRSI2 is a candidate driver oncogene on 13q34 in colorectal cancer

Elizabeth Day*, George Poulogiannis†,‡, Frank McCaughan*§, Shani Mulholland§, Mark J. Arends§, Ashraf E. K. Ibrahim* and Paul H. Dear*

*MRC Laboratory of Molecular Biology, Cambridge, UK, †Department of Systems Biology, Harvard Medical School, Boston, MA, USA, ‡Department of Medicine-Division of Signal Transduction, Beth Israel Deaconess Medical Center, Boston, MA, USA and §Department of Pathology, University of Cambridge, Cambridge, UK

SUMMARY

Copy number alterations are frequently found in colorectal cancer (CRC), and recurrent gains or losses are likely to correspond to regions harbouring genes that promote or impede carcinogenesis respectively. Gain of chromosome 13q is common in CRC, but, because the region of gain is frequently large, identification of the driver gene(s) has hitherto proved difficult. We used array comparative genomic hybridization to analyse 124 primary CRCs, demonstrating that 13q34 is a region of gain in 35% of CRCs, with focal gains in 4% and amplification in a further 1.6% of cases. To reduce the number of potential driver genes to consider, it was necessary to refine the boundaries of the narrowest copy number changes seen in this series and hence define the minimal copy region (MCR). This was performed using molecular copy-number counting, identifying IRS2 as the only complete gene, and therefore the likely driver oncogene, within the refined MCR. Analysis of available colorectal neoplasia data sets confirmed IRS2 gene gain as a common event. Furthermore, IRS2 protein and mRNA expression in colorectal neoplasia was assessed and was positively correlated with progression from normal through adenoma to carcinoma. In functional in vitro experiments, we demonstrate that deregulated expression of IRS2 activates the oncogenic PI3 kinase pathway and increases cell adhesion, both characteristics of invasive CRC cells. Together, these data identify IRS2 as a likely driver oncogene in the prevalent 13q34 region of gain/amplification and suggest that IRS2 over-expression may provide an additional mechanism of PI3 kinase pathway activation in CRC.

Keywords
colorectal cancer, copy number alterations, insulin receptor substrates, minimal copy regions, molecular copy-number counting, oncogene

Genomic instability is a common feature of the cancer genome and facilitates tumour progression (Lengauer et al. 1998; Hanahan & Weinberg 2000). Colorectal cancers (CRC) display two types of genomic instability: microsatellite instability (MSI) and chromosomal instability (CIN). Microsatellite instability is characterized by the expansion or contraction of microsatellite repeats and CIN by loss of heterozygosity and by aneuploidy. Here, the genomic profiles of predominantly CIN tumours were analysed to identify genes involved in cancer development and progression.

Cancer progression can be viewed as the successive clonal expansion of advantageous genotypes (Nowell 1976). In a dividing population of tumour cells, genomic alterations – in this case copy number alterations (CNAs) – will be subject to a number of selective pressures. In this model, alterations that confer a selective advantage, by promoting cell growth and survival, will be maintained, whereas disadvantageous changes will not. The pattern of recurrent copy number changes within a given cancer type is therefore likely to reflect the balance between advantageous and disadvantageous alterations (Stratton et al. 2009). These recurrent changes will be superimposed upon selectively neutral events peculiar to each cancer genome (Chin et al. 2011).
A number of studies have shown that the more prevalent CNAs often involve genes that are critical to cancer progression: the so-called driver oncogenes (in regions of gain) or tumour suppressors (in regions of loss; Martin et al. 2007; Wood et al. 2007; Stratton et al. 2009). The analysis of CNAs should therefore be able to guide the search for these critical genes. This approach is challenging, though, because such CNAs are often large, encompassing tens or hundreds of genes, the majority of which are presumed to be selectively neutral ‘passengers’ (Tonon et al. 2005). A minority of cancers, however, show small, focal copy number changes which fall within the larger, more commonly affected region of interest (Martin et al. 2007; Leary et al. 2008). Analysis of amplicon structure in these rare samples can more precisely define the ‘minimal affected region’ or ‘minimal copy region’ (MAR or MCR), reducing the number of candidate genes that must be considered (Kendall et al. 2007; McCaughan et al. 2010; Poulogiannis et al. 2010a,b).

Most approaches using the concept of MCRs have relied on a single stage of copy number analysis using array-based platforms (aCGH/SNP; Martin et al. 2007; Chin et al. 2011). However, such methods typically leave some uncertainty as to the precise boundaries of the MCR and therefore fail to fully exploit its potential to pinpoint candidate driver genes. We have previously shown that molecular copy-number counting (MCC) is ideally suited to pinpointing the boundaries of copy number changes, even using very limited DNA samples (Daser et al. 2006; McCaughan et al. 2008). In the current study, we focus on the MCR on chromosome 13q34 in CRC, identified by us and others using array CGH (Lips et al. 2007; Martin et al. 2007). We use MCC to demonstrate that the smallest MCR contains a single whole gene – IRS2. We go on to perform the first comprehensive analysis of IRS2 gene copy number and expression in the different stages of CRC progression: from normal colonic mucosa to adenoma to adenocarcinoma. We also assess some functional properties of deregulated IRS2 expression in vitro. These data strongly implicate IRS2 as a driver oncogene within the recurrent 13q34 gain/amplicon in colorectal adenocarcinoma.

**Material and methods**

**Colorectal clinical samples**

Two independent sample sets were collected from colectomy surgical specimens. The first set of samples (CRC1, $n = 119$) was obtained from 94 patients with invasive colorectal primary carcinoma with or without evidence of metastatic cancer deposits. The CRC1 sample set comprised normal colonic mucosa ($n = 22$), primary adenocarcinoma ($n = 65$) and liver metastatic deposits ($n = 32$). The second set (CRC2, $n = 133$) comprised samples from normal colonic mucosa ($n = 62$), hyperplastic polyps ($n = 7$), adenomatous polyps ($n = 27$) and primary adenocarcinoma samples ($n = 37$) from a set of 44 patients presenting with synchronous adenoma and invasive carcinoma.

**Ethical approval**

Ethical approval for all the work conducted was obtained from Cambridgeshire local research ethics committee (LREC ref. 04/Q0108/125 and 06/Q0108/307).

**Array comparative genome hybridization**

Details of the array platform and statistical analysis have been previously described (Poulogiannis et al. 2010a,b). The array platform has a mean resolution of 0.97 Mb. The reference DNA consisted of a pool of mixed female and male normal leucocyte DNA from 20 unrelated individuals.

**Molecular copy-number counting**

Molecular copy-number counting has been described previously in detail (Daser et al. 2006; McCaughan et al. 2008). Molecular copy-number counting is a digital PCR technique wherein the sample of interest is dispensed at limiting dilution into multiple aliquots, so that each aliquot contains less than one haploid genome’s worth of DNA. A multistep, multiplexed, single-molecule PCR is then used to count the number of aliquots containing each sequence of interest. By using reference sequences, which are assumed to be at normal copy number, the degree of copy number change can be estimated. Details of primer design can be found in Data S1.

**Bioinformatic meta-analysis of IRS2 gene copy number and mRNA expression**

The significance of copy number alteration across 161 colon cancer samples, including 33 CRC cell lines, was determined using the GISTIC algorithm with methods described previously (Beroukhim et al. 2010), using the data deposited at http://www.broadinstitute.org/tumorscape. Microarray expression data from The Cancer Genome Atlas (TCGA) and three previously published data sets (Kaiser et al. 2007; Hong et al. 2010; Skrzypczak et al. 2010) were downloaded from TCGA portal (https://tcga-data.nci.nih.gov/tcga/) and the Oncomine repository (http://www.oncomine.org/). The distributions of log2 median-centred signal intensities were plotted using box plots.

**Tissue microarray**

The construction of the tissue microarray (TMA) has previously been described (Ibrahim et al. 2011). The TMA contains 419 tissue samples from 64 cases: 23% normal colonic mucosa, 7% hyperplastic polyps, 15% adenomas, 34% primary colorectal carcinomas and 21% metastatic CRC samples. The strength of immunohistochemical staining was semiquantitatively scored as negative (0), weakly positive in some cells (1), moderately positive in most cells (2) or strongly positive in most cells (3) by a consultant histopathologist (AI), who rescored a proportion of samples to confirm reproducibility of scoring. The data were analysed...
Hyperplastic polyp 0 15 17 2 1 3
Carcinoma 2 78 36 6 2 5

were defined as DNA copy number increases (three copies or more) affecting DNA copy number cut-offs: Loss

OR, USA) using the cell cycle analysis protocol (Watson

70% ice-cold ethanol and placed at

lies combined; a higher prevalence of 13q34 gains has been number increase (031; Table 1, Figures 1a,b and S1). These focal

amplicon and the case 031 due to the degree of copy num-

region on chromosome 13q34, which was focally gained in five cases. This analysis highlighted a region on chromosome 13q34, which was focally gained in five cases. Two of these cases were then chosen for further analysis: case 026 because it showed the narrowest identified amplicon and the case 031 due to the degree of copy number increase (031; Table 1, Figures 1a,b and S1). These focal gains fell within a larger region that showed DNA copy number gains in 35% (44/124) of tumours in these two ser-

number gains in 35% (44/124) of tumours in these two ser-

five cases. Two of these cases were then chosen for further

region on chromosome 13q34, which was focally gained in five cases. Two of these cases were then chosen for further analysis: case 026 because it showed the narrowest identified amplicon and the case 031 due to the degree of copy number increase (031; Table 1, Figures 1a,b and S1). These focal gains fell within a larger region that showed DNA copy number gains in 35% (44/124) of tumours in these two series combined; a higher prevalence of 13q34 gains has been described in the literature (Lips et al. 2007; Martin et al. 2007, and see Discussion).

The smallest amplicon defined by aCGH (case 026; Figure 1a) in the 13q34 region spanned 2.5 Mb and con-

10 genes. A slightly larger (3.2 Mb) focal gain in this region was also seen in case 031, encompassing 13 genes; this case was also notable in that the region was present at eight copies, relative to normal (diploid) copy number (Figure 1b). The relatively low resolution of the array CGH data made it likely that one or both of these small ampli-

mRNA expression. It is difficult to predict the impact of the copy number changes, making it necessary to assess the correlation of IRS2 gene copy number with expression. To achieve this, a comprehensive meta-analysis of the gene copy number and mRNA expression levels in TCGA colon cancer data set was undertaken. This showed that IRS2 was frequently amplified in CRC and, importantly, that IRS2 gene copy number was significantly positively correlated with IRS2 mRNA expression (Figure 2a–c).

IRS2 expression is not solely a function of gene copy number and is likely to be influenced by a number of feedback loops and other mutations. Therefore to further

Table 1 The copy number of IRS2 in 124 sporadic colorectal cancers and 35 hyperplastic polyps

<table>
<thead>
<tr>
<th>Number of cases</th>
<th>Loss (2 copies)</th>
<th>3 copies</th>
<th>4–5 copies</th>
<th>Amplification (&gt;5 copies)</th>
<th>Focal gains (&lt;5 Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carcinoma</td>
<td>2</td>
<td>78</td>
<td>36</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Hyperplastic polyp</td>
<td>0</td>
<td>15</td>
<td>17</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

DNA copy number cut-offs: Loss – <1.75, Normal – 1.75–2.5, three copies – 2.5–3.5, 4–5 copies – 3.5–5.5, Amplification – >5.5. Focal gains were defined as DNA copy number increases (three copies or more) affecting <5 Mb of DNA and encompassing IRS2.
explore the transcriptional status of IRS2 in CRC, the distributions of IRS2 mRNA levels in two independent microarray data sets were plotted to show that IRS2 expression is significantly higher in colon carcinomas than in normal mucosa samples (Figure 2d,e; Hong et al. 2010).

The copy numbers and/or expression levels of driver oncogenes are likely to correlate with disease stage (Bertucci et al. 2004; Garnis et al. 2004). To assess this, bioinformatic analysis and TMA analysis of IRS2 expression during tumour progression from normal through adenoma to carcinoma were carried out. A single expression data set containing samples from all three stages of colorectal carcinoma formation was available (Skrzypczak et al. 2010). In this data set, IRS2 mRNA levels showed significant increases through the normal-to-adenoma-to-carcinoma sequence (Figure 2e).

Immunohistochemical analysis of IRS2 protein expression. The TMA-IHC analysis showed that all pathological groups differed significantly from normal colon mucosa with respect to IRS2 expression and demonstrated an increase in IRS2 expression with progression through the stages of colorectal carcinoma formation (Figure 3a,b).

Tissue sections from cases 026 and 031 were also stained and showed an increase in IRS2 expression in the carcinoma in comparison with the adjacent normal tissue, consistent with the specific amplification of the IRS2 gene in these cases (Figure 3c).

Functional analysis of IRS2 over-expression in vitro. IRS proteins are adaptors in the insulin and IGF signalling cascades (Dearth et al. 2007). These cascades use the oncogenic PI3K and MAPK pathways and control many characteristics important to tumour progression including cell proliferation, adhesion and migration (Figure 4a; Yuan & Cantley 2008). Differential adaptor protein activation is likely to be one mechanism by which a single extracellular signal can have distinct phenotypic effects (Csizsar 2006). In breast cancer, IRS2 is implicated in metastasis, but not in growth and proliferation (Jackson et al. 2001; Zhang et al. 2004; Byron et al. 2006; Gibson et al. 2007).

To begin to assess some of the functional implications of IRS2 over-expression in CRC, three functional in vitro assays were used. The CRC cell line SW480, derived from a
primary Dukes’ stage B colon carcinoma, was chosen because endogenous levels of IRS2 expression were undetectable. Transient over-expression was achieved through transfection with pcDNA.IRS2-HA (Jackson et al. 2001) and expression of IRS2 confirmed by immunoblotting (Figure 4b).

PI3 kinase pathway activation. IRS2 over-expression led to PI3K pathway activation as evidenced by phosphorylation of AKT (pAKT) in the absence of upstream activation (e.g. by IGF-1; Figure 4b, Lane 1). It is challenging to remove all traces of upstream activators in cultured cells; however, the cells grown in serum-free media demonstrate low levels of pAKT (Figure 4b, Lane 2), thus suggesting minimal upstream activation and therefore ligand-independent activation of the PI3K pathway by IRS2 over-expression. Treatment of the mock-transfected cells with IGF-1 alone led to the phosphorylation of AKT. These cells have no detectable IRS2 expression, which therefore suggests signalling is taking place via another adaptor, possibly IRS1.

MAPK pathway activation was also assessed in a similar manner, and results suggested constitutive activation of this pathway in the cell line independent of IGF-1, epidermal growth factor and IRS2 expression (data not shown).

Cell cycle and adhesion assays. To assess whether IRS2 over-expression influenced cell cycle and adhesion, in vitro assays were set up using the transiently transfected SW480 cells. Cell proliferation was assessed by flow cytometry and cell cycle analysis. Adhesion was determined by the ability of the cells to adhere to a collagen-coated plate (Zhang et al. 2005). Migration studies were also carried out using SW480 cells but yielded inconclusive results (data not shown).
shown). Consistent with prior data in breast cancer, the assays demonstrated that IRS2 over-expression does not modify cell proliferation but does increase cell adhesion, whereas IGF-1 stimulation of mock-transfected cells increased both proliferation and adhesion (Figure 4c,d). Prior to these assays, the cells were serum-starved to remove potential upstream activators of the PI3K pathway, suggesting that increased adhesion in IRS2-over-expressing cells is ligand independent.

**Discussion**

Minimal copy region analysis pinpoints IRS2 as a candidate driver gene in CRC.

Minimal copy regions are rare sites of focal DNA copy number changes seen within larger regions that more frequently show such changes. It is likely that MCRs contain the genes driving the larger, more prevalent copy number
changes, and data presented here demonstrate how MCRs can be used to identify candidate driver genes. This analysis focuses on an MCR identified on chromosome 13q34 in CRC.

The 13q34 region, within which IRS2 is found, frequently shows gains in CRC, with a prevalence of up to 85% reported in the literature (Lips et al. 2007; Martin et al. 2007). Our aCGH analysis shows a lower prevalence of this gain, with IRS2 copy number increased in 35% of CRCs analysed, with 4% of cases showing focal gains of <5 Mb. A further 1.6% of cases contained amplifications of this region, which is in keeping with other reports (Parsons et al. 2005; Table 1). Despite their lower prevalence, focal changes and amplifications are a strong indication of a functional copy number change (Poulogiannis et al. 2010a; Chin et al. 2011).

The smallest focal change in the region identified by aCGH still contained 10 genes, making it difficult to propose a candidate driver gene. It was therefore necessary to refine the boundaries of the amplicon. Molecular copy-number counting is ideally suited to this application, both because it can be applied iteratively to refine amplicon boundaries and because it requires very little sample DNA. Through this refinement, IRS2 was pinpointed as a potential driver oncogene in CRC.

Copy number and mRNA and protein expression analysis support IRS2 as a candidate oncogene

To assess IRS2 as a candidate driver gene, its copy number and expression were assessed across separate series of tumours. Our meta-analysis using TCGA demonstrates IRS2 gains in 52.2% of CRCs, with 11.8% showing high-level amplification in this region (Figure 2a). This supports previous studies, which have proposed IRS2 as a candidate based on copy number analysis (Lips et al. 2007; Martin et al. 2007; Beroukhim et al. 2010). However, none of these studies correlated these observations with an assessment of IRS2 expression in CRC. Previous studies characterizing IRS2 expression across human cancers, including CRC, have failed to show a clear correlation between expression and the progression from normal tissue to carcinoma (Mardilovich et al. 2009; Zha et al. 2009). In contrast, our analysis correlates IRS2 copy number with mRNA expression (Figure 2c) and demonstrates that expression at both the mRNA and protein level increases with progression (Figures 2e and 3a). This suggests that expression of IRS2 may influence the progression of CRC and supports data from Szabolcs et al. (2009) showing that IRS2 inactivation suppressed tumour progression in Pten+/− mice.

IRS2 gene copy number and expression in hyperplastic polyps do not correlate, suggesting feedback mechanisms may be important

The analysis of hyperplastic polyps in this study shows that 57% of cases demonstrate a gain in IRS2, compared to 35% of carcinomas. This is unexpected given that these lesions are not considered immediate precursors of malignancy, although a very small proportion may progress along the serrated/MSI pathway of colorectal neoplasia. However, the copy number does not correlate with IRS2 protein levels seen in this type of lesion – no hyperplastic polyps demonstrated significant over-expression (a score of 3 on the TMA-IHC analysis; Figure 3a, Table 1). It is therefore possible that feedback mechanisms, controlling IRS2 protein levels, are still functioning in these lesions, thus preventing an increase in IRS2 expression. IRS2 protein levels are controlled by negative feedback from downstream components of the PI3K pathway, specifically via mTOR/S6 kinase (Manning 2004). This is of particular importance in the context of cancer therapy, as feedback will be lost in mTOR-targeted chemotherapy (McCampbell et al. 2010). Therefore, in the light of data presented here and by others showing IGF-1-independent phosphorylation of AKT by IRS2 over-expression, mTOR inhibition has the potential to activate components of the oncogenic PI3 pathway (Dearth et al. 2006).

IRS2 over-expression in vitro results in AKT phosphorylation independent of IGF-1 stimulation and may confer an advantageous phenotype to CRC

The insulin receptor substrate (IRS) family of proteins are adaptors linking upstream activators, canonically insulin and IGF-1, to multiple downstream effectors with roles in normal growth, metabolism and differentiation (Figure 4a; Dearth et al. 2007). The role of IRS proteins in CRC is unclear, but a number of studies in breast cancer have suggested a role in proliferation and metastasis of cancer cells (Zhang et al. 2004; Byron et al. 2006; Gibson et al. 2007; Mardilovich et al. 2009).

We have shown that IRS2 over-expression in the absence of an upstream activator, in the CRC cell line SW480, leads to AKT phosphorylation (Figure 4b). This observation is supported by previous studies of IRS2 over-expression in both breast cancer and mouse fibroblasts showing AKT phosphorylation independent of IGF-1 (Hennige et al. 2000; Mardilovich & Shaw 2009). This is significant as inappropriate AKT activation is oncogenic and a potential therapeutic target (Hsieh et al. 2011). IRS2 over-expression also increases CRC cell adhesion to a similar extent to IGF-1 stimulation, but has no effect on cell cycle in this cell line; this is in agreement with observations in breast cancer and supports a role of IRS2 in adhesion, but not in proliferation (Jackson et al. 2001; Gibson et al. 2007). Changes to adhesion, both increases and decreases, are important properties of metastasizing cancer cells and are involved in invasion, migration, arrest within the circulation and distant ‘seeding’ of a tumour (Hewitt et al. 2000; Schluter et al. 2006). It is possible that proliferation stimulated by IGF-1 is mediated by IRS1 in the SW480 cell line. Although the physiological relevance of the levels of expression seen in these assays remains to be proven, the results indicate that IRS2
over-expression has the potential to confer an advantageous phenotype upon cancer cells.

**IRS2 and PI3 kinase pathway in CRC**

The PI3K pathway is frequently dysregulated in CRC, and this can be a consequence of alterations at any level of the pathway – from receptor mutations to the deletion of negative regulators such as PTEN and mutations and subsequent over-expression of effectors such as AKT (Parsons et al. 2005). Recently, TCGA Network demonstrated in CRC that high levels of IRS2 expression are mutually exclusive with IGF2 over-expression and with other mutations in the PI3K pathway (TCGA 2012). This suggests that IRS2 over-expression is one mechanism by which the PI3K pathway may be dysregulated in CRC, and our data support this conclusion.

In summary, we have used high-resolution analysis of the 13q34 amplicon in two CRC cases to pinpoint the gene IRS2 as a potential driver oncogene in this amplicon. We also show the first focused analysis of IRS2 protein and mRNA levels in colorectal neoplasms, demonstrating its strongly positive correlation with CRC progression. Preliminary functional in vitro analysis suggests that, as has been seen in other cancer types, IRS2 over-expression confers an advantageous phenotype on CRC cells. This highlights the need for further work to address the circumstances in which IRS2 contributes to CRC progression. Together, these data indicate IRS2 as a strong candidate oncogene and support recent suggestions that the IGF1R–IRS2–PI3K axis may be an important therapeutic target in a subset of CRC (TCGA 2012).

**Funding source**

This study was supported by the Medical Research Council and Cancer Research UK. AEKI is Cancer Research UK Clinician Scientist Fellow (WT097143MA). FM is a Wellcome Trust Intermediate Clinical Fellow. GP is a Pfizer fellow of the Life Sciences Research Foundation. The Human Research Tissue Bank is supported by the NIH’s Cambridge Biomedical Research Centre.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**References**


Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Regions of DNA copy number alteration identified along chromosome 13 by the aCGH analysis of the second CRC series of 50 CRCs, showing the percentage of samples with gain of copy (copy number > 1.25, where normal copy = 1) against chromosomal position (NCBI36).

Figure S2. Confocal microscopy of IRS2-positive and negative cells.

Appendix S1. For each marker, the sequences of the forward (Fex) and reverse (Rvs) primers are given (5′–3′).

Appendix S2. Adhesion analysis.

Table S1. Antibodies.

Data S1. Functional analysis.