Lehrstuhl für Biotechnologie der Nutztiere

Analysis and optimisation of a porcine model for colorectal cancer

Carolin Perleberg

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende(r):	Prof. Dr. Harald Luksch
Prüfer der Dissertation:	1. Prof. Angelika Schnieke, Ph.D.
	2. apl. Prof. Dr. Dieter Saur

Die Dissertation wurde am .....17.01.2019.....bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am .....06.06.2019..... angenommen.

Abstract

Colorectal cancer (CRC) is the fourth most common cancer worldwide in both sexes and as it is often diagnosed late it has a high mortality. CRC is a very heterogeneous group of malignancies where a multitude of molecular changes such as chromosomal instability, microsatellite instability, mutations, deletions of chromosomal regions, aberrant epigenetic modifications and dysregulated microRNAs (miRNAs) disrupt the signalling pathways WNT, Ras/ MAPK, PI3K, TGFβ/SMAD and p53. Animal models for CRC are essential to better understand the diseases and to identify novel treatment opportunities. Although mice are frequently used to model human disease, they do not replicate key aspects of human CRC pathology. The chair of livestock biotechnology at the University of Munich has generated pigs that carry a translational stop signal at codon 1311 in porcine APC that is orthologous to the human APC^{1309} mutation that is most frequently diagnosed in both CRC and familial adenomatous polyposis (FAP) a hereditary autosomal dominant disorder causing development of multiple polyps during puberty, resulting in a strong CRC risk. APC^{1309} germline mutation in human FAP is associated with very severe polyposis. The APC^{1311} pigs replicate hallmarks of human FAP and CRC including adenomatous polyps in the colorectum with loss of APC heterozygosity, β-catenin accumulation, upregulation of c-MYC, MAPK pathway activation and progression to carcinoma *in situ* and phenotypic variation in polyposis severity ranging from ≥100 (high polyp animals (HP)) to only 1-10 polyps (low polyp animals (LP)) in the distal colorectum (last 40 cm).

The aim of this project was to identify elements in the genetic background such as modifier genes, single-nucleotide polymorphisms, gene sets and miRNAs that may contribute to severe polyposis using next generation mRNA and miRNA sequencing. The gene *CYP7A1*, miRNAs miR-215 and 194b-5p and gene sets associated with oestrogen response were found to be highly expressed and targeted by miRNAs differentially expressed in animals with severe polyposis (HP). *Cis*-regulation of *CYP7A1* was analysed for CpG island methylation and SNPs in promoter region (>5000 bp prior to the ATG). One CpG site that is included in the binding sequence of STAT3, a known tumour-promoting factor, was significantly lower methylated in HP animals, likely allowing better binding of STAT3. Increased CYP7A1 expression could be traced to stromal cells of the normal mucosa rather than crypt cells.

Analysis of tumour progression from low grade (LG) to high grade (HG) intraepithelial neoplasia (IEN) revealed an increase in the expression of immune associated genes originating from HG tumour stroma, the genes *PLXND1* and *GBP6* in laser microdissected HG-IEN and miRNAs let-7e, miR-146a-5p, 146b, 183, 196a in HG-IEN bulk samples. High expression of these genes and miRNAs has been associated with tumour-promoting capacities in humans. Gene set enrichment showed gene sets such

as MYC targets and cell cycle related gene sets enriched that were also found enriched in human CRC, replicating broad molecular pathways of human CRC in the *APC*^{1311} pig model.

CRC in the *APC*^{1311} pigs progressed very slowly, just like in humans. Acceleration and increase of tumour progression of the model was aimed by introduction of oncogenic mutations. Therefore, the generation of *APC*^{1311} pigs with ubiquitous *Cas9* expression in the *ROSA26* locus was aimed to allow local genome editing with high efficiency of introducing sequential oncogenic mutations by *in vivo* administration of guide RNAs (gRNAs) via *in vivo* electroporation or adeno-associated viral vectors. Primary cells from *APC*^{1311} pigs were *Cas9* (isolated from *Streptococcus pyogenes*) targeted into the porcine *ROSA26* locus via homologous recombination using promoter trap strategy. Correctly targeted clones were analysed on genomic, RNA and protein level, validating correct targeting, expression and functional nuclease activity. However, the cells failed to generate viable offspring when used as nuclear donors for somatic cell nuclear transfer.

Zusammenfassung

Kolorektaler Krebs ist die vierthäufigste Krebserkrankung weltweit und weist aufgrund später Diagnosen eine hohe Sterblichkeitsrate auf. Die Betrachtung der molekularen Karzinogenese kolorektalen Krebses zeigt, dass es sich um eine sehr heterogene Gruppe von Krebserkrankungen handelt. Verschiedenste molekulare Veränderungen wie chromosomale Instabilität, Mikrosatelliteninstabilität, Mutationen, Deletionen chromosomaler Abschnitte, anormale epigenetische Modifikationen und dysregulierte microRNAs (miRNAs) stören die Signalwege WNT, Ras/MAPK, PI3K, TGFβ/SMAD und p53. Tiermodelle für kolorektalen Krebs sind essenziell um die Erkrankung besser zu verstehen und neue Behandlungsmöglichkeiten identifizieren zu können. Obwohl hierbei sehr häufig auf Mäuse zurückgegriffen wird, um humane Erkrankungen zu modellieren, können sie Schlüsselaspekte humaner kolorektaler Krebspathologie nicht replizieren. Der Lehrstuhl für Biotechnologie der Nutztiere an der Technischen Universität München hat Schweine mit einem translationalen Stopsignal an Kodon 1311 des porzinen *APC* gens generiert, welches ortholog zur humanen APC^{1309} Mutation ist. APC^{1309} ist die am häufigsten diagnostizierte Mutation in sporadischem kolorektalen Krebs sowie der erblich autosomal dominanten Erkrankung familiäre adenomatöse Polypose (FAP) bei welcher sich unzählige Polypen bereits während der Pubertät im Darm entwickeln und so ein enormes Krebsrisiko darstellen. Die Keimbahnmutation APC^{1309} ist zudem mit einer sehr schweren Polypose der FAP Patienten assoziiert. Die APC^{1311} Schweine replizieren Schlüsselaspekte humaner FAP und kolorektalen Krebs wie adenomatöse Polypen im Kolorektum mit Verlust des APC wildtyp Alleles, β-catenin Akkumulierung, Hochregulierung von c-MYC, MAPK Signalwegaktivierung sowie die Progression zum *Carcinoma in situ* und phentopische Variation der Polypose von ≥100 Polypen (high polyp animals (HP)) zu 1-10 Polypen (low polyp animals (LP)) im distalen Kolrektum (letzten 40 cm).

Ziel dieses Projektes war die Identifizierung von Elementen im genetischen Hintergrund wie Modifier Gene, einzelner Nukleotidpolymorphismen (SNPs), Gen-Sets und miRNAs, welche zur Empfänglichkeit einer schwerwiegenden Polypose beitragen, mittels “next generation” RNA und miRNA Sequenzierung. Das Gen *CYP7A1*, miRNAs miR-215 und 194b-5p sowie Gen-Sets assoziiert mit Östrogen-Antwort wurden hoch exprimiert oder von dysregulierten miRNAs anvisiert in Tieren mit einer sehr schwerwiegenden Polypose (HP) nachgewiesen. *Cis*-Regulation von *CYP7A1* wurde mittels CpG-Insel Methylierungsevaluation und SNP Detektion >5000 bp vor dem ATG analysiert. Eine CpG Position welche sich in der Bindungssequenz von STAT3, ein bekannter tumorfördernder Faktor, befindet, war signifikant weniger methyliert in HP Tieren, welches eine mögliche STAT3 Bindung

begünstigen kann. Darüber hinaus konnte die erhöhte *CYP7A1* Expression Stromazellen der normalen Mukosa statt Kryptzellen zugeordnet werden.

Analysen der Tumorprogression von niedriggradigen intraepithelialen Neoplasien (LG-IEN) zu hochgradigen intraepithelialen Neoplasien (HG-IEN) offenbarten eine erhöhte Expression immunassoziierter Gene, welche dem Tumorstroma zugeordnet werden konnten, sowie der Gene *PLXND1* und *GBP6* in laser mikrodissektiertem intraepithelialen Neoplasien und miRNAs let-7e, miR-146a-5p, 146b, 183, 196a in Bulkproben. Im Menschen wurden hohe Expressionen dieser Gene und miRNAs tumorfördernden Fähigkeiten zugeordnet. “Gene set enrichment” Analysen zeigten Anreicherungen von MYC-Zielgenen und Zellzyklus assoziierte Gen-Sets, wie sie auch bei humanen kolorektalem Krebs gefunden wurden. Somit repliziert das Schweinemodell wesentlich molekulare Signalwege humanen kolorektalen Krebses.

Die Progression zu kolorektalem Krebs in den *APC*^{1311} Schweinen verlief, wie auch im Menschen, sehr langsam. Um eine Tumorbildung zu beschleunigen, sollte durch das Einfügen weiterer onkogener Mutationen erzielt werden. Hierfür, sollte die Generation eines APC^{1311} Schweins mit ubiquitärer *Cas9* Expression im *ROSA26* Lokus generiert werden. In Vivo Verabreichung von guide RNAs (gRNA) mittels in vivo Elektroporation oder adeno-assoziierter viraler Vektoren sollte hierbei das sequentielle Einfügen onkogener Mutationen ermöglichen. *Cas9* (isoliert aus *Streptococcus pyogenes*) wurde in den porzinen *ROSA26* Lokus primärer Nierenzellen eines *APC*^{1311} Schweins mittels homologer Rekombination und „promoter trap” Strategie eingebracht. Die korrekte Position der *Cas9* Insertion sowie Expression und Nukleaseaktivität wurden auf genomischer, RNA- und Protein-Ebene erfolgreich transfizierter Klone validiert. Als Nukleus-Donator für Kerntransfer somatischer Zellen konnten die Zellklone jedoch keine lebenden Nachkommen generieren.

Content

1. Introduction

1.1 Colorectal cancer

Cancers are the second leading cause of death worldwide, with more than 14 million new cases and 8 million mortalities in 2012 (Ferlay et al, 2015).

Colorectal cancer (CRC) is the fourth most common cancer in both sexes worldwide with 1.36 million cases and 51 % mortalities in 2012 (Ferlay et al, 2015). Despite increasing efforts in research and healthcare to counter act cancer, the number of CRC cases is estimated to grow over time (Ferlay et al, 2010; Ferlay et al, 2015). The increasing number of CRC cases is associated with the increased age of people (World Health Statistics 2017: Monitoring health for the SDGs) but also with the expansion of western lifestyle and diet. This was evidenced by 51,508 more CRC incidents in first-world regions compared to those in developing areas in 2012 (Ferlay et al, 2015). High risk factors for CRC are high consumption of red meat, alcohol and tobacco, obesity and a lack of exercise which represent aspects of a western lifestyle (Jemal et al, 2011).

70-80 % of all CRC cases are sporadic while 20-30 % are caused by hereditary mutations (Whiffin et al, 2014). Both sporadic and hereditary CRC arise from mutations causing histological changes of the colorectal epithelium. Through early mutations of tumour suppressor genes, the stem cells or progenitor cells in the colorectal crypts give rise to aberrant crypt foci that progress to conventional adenomatous or serrated adenomas in colon and rectum. Only 10-15 % of conventional adenomas will develop over a span of a decade to CRC (Fearon, 2011; Yang et al, 2004). No matter how slow this process is, CRC accounted for 8 % of all cancer deaths worldwide in 2012, making it the fourth most deathly cancer (Ferlay et al, 2015). This is in large part due to late diagnosis of the disease as early stage CRC is often asymptomatic.

The survival rate of CRC is strongly dependent on the stage of the disease at the time of diagnosis (Tomlinson et al, 2012). According to the American cancer society CRC diagnosed in stage 1 has a 5-year survival rate of 89.9 % while later stages, with or without lymph node metastasis, range from 71.3 % to 13.9 % respectively (US national cancer institute). However, only about 12 % of all CRC cases are diagnosed in stage 1 (Guinney et al, 2015). This emphasises the importance of early diagnostic methods and better implementation of screening methods.

1.2 Molecular pathology of colorectal cancer

The more research efforts were addressed to CRC, the more complex pathological pathways were uncovered, identifying CRC as a heterogeneous group of malignancies. The two best characterised CRC carcinogenesis pathways are the classic adenoma-carcinoma sequence (Fearon & Vogelstein, 1990) and the microsatellite instability (MSI) pathway. The prognosis and strategies for treating CRC are dependent on the molecular mechanisms underlying the disease (Huth et al, 2014).

Genomic instability is a crucial feature of many cancer types. In CRC at least three main mechanisms are well known that cause genomic instability. One mechanism is microsatellite instability (MSI), which is the driving force of the MSI pathway. The second mechanism , aberrant DNA methylation is also strongly associated with the MSI pathway (Guinney et al, 2015; Hinoue et al, 2012). The third mechanism, chromosomal instability, is the driving force of the classic adenoma-carcinoma sequence. So far the MSI pathway and adenoma-carcinoma-sequence have been mainly characterised by the acquisition of specific mutations modulating the signalling pathways WNT, Ras/MAPK, PI3K, TGFβ/SMAD and TP53. Integrative analysis revealed that 94 % of all CRC carry an alteration (including mutations, amplification and deletions) in the WNT, 61 % in the MAPK, 50 % in the PI3K, 36 % in the TGFβ and 60 % in the p53 signalling pathways (Guinney et al, 2015). However, the improved insight into epigenetic modifications with their genomic instability promoting capacity, microRNA (miRNA) interference and their influence on gene expression will lead to an increase of these numbers and expansion of the CRC carcinogenesis pathways.

1.2.1 Epigenetic modifications in cancer

The field of epigenetics is concerned with the study of the regulation of DNA-templated processes (transcription, repair and replication of DNA) by chromatin based events without changes to the DNA sequence (Dawson & Kouzarides, 2012).

Chromatin is composed of nucleosomes that consist of an octamer containing two of each H2A, H2B, H3 and H4 histones with 147 bp DNA wrapped around (Luger et al, 1997). Tightly packed chromatin, rendering the DNA inaccessible for transcription factors (TFs), is termed heterochromatin and mainly contains inactive genes. Euchromatin on the other hand, is packed very loosely allowing the TFs to access the DNA and thus contains most of the active genes. The chromatin conformation can be modulated by two different mechanisms, DNA and histone modifications. So far, four different types of DNA and 14 different classes of histone modifications have been identified (Bannister & Kouzarides, 2011; Dawson & Kouzarides, 2012; Pfaffeneder et al, 2011; Tahiliani et al, 2009; Tan et al, 2011).

Methylation and acetylation at lysine and arginine residues of histones 3 and 4 are the best understood histone modifications. Histone acetylation neutralises the positive charge of the bound lysine residue at the amino terminal of the histone tail, reducing the affinity between negatively charged DNA and the histone complex (Luger et al, 1997). This results in accessible euchromatin upon acetylation by histone acetyltransferases (HATs). Histone deacetylases (HDACs) remove the acetyl groups from lysine residues of the histones, rendering the histone tail positively charged with higher affinity towards the negatively charged DNA, resulting in heterochromatin.

Histone methylation is more complex. Methylation occurs on both arginine and lysine residues governed by methyltransferases and demethylases. Methylation of arginine can be asymmetrical or symmetrical with one or two methyl groups (Blanc & Richard, 2017). Methylation of lysine residues with one, two or three methyl groups can have repressing or enhancing effects on gene expression depending on the position of the lysine residue, the histone and the amount of methyl groups added. Even though the system is bivalent, active euchromatin is associated with H3K9me, H3K27me, H4K20me H3K4me/2/3, H3K36me/3 and H3K79me/2/3. While heterochromatin is associated with H3K9me2/3, H3K27me2/3, H3K4me3 and H4K20me3 (Black et al, 2012).

DNA methylation was the first modification discovered and is therefore the best characterised. In higher eukaryotes, three types of DNA methyl transferases (DNMTs) target the DNA directly and catalyse the conjugation of a methyl group to the 5C position of a CpG dinucleotide, generating a 5-methylcytosine. DNMT1 is a maintenance methyltransferase catalysing the methylation of hemi methylated DNA after replication by copying the methylation patterns from the parental to the newly synthesised DNA strand (Li et al, 1992). De novo methylation during embryogenesis, for instance, is performed by de novo methyltransferases DNMT3a and DNMT3b (Okano et al, 1999).

Aberrant DNA methylation in the shape of global hypomethylation and local hypermethylation has been consistently found in all cancers including CRC (Feinberg & Vogelstein, 1983a; Gama-Sosa et al, 1983). Global hypomethylation can be caused by loss of function of the maintenance methyltransferase DNMT1 or by the TET mediated oxidation of the 5-methylcytosine and subsequent base excision repair (Cortellino et al, 2011; Karpf & Matsui, 2005; Li et al, 1992; Tahiliani et al, 2009). Global hypomethylation leads to loss of methylation of repetitive sequences that are mainly composed of transposable elements (TE) that make up about 50 % of the human genome (Estecio et al, 2007; Szpakowski et al, 2009). The group of TEs contains long terminal repeats (LTRs), long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs) including Alu sequences (Anwar et al, 2017) (Kapitonov & Jurka, 2008). In normal cells TEs are silenced by DNA methylation and histone modifications. By DNA hypomethylation, TEs are activated and can translocate within the genome, rendering them highly mutagenic (Criscione et al, 2014; Lee et al, 2012). This causes genomic

instability, gene disruption of tumour suppressor genes such as APC, oncogenic activation and even chromosomal breakage (Daskalos et al, 2009; Feinberg & Vogelstein, 1983b; Hur et al, 2014; Lee et al, 2012; Liu et al, 1997; Miki et al, 1992; Wolff et al, 2010).

Local hypermethylation of CpG islands (CGI) in promoter regions termed the CGI methylator phenotype (CIMP) is closely associated with the silencing of tumour suppressor genes (Esteller et al, 2000; Herman et al, 1998; Tse et al, 2017). The silencing can be caused by steric blockage of the binding of TFs when the consensus sequence of a TF is methylated (Klose & Bird, 2006) or by methylation-mediated recruitment of methyl-CpG-binding proteins (MBPs) which attract corepressor molecules to silence the expression and modulate chromatin conformations (Deaton & Bird, 2011). DNA methylation-mediated silencing in CRC of genes from all five crucial signalling pathways WNT, MAPK, p53, PI3K and TGFβ/SMAD have been reported.

1.2.2 MicroRNAs in cancer

The human haploid genome of the size of ~3 x 10^9 bp contains about 21 000 protein-coding genes that make up only 1.5 % of our entire genome. However, almost the entire genome is transcribed, leaving a vast amount of RNA that is not translated. These non-coding RNAs (ncRNAs) are subdivided into small ncRNAs, below 200 bp in size, and long ncRNAs. MicroRNAs (miRNA) are part of the small ncRNAs (Dawson & Kouzarides, 2012) and are evolutionarily highly conserved 21-25 bp RNA sequences. 25 % of all miRNA genes do not have their own regulatory elements including a promoter but are positioned in the intron region of other genes, dependent on host gene expression (Lin et al, 2006; Wu et al, 2011). MiRNA genes transcribed (Hayes et al, 2014) (pre-miRNAs) (Shirafkan et al, 2018) into ~22 nucleotide long mature miRNAs (Ha & Kim, 2014). With assembly of the RNA-induced silencing complex (RISC) the mature miRNA binds to complementary sequences mainly of the 3' untranslated region (UTR) of its mRNA targets (Easow et al, 2007; Fabian et al, 2010; Orom et al, 2008; Shirafkan et al, 2018). This will induce translational repression reducing protein numbers or mRNA degradation reducing both mRNA and subsequently protein numbers (Shirafkan et al, 2018). Although mRNA degradation is believed to mainly contribute to reduced protein expression (75-84%) (Guo et al, 2010; Hendrickson et al, 2009). Just like genes, mi-RNAs have been found dysregulated in cancer with oncogenic (onco-miR) or tumour suppressive (ts-miRs) effects and in some cases even both.

The reasons for miRNA dysregulation are still poorly understood but more than 50 % of miRNAs are localised in so called fragile regions of the genome, where genomic instability will lead to loss or amplification of miRNAs (Calin et al, 2004). Aberrant epigenetic modifications including DNA methylation of the genome also influence the expression of miRNA genes, resulting in aberrant miRNA

expression (Bandres et al, 2009; Toyota et al, 2008; Vogt et al, 2011). Therefore, dysregulated miRNA regulation can be observed in both the MSI and the classic adenoma carcinoma pathway modulating all 5 pathways WNT, MAPK, p53, PI3K and TGFβ/SMAD (see 1.2.3 and 1.2.4).

1.2.3 Microsatellite instability pathway

The MSI pathway is responsible for about 13-16 % of all human sporadic CRCs and mostly associated with lesions in the right colon (Bettington et al, 2013; Guinney et al, 2015). It is characterised by a defective DNA mismatch repair (MMR) system frequently caused by promotor methylation-mediated silencing or mutations of the gene *MLH1* (Deng et al, 1999; Herman et al, 1998; Kane et al, 1997; Veigl et al, 1998). However, there are also MSI CRC cases, where other members of the MMR system, such as *MSH2*, *MSH6*, *MGMT* and *MBD4* (Esteller et al, 1999; Lahtz & Pfeifer, 2011) are aberrantly methylated and silenced, or downregulated by overexpressed miR-155 and miR-21 that target *MLH1*, *MSH2* and *MSH6* (Valeri et al, 2010; Volinia et al, 2006). The defective MMR system is incapable of repairing errors made during DNA replication leading to the accumulation of insertions, deletions and substitutions in stretches of short tandem repeats. These 1-6 bp repeats, called microsatellites, can be found throughout the entire genome of most mammals in both coding and non-coding regions (de la Chapelle & Hampel, 2010). Therefore, microsatellite mutations can affect gene functions. Mutations of other MMR-related genes *MSH3* and *MSH6* (Duval & Hamelin, 2002; Malkhosyan et al, 1996), the genes *APC* (Guinney et al, 2015), *AXIN2* (Liu et al, 2000; Shimizu et al, 2002), $BRAF^{V600E}$ (oncogenic mutation) (Guinney et al, 2015), *PTEN* (Guanti et al, 2000), *TGFβR2* (Markowitz et al, 1995), *ACVR2* (Jung et al, 2004) and proapoptotic factors *BCL10* (Yamamoto et al, 2000) and *BAX* (Rampino et al, 1997) are also often diagnosed in MSI tumours, activating proliferation via the WNT and MAPK pathways and inactivating the growth inhibition via TGFβ signalling and blocking apoptosis. Additionally, aberrant miRNA expression of miR-31, 223 and 26b was detected (Earle et al, 2010). MiR-31 has been shown to target AXIN1, activating the WNT pathway (Slaby et al, 2007). The functions of miR-223 and 26b have not specifically been determined in the MSI pathway, but miR-223 has been associated with an oncogenic effect on the RAS pathways by targeting RASA1 (Sun et al, 2015). MiR-26b has been associated with oncogenic effects on metastasis targeting both PTEN and WNT5A (Fan et al, 2018).

As the MSI pathway is highly associated with CIMP high phenotype (Guinney et al, 2015; Hinoue et al, 2012), methylation-mediated silencing of WNT pathway components such as *APC* (Esteller et al, 2000; Segditsas et al, 2008), negative regulators *SFRP1*, *2*, *4* and *5* (Suzuki et al, 2004), *APCDD1*, *DKK1*, *AXIN2* (de Sousa et al, 2011),WNT targets *LGR5* and *ASCL2* and RAS/MAPK components *RASSF1* and *2* (80%

of CRC) (Akino et al, 2005; Fernandes et al, 2013; Harada et al, 2007), PI3K pathway elements *PPP2R2B* and *PTEN* (in 20% of CRCs) (Goel et al, 2004; Tan et al, 2010), TGFβ factors *TSP1* and *RUNX3* (20% of CRCs) (Imamura et al, 2005; Rojas et al, 2008), p53 pathway components *P14* (20% of CRCs) and *IGFBP7* (Nyiraneza et al, 2012; Shen et al, 2003; Suzuki et al, 2010), genes involved in cell cycle control *P16* (Shima et al, 2011)and TFs important for differentiation *CDX1*, *CDX2*, *GATA4* and *GATA5* (Baba et al, 2009; Hellebrekers et al, 2009; Hryniuk et al, 2014; Suh et al, 2002; Wong et al, 2004) are possible.

1.2.4 Classic adenoma-carcinoma sequence

The mechanism of chromosomal instability (CIN) is not quite understood, although it has been found to behave like a dominant trait causing deletions, inversion, translocations and duplications (Lengauer et al, 1997). Defects in the mitotic spindle formation, alignment and segregation of chromosomes during mitosis may contribute to CIN (Barber et al, 2008; Grady, 2004). Mutated APC, which occurs in more than 80 % of all CRC cases, has also been discussed as potential cause of CIN, however some MSI tumours with APC mutations do not exhibit any CIN characteristics (Alberici & Fodde, 2006; Guinney et al, 2015; Moran et al, 2010).

The classic adenoma-carcinoma sequence, first described by Fearon and Vogelstein and closely associated with the left colon and rectum (Bettington et al, 2013; Guinney et al, 2015), described CRC as a disease that acquires oncogenic mutations in *APC*, *KRAS*, *DCC*, *SMAD* and *TP53* (Fearon & Vogelstein, 1990). However, only very few cases of CRC show mutations in all of these genes collectively (Guinney et al, 2015) and new insights have shown that other mechanisms can modulate the crucial signalling pathways WNT, MAPK, PI3K, TGFβ/SMAD and p53 (Figure 1).

The adenoma-carcinoma sequence in the human gut is initiated by dysregulation of the WNT pathway due to loss or reduction of *APC* (> 80 % of all CRCs). The activation of the WNT pathway can also occur by mutations of *CTNNB1* or *AXIN2* (Fodde et al, 2001; Korinek et al, 1997; Liu et al, 2000; Morin et al, 1997), aberrant methylation and downregulation of the *APC* promotor (Esteller et al, 2000; Segditsas et al, 2008), negative regulators *SFRP1*, *2*, *4* and *5* (Suzuki et al, 2004), *APCDD1*, *DKK1*, *AXIN2* and downstream effectors such as *ASCL2* and *LGR5* (de Sousa et al, 2011). Increased expression of the onco-miRs miR-135a and b repressing APC translation (Nagel et al, 2008) and miR-146a stabilising β-catenin (CTNNB1) (Hwang et al, 2014; Lu et al, 2017) have also been detected. Downregulation of the ts-miRs miR-145, disrupting translocation of β-catenin (Michael et al, 2003; Yamada et al, 2013) and miR-506 and miR-101 that both target the polycomb group protein EZH2 (Strillacci et al, 2013; Zhang et al, 2015) have been identified. The activation of the WNT pathway leads to proliferation and the

development of a micro adenoma. CIN-mediated loss of heterozygosity of the wildtype *APC* allele causes further progression to an early adenoma.

Successive activation of the MAPK signalling pathway of EGFR signalling through oncogenic *KRAS* mutation in 30-60 % of all CRCs and rarely *NRAS* mutations (Cekaite et al, 2016) (Irahara et al, 2010) leads to further proliferation and the development of an intermediate adenoma. RAS/MAPK activation can also occur through aberrant methylation of pro apoptotic and growth inhibition modulators *RASSF1* and *2* (80% of CRC) (Akino et al, 2005; Fernandes et al, 2013; Harada et al, 2007). In normal non-mutant *KRAS* the ts-miRs miR-143 (Chen et al, 2009), miR-145 (Kent et al, 2010) and let-7 (Akao et al, 2006) target *KRAS* directly and have been found downregulated, also allowing activation of the RAS/MAPK signalling and subsequent proliferation.

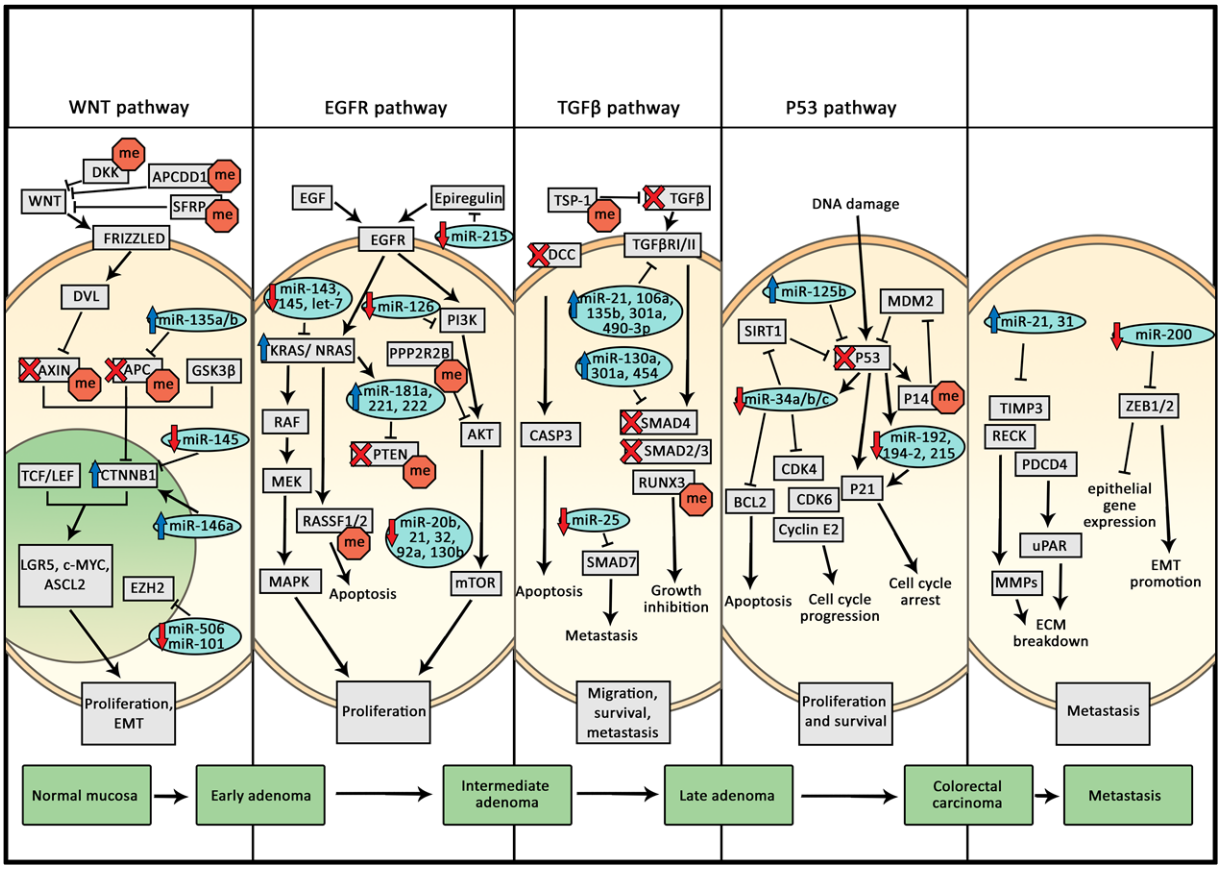


Figure 1 Display of the classic adenoma-carcinoma sequence incorporating mutations, aberrant methylation and miRNA dysregulation in the WNT, MAPK, PI3K, TGFβ and p53 pathway observed in CRC (Adapted from (Fearon, 2011; Shirafkan et al, 2018). Red crosses indicate loss of protein by loss of function mutations or CIN mediated loss of alleles. Red octamers with me mark reduced protein level caused by methylation-mediated gene silencing. Blue arrows indicate oncogenic mutations causing increased protein function or increased expression of miRNAs. Red arrows symbolise reduced miRNA expression.

EGFR activation can also lead to signalling of the PI3K survival promoting pathway which is regulated via the tumour suppressor PTEN. *PTEN* has been found to be mutated in CRC (Molinari & Frattini, 2013),

both in MSI and CIN marked tumours (Guinney et al, 2015). Altered methylation of *PTEN* (in 20 % of CRCs) (Goel et al, 2004) and upregulation of a vast number of miRNAs targeting *PTEN* have been identified to reduce its tumour suppressor function in CRC. The upregulation of miR-181a, 221 and 222 induced by oncogenic KRAS (Ota et al, 2012; Tsunoda et al, 2011) combined with miR-21, 32, 92a, 20b and 130b (Sarver et al, 2009; Wu et al, 2013; Zhang et al, 2014a; Zhu et al, 2014) that all target and repress *PTEN* (Garofalo et al, 2009; Nishimura et al, 2012; Tsunoda et al, 2011; Zhang et al, 2017), highlight its significance in CRC progression and PI3K pathway control. Activation of the PI3K pathway has been observed by methylation-mediated downregulation of *PPP2R2B* (Goel et al, 2004; Tan et al, 2010) and downregulation of the ts-miR miR-126 repressing expression of p85β, a regulatory subunit that stabilises PI3K itself (Guo et al, 2008) and miR-215 repressing the EGFR ligand epiregulin, which also influences the RAS/MAPK pathway (Vychytilova-Faltejskova et al, 2017).

The TGFβ response is lost due to mutations or loss of *DCC* and *SMAD4* by CIN, promoting development of a late adenoma (Mehlen & Fearon, 2004; Mehrvarz Sarshekeh et al, 2017; Takayama et al, 2006). Mutations in *SMAD2* and *3* but also in *TGFβ* were observed in CRC (Grady et al, 1999; Leary et al, 2008; Wood et al, 2007). The pathway components *TSP1* and *RUNX3* (20% of CRCs) have been found to be downregulated by aberrant methylation (Imamura et al, 2005; Rojas et al, 2008). Inactivation of the TGFβ response in CRC was observed by upregulation of miR-21, 106a, 301a and 135b, that target and repress the TGFβR2 and miR-490-3p, targeting TGFβR1 (Feng et al, 2012; Li et al, 2015a; Xu et al, 2015; Yu et al, 2012; Zhang et al, 2014b). Downstream miR-130a, 301a and 454 have also been found to be upregulated in CRC and inversely correlated with the expression of their target SMAD4 (Liu et al, 2013). The ts-miR miR-25 downregulated in CRC targets *SMAD7* (Li et al, 2013). The modulation of the TGFβ pathway enables survival, migration and evasion of growth inhibition and the progression to a late adenoma.

The p53 pathway plays a significant role in almost all cancers including CRC where *TP53* itself is often mutated and the wildtype allele even lost (loss of heterozygosity (LOH)) resulting in evasion of apoptosis and carcinoma formation (Fearon & Vogelstein, 1990). Additionally, miR-125b, upregulated in CRC, targets *TP53* directly and reduces its expression (Nishida et al, 2011). Aberrant methylation of p53 target genes *IGFBP7* and *P14* (20 % of CRCs), an inhibitor of p53 degradation, was found to inactivate this proapoptotic pathway (Nyiraneza et al, 2012; Shen et al, 2003; Suzuki et al, 2010). Activated p53 can induce the upregulation of ts-miRs miR-34a/b/c, that inhibit the p53-inhibiting factor SIRT-1 and essential factors for cell cycle progression and apoptosis inhibitors such as BCL2 (He et al, 2007; Misso et al, 2014; Rokavec et al, 2014). MiR-34a/b/c and miR-192, 194-2 and 215 that induce P21-mediated cell cycle arrest (Akao et al, 2011; Yamakuchi et al, 2008) were found to be

downregulated in CRC (Braun et al, 2008; Chiang et al, 2012) inactivating the p53 pathway of cell cycle arrest and apoptosis.

The activated WNT and inactivated TGFβ pathway contribute to epithelial mesenchymal transition (EMT) (by MYC activation in WNT), essential for metastasis (Weinberg, 2007). Additionally, upregulation of the miR-21 and 31 that target and repress *TIMP3* and *RECK*, increases in turn matrix metalloproteinases (Bandres et al, 2006; Slaby et al, 2007). Downregulation of miR-200 inhibits TGFβ1 and also targets ZEB1/2 (Gregory et al, 2008). By increased ZEB1/2, E-cadherin is reduced and vimentin increased, allowing EMT and Metastasis (Mongroo & Rustgi, 2010).

It must be noted that the classic adenoma-carcinoma sequence is associated with mainly CIMP negative and CIMP low, rendering local hypermethylation-mediated silencing rarer than in the MSI pathway (Guinney et al, 2015). However, there is often no clear distinction between the MSI pathway and classic adenoma-carcinoma sequence as they can blend into each other (Guinney et al, 2015). Therefore, the understanding of the importance of the five crucial signalling pathways in both carcinogenesis and the multitude of modulation mechanisms is essential to comprehend disease progression and to enable development and improvement of therapy (Figure 1).

1.3 Adenomatous polyposis coli

The gene adenomatous polyposis coli (APC) spans 21 exons in humans and encodes a very important tumour suppressor of 2843 amino acid length (Fearnhead et al, 2001). It plays an essential role in the regulation of the WNT pathway. Under normal conditions APC together with GSK3β and AXIN regulates the proteasomal degradation of β-catenin (CTNNB1) (Rubinfeld et al, 1996) (Figure 2). When a WNT molecule binds its receptor frizzled (FZD), the protein dishevelled (DVL) becomes activated and dephosphorylates AXIN and thus inactivates the whole β-catenin degradation complex (Fagotto et al, 1999; Willert et al, 1999; Yamamoto et al, 1999). This leads to the accumulation of β-catenin in both the cytoplasm and the nucleus (Fearnhead et al, 2001; Kobayashi et al, 2000; Smalley et al, 1999). Together with the TFs T-cell factor (TCF) and lymphoid enhancer factor (LEF) (Behrens et al, 1996), β-catenin induces the transcription of genes facilitating proliferation and migration (Roose & Clevers, 1999).

More than 80 % of all sporadic CRC cancers carry an APC mutation (Fearnhead et al, 2001; Moran et al, 2010). 70-80 % of those mutations cause a premature stop resulting in translation of a truncated APC protein. 60 % of these mutations are located between amino acid codon 1286 and 1513, the

mutation cluster region (MCR) (Fearnhead et al, 2001; Miyoshi et al, 1992). Mutations in this region specifically cause loss of many AXIN and β-catenin binding sites, resulting in reduced β-catenin degradation (Polakis, 1997). When both APC alleles carry such a mutation, or the wildtype allele is lost by CIN (Solomon et al, 1987) or hypermethylation (Esteller et al, 2000; Segditsas et al, 2008), there is no more proteasomal degradation of β-catenin (Figure 2). Thus, the WNT signalling pathway is constantly active, independent of a WNT signal, by constant β-catenin accumulation in cytoplasm and nucleus (Korinek et al, 1997; Munemitsu et al, 1995). Successively cancer promoting genes, facilitating proliferation, migration and EMT are constantly expressed (Boon et al, 2002; Mann et al, 1999). Additionally, the β-catenin bound by e-cadherins will be released upon EMT by reduction of e-cadherin, further magnifying the signal (Heuberger & Birchmeier, 2010; Lu et al, 2003). The codons 1309 and 1450 are by far the most frequently mutated amino acids in APC in sporadic CRC (Beroud & Soussi, 1996).

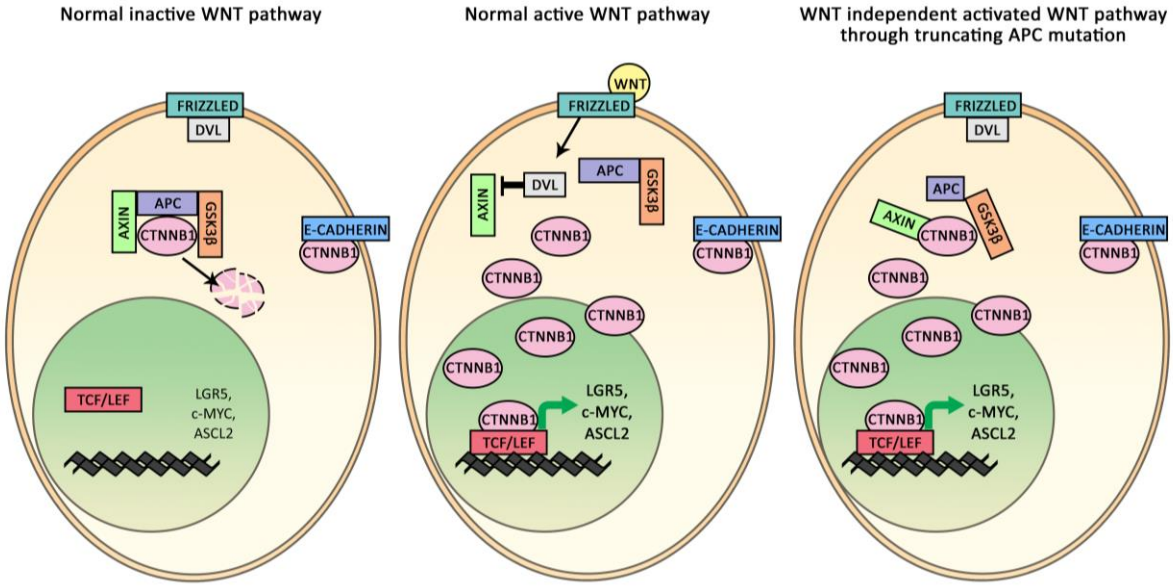


Figure 2 A display of the inactive, active and WNT independent WNT pathway (Adapted from (Pennisi, 1998). APC, adenomatous polyposis coli; CTNNB1, β-catenin; DVL, dishevelled; GSK3 β, glycogen synthase kinase 3β.

1.4 Familial adenomatous polyposis

Familial adenomatous polyposis (FAP) is an autosomal dominant disease caused by germline mutations of the *APC* gene (Kinzler et al, 1991).

FAP can vary in severity but patients typically develop hundreds of adenomatous polyps already during puberty or their twenties (Croner et al, 2005). If these polyps are not removed surgically or medically treated, statistically 10 % of them will proceed to adenocarcinomas by the average age of 36 (Croner

et al, 2005; Fearon, 2011; Fodde & Smits, 2001). Just like in sporadic CRC, mutations in the *APC* gene at position 1309 are the most diagnosed in FAP patients and associated with a very severe phenotype of polyposis (Crabtree et al, 2002; Miyoshi et al, 1992). Just like in the adenoma-carcinoma sequence of sporadic CRC the initiating event of FAP is dysregulation of the WNT pathway by APC mutations. Thus FAP replicates many aspects of adenoma carcinoma sequence pathology (Bettington et al, 2013).

1.5 Modifier genes

CRC like other cancers is a systemic multifactorial disease. The two main carcinogenic pathways outlined here illustrate the variety of mechanisms and alterations that contribute to tumorigenesis (Figure 1). Therefore, it is not surprising that there are FAP patients carrying the same germline *APC* mutation but suffer different extents of polyposis severity (Crabtree et al, 2001). This difference is believed to be caused by genes and loci that pose a genetic low risk predisposition, so called modifier genes that modulate the polyposis severity (Crabtree et al, 2002; Houlston et al, 2001). Genome-wide association studies have been employed to detect SNPs that are associated with sporadic CRC (Broderick et al, 2007; Crabtree et al, 2002; Dunlop et al, 2012; Houlston et al, 2001; Tomlinson et al, 2007; Tomlinson et al, 2008; Whiffin et al, 2014). A study focussed on FAP patients showed, that two of the SNPs associated with sporadic CRC risk are also associated with severe polyposis in FAP (Ghorbanoghli et al, 2016). Thus the identification of modifier genes or loci mediating severe polyposis in FAP may also mediate severe polyposis and thus higher risk of CRC in sporadic CRC with *APC* mutations. The identification of such modifier genes with the help of next generation sequencing (NGS) methodology is of great value to determine susceptibility to CRC very early. The early identification of patients genetically susceptible to CRC may allow improved preventive screening.

1.6 Next generation sequencing and colorectal cancer

Sequencing, determining the order of nucleotides of DNA, was first aimed to identify the sequence of the genes and the entire genome. The first human genome was sequenced using the improved Sanger or termination sequencing method, where the chain-terminating dideoxynucleotides (ddNTPs) were labelled with four different fluorescent dyes (Sanger et al, 1977; Smith et al, 1986) in a shotgun sequencing approach (Anderson, 1981; Gardner et al, 1981; Staden, 1979) and paired end setting to eliminate and reduce sequencing gaps (Edwards et al, 1990; Ewing & Green, 1998; Ewing et al, 1998; Roach et al, 1995; Venter, 2003; Venter et al, 2001).

NGS or high-throughput sequencing incorporate different technologies that all allow the sequencing of DNA and RNA in a cheaper, faster and larger scale (1000-1 x 10^6 DNA molecules simultaneously) than Sanger sequencing.

One of NGS technologies is Illumina (Solex) sequencing. This technique is a combination of the Sanger termination method and pyrosequencing by synthesis (Nyren & Lundin, 1985; Ronaghi et al, 1998) by using reversible terminator dye-labelled dNTPs (Canard & Sarfati, 1994). Also the system is based on DNA colony sequencing (Kawashima, E. H. et al, 1998, International patent no. WO1998044151A1; Kawashima, E. H. et al. (1998) International patent no. WO1998044152A1).

Now that the genomes of many organisms were widely elucidated NGS whole genome sequencing was used to identify the genomic landscape of human cancers including CRC (Vogelstein et al, 2013; Wood et al, 2007). However, NGS methodology in oncology and CRC does not only help characterise the disease further (Cancer Genome Atlas, 2012), but is used to identify a consensus molecular classification (Guinney et al, 2015), to perform molecular diagnosis that aid treatment prognosis (Hsu et al, 2016; Jesinghaus et al, 2016), to identify genetic high, medium and low risk factors (Broderick et al, 2007; Fernandez-Rozadilla et al, 2014; Palles et al, 2013; Tomlinson et al, 2008; Whiffin et al, 2014) and to find biomarkers of disease, progression and drug vulnerability (Garnett et al, 2012). For the detection of specific known SNPs, mutations, genomic amplifications or deletions or gene or miRNA expression differences microarrays are useful (Malapelle et al, 2015; Serrati et al, 2016). However, also in diagnostics, there are cases where no gene of the microarray panels is tested positively (LaDuca et al, 2014; Susswein et al, 2016). Therefore, and for search of unknown mechanisms in CRC progression whole genome, whole exome, whole transcriptome, mRNA or miRNA sequencing is required. Both DNA and RNA sequencing allow comparative analysis between groups. RNA sequencing allows comparative quantitative expression analysis of genes, miRNAs and entire gene sets. This is important as shown above molecular changes in CRC are not only based on genomic changes but also by miRNAs dysregulation, epigenetic modifications and TFs.

1.7 Early diagnosis

The survival rate of CRC is strongly dependent on the stage of the disease at the time of diagnosis (Tomlinson et al, 2012). CRC diagnosed at stage 1 has a 5-year survival rate of 89.9 % according to the American cancer society, while later stages without and with lymph node metastasis range from 71.3 - 13.9 % respectively (US national cancer institute). However, only about 12 % of all CRC cases are diagnosed at stage 1 (Guinney et al, 2015). Early diagnosis is challenging due to the fact that symptoms of CRC are often not very specific. Primary and most frequently observed symptoms are altered bowel

habits, such as diarrhoea or constipation, abdominal pain, change in stool size, weight loss, weakness, iron deficiency and anaemia, which can be mistaken as symptoms of digestive irregularities or other diseases (Dziki et al, 2015; Tomlinson et al, 2012). Only late very severe symptoms such as rectal bleeding and abdominal mass showed symptom specificity of >95 % (Labianca et al, 2013).

The current standard method of diagnosis is colonoscopy (Kim et al, 2014). However, due to high invasiveness and high costs, many patients do not undergo colonoscopy, rendering it unsuitable as preventive screening. A less invasive and more frequently performed screening method is the group of faecal tests including faecal occult blood test (FOBT), the faecal immunochemical test (FIT), DNA- and RNA-based biomarker tests, Protein biomarker test and faecal microbiome-based biomarker test (Schreuders et al, 2016). Even though the DNA based biomarker test showed higher sensitivity, the FIT displayed higher specificity and lower costs (Imperiale et al, 2014). Therefore, the FIT that detects traces of blood in the stool via antibody-based haemoglobin detection, is the most sensitive cost effective screening method (Brenner & Tao, 2013; Hol et al, 2009; Song & Li, 2016). However, these tests are not compulsory and many patients who are not aware of a disease risk are not inclined to undergo such tests voluntarily.

All cells in the organism and especially cancer cells secrete nucleic acids into the blood (O'Driscoll, 2007) which presents another non-invasive screening opportunity. The detection of oncogenic mutations and other biomarkers of CRC in the blood has already been shown (Cassinotti et al, 2013; Chen et al, 2015; Lim et al, 2013). The presence and detection of nucleic acids in blood could allow non-invasive screening for susceptibility factors for CRC. The knowledge of a susceptibility status towards CRC, would make patients more aware to the risk of CRC and might encourage them to undertake more screening procedures.

1.8 Animal models for CRC

No *in vitro* system can model the whole-body pathophysiology of a systemic disease as CRC. Therefore, to identify markers of susceptibility and to fully understand the heterogeneous aspects of CRC, well defined disease animal models, that replicate relevant aspects of human CRC pathology as closely as possible, are essential.

1.8.1 Mouse

The mouse is the best studied and most commonly used mammalian model organism in biomedical research. This is mainly due to convenient and cheap housing and the advances and ease of modifying them genetically (Chu et al, 2016; Skarnes et al, 2011).

Mice have provided valuable information on the molecular basis of many human diseases and facilitated multiple proof of principle studies. However, mice do not always replicate human disease pathology accurately, reducing their predictive value for preclinical studies (Mak et al, 2014). To improve the prediction of safety and effectiveness of novel drugs in clinical trials, non-rodent species may prove valuable in preclinical studies (Bahr & Wolf, 2012; Justice & Dhillon, 2016; Ledford, 2011). In mice many different approaches, mostly targeting *Apc*, have been applied to model human FAP, the hereditary predisposition to CRC (Karim & Huso, 2013). The sole mutation of *Apc* however, did not achieve full replication of human polyposis. In contrary to human pathology where the polyps develop in the colon, the widely used Apc^{Min} mouse develops polyps mainly in the small intestine (Karim & Huso, 2013). The combination of *Apc* mutations with tissue-specific and locally activated oncogenes, has allowed more successful modelling of FAP (Fearon, 2011; Hung et al, 2010; Tetteh et al, 2016). Engineered addition of extra mutations in mice does not allow accurate modelling and study of the spontaneous alterations and mutation events that occur in the human gut subsequent to an APC initiating event and drive disease progress towards cancer.

Therefore, other model organisms are required to model the disease more accurately and to provide better insight into the progression, to identify better diagnosis and therapy opportunities and to provide non-rodent preclinical data (Bahr & Wolf, 2012)

1.8.2 Pig

Pigs share many key similarities with humans including anatomical features, physiology, body size and pathophysiological responses and are already used for research, development and refinement of medical equipment and biomedical procedures (Heinritz et al, 2013; Kararli, 1995; Schubert et al, 2016; Swindle et al, 1988). Further advantages are that they mature relatively quickly (6-7 months), produce large litters (~10 piglets) and have a short gestation time of ~114 days (Sachs, 1994). Where necessary specified pathogen free housing is easily adapted from the established domestication of the pig. In contrast to primates the use of pigs for research is ethically and socially more acceptable due to their use as food animals. Additionally, just like humans, pigs are omnivores which is a marked advantage for the study of CRC, facilitating dietary studies such as the effects of a western diet. The gastrointestinal structure and size allows the use and refinement of human diagnostic and operational

equipment. The relatively long lifespan of pigs (~10 years) enables the performance of longitudinal studies. Most importantly however, genetic modification technology was extended to the pig, allowing genetically modified pigs for biomedical research.

1.9 Genetic modification of pigs

Genetically modified pigs were first generated by pronuclear DNA microinjection (Hammer et al, 1985) (Figure 3). This rather inefficient procedure (Hammer et al, 1985; Logan & Martin, 1994; Uchida et al, 2001) was technically challenging due to the opacity of porcine oocytes which required centrifugation to visualise the pronuclei (Kikuchi et al, 2002). The establishment of refined methods of microinjection including transposon systems and lentiviral vectors has increased transgenesis efficiency in pigs markedly (Clark et al, 2007; Garrels et al, 2011; Hofmann et al, 2003; Ivics et al, 2014; Whitelaw et al, 2004). Even though, pronuclear DNA microinjection allowed only random transgene integration but no gene targeting and remained the only method for genetic modification of livestock until 1997, when Schnieke et al. performed nuclear transfer of primary somatic ovine cells into ovine enucleated oocytes (Schnieke et al, 1997).

The generation of genetically modified mice by gene targeting via homologous recombination (HR) first performed and established in murine embryonic stem (ES) cells (Evans & Kaufman, 1981; Smithies et al, 1985; Thomas & Capecchi, 1987), could not be performed in pigs, as ES, embryonic germ or induced pluripotent stem cells, capable of germline transmission were elusive in pigs (Nowak-Imialek & Niemann, 2012). Nuclear transfer of cultured cells, however, enabled gene targeting of primary somatic cells first in sheep (McCreath et al, 2000) and later in pigs (Dai et al, 2002; Lai et al, 2002). Genetic modification of primary porcine cells is more challenging than the modification of murine ES cells due to their limited lifespan (Schnieke et al, 1997). The efficiency of HR is also lower compared to mice, although loci such as the porcine *ROSA26* allow more efficient gene targeting (Li et al, 2014). Additionally, the procedure of nuclear transfer requires time and skill and has a low efficiency in generating viable healthy offspring (Callesen et al, 2014; Kurome et al, 2013). Although nuclear transfer is the standard method for generating genetically tailored pigs, it is a very challenging process especially when compared to producing gene targeted mice using ES cells (Perleberg et al, 2018).

The optimisation of gene targeting in pigs has been inspired by techniques performed in mice. Therefore, methods such as site specific recombinase technology (Leuchs et al, 2012; Li et al, 2014), recombinase-mediated cassette exchange (RMCE) (Clark et al, 2007; Jakobsen et al, 2013) and adeno-associated viral (AAV) vectors (Luo et al, 2011) have all been extended to pigs.

The development of synthetic engineered nucleases, however, changed the procedure of both random transgene integration and gene targeting in all model organisms (Figure 3).

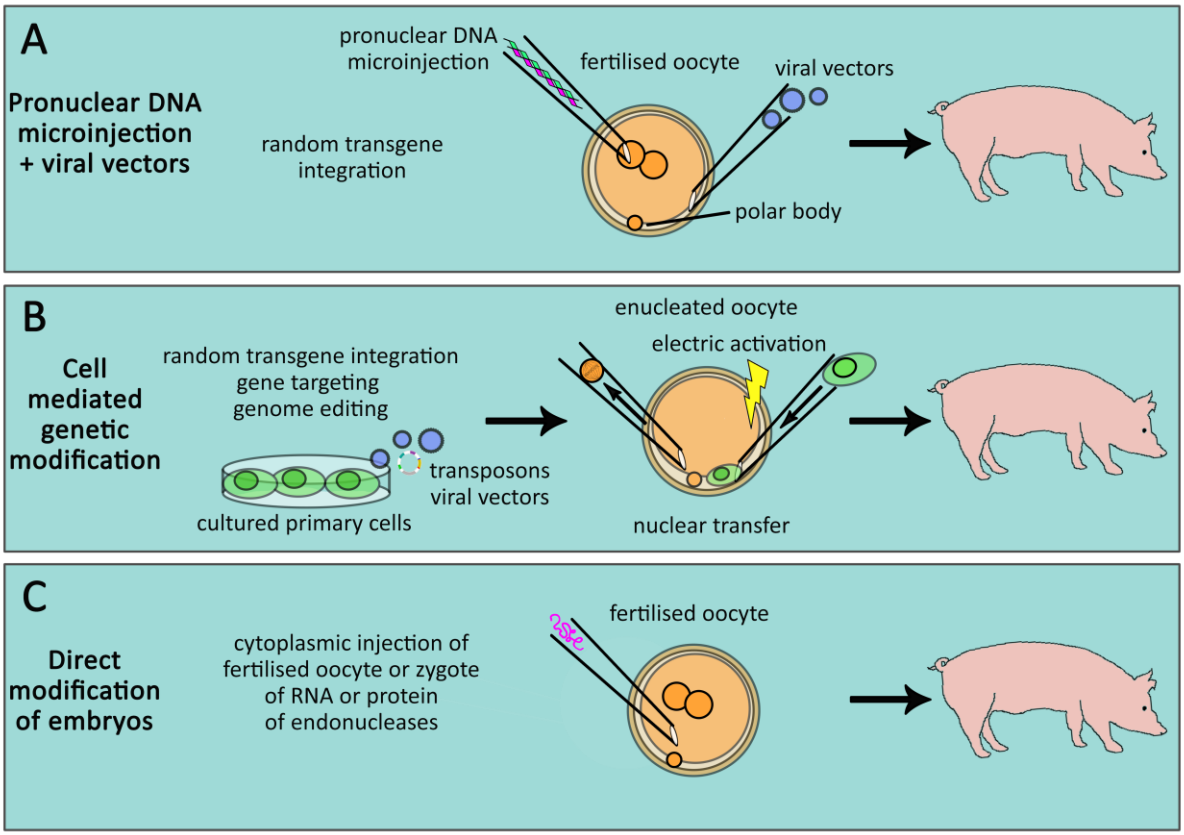


Figure 3 A collection of the different methods applicable for the generation of genetically modified pigs (adapted from (Perleberg et al, 2018).

1.9.1 Genome editing in pigs

Engineered endonucleases can induce a double strand break (DSB) at a unique targeted genomic location (Figure 4). The error prone repair mechanism of non-homologous end joining (NHEJ) introduces deletion or insertions of several bases, which can result in gene inactivation (Gabriel et al, 2011). The homology-directed repair uses templates, such as sister chromatids or exogenous homologous DNA fragments, for accurate repair, facilitating transgene introduction or sequence replacement (Joung & Sander, 2013).

Zinc finger nucleases (ZFN) (Hauschild et al, 2011; Kwon et al, 2013) and transcription activator-like effector nucleases (TALENs) (Carlson et al, 2012) compose the second and third generation of engineered nucleases after meganucleases, respectively. Both bind the target via protein-DNA interaction (Font & Mackay, 2010; Mussolino & Cathomen, 2012) which hampers design, target and off-target prediction (Bogdanove & Voytas, 2011; Jinek et al, 2013; Urnov et al, 2010). The newest

system is the RNA-guided endonuclease. The most famous representative is the clustered regularly interspaced short palindromic repeat (CRISPR) system (Tan et al, 2013), first identified as the adaptive immune system of bacteria and archaea (Barrangou et al, 2007). Extrachromosomal DNA incorporated into the genome transcribed into CRISPR RNA (crRNA) and the repeats transcribed into transactivating CRISPR RNA (tracrRNA) (Garneau et al, 2010) bind the endonuclease CRISPR-associated 9 (Cas9) and guide it to a sequence complementary to the crRNA (Wei et al, 2013). Cas9 cleaves the double stranded DNA if the crRNA target (protospacer) is followed by a protospacer adjacent motif (PAM) of NGG (Wei et al, 2013). This RNA-DNA interaction-based targeting system is much easier to design and to predict (Cho et al, 2013). Therefore, this system was simplified to a two-component system for genome editing, where the crRNA and tracrRNA were fused to a crRNA-tracrRNA chimera called guide RNA (gRNA) (Jinek et al, 2012) (Figure 4). Other RNA-guided endonucleases with different PAMs are known, for example Cpf1 (Zetsche et al, 2015) and of smaller sized CjCas9 (Kim et al, 2017).

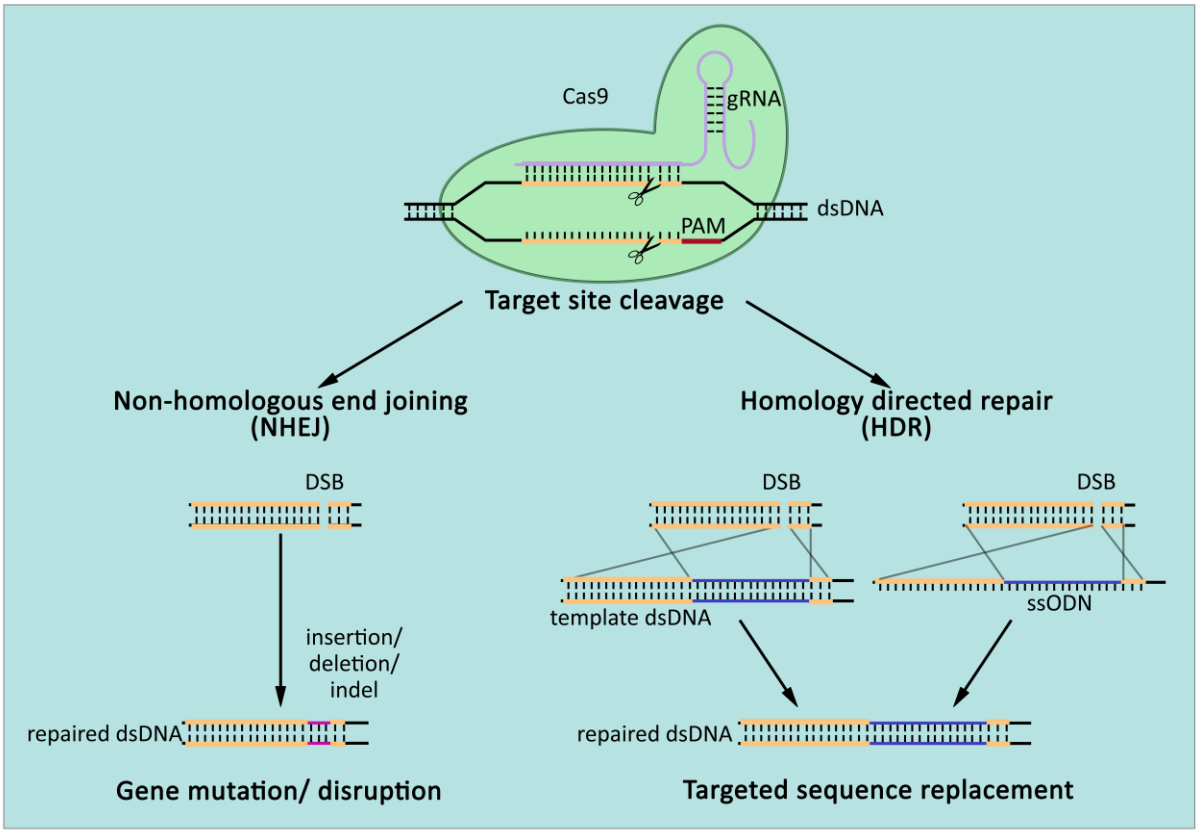


Figure 4 The CRISPR/Cas9 system and its potential utilised for genome editing (adapted from (Perleberg et al, 2018). DSB, double strand break; dsDNA, double-stranded DNA; gRNA, guide RNA; PAM, protospacer adjacent motif; ssODN, single-stranded oligodeoxynucleotide.

Genome editing in general has not only improved efficiency of targeted genetic modifications of both knock-out and knock-ins in mice (Chu et al, 2016; Miyaoka et al, 2016) and pigs (Zhou et al, 2016) but

also facilitated direct zygote and early stage embryo modifications in mice (Meyer et al, 2010). HR in zygotes, which is very inefficient (<0.1 %) (Brinster et al, 1989), was increased to an efficiency of 1.7-4.5 % by the use of ZFN. This procedure has been extended to pigs avoiding nuclear transfer (Lillico et al, 2013).

The CRISPR/Cas9 system superseded ZFN and TALENs by easier design, higher efficiency, specificity and equal to better prediction of targets and off-targets (Cho et al, 2014; Fu et al, 2014; Mali et al, 2013; Wu et al, 2016). Therefore, it has been used for direct zygote modification to generate genetically modified pigs (Hai et al, 2014; Wang et al, 2015; Whitworth et al, 2014; Yu et al, 2016).

1.10 Porcine cancer models

Table 1 An overview of genetically modified porcine models of human cancers, the most promising and clinically relevant are marked with * (Perleberg et al, 2018). AAV, adeno-associated viral vector; MMTV, mouse mammary tumour virus; TALENs, transcription activator-like effector nucleases.

Human disease	Genetic modification	Produced by	Reference
Basal cell carcinoma	human truncated *GLI2*, with bovine keratin 5 promoter	random transgene integration + nuclear transfer	(McCalla-Martin et al, 2010)
Breast cancer	*V-H-Ras*, with MMTV promoter	Pronuclear microinjection	(Yamakawa et al, 1999)
	BRCA1	gene targeting by AAV + nuclear transfer	(Luo et al, 2011)
Colorectal cancer	heterozygous APC^{1311} + APC^{1061} mutations	gene targeting + nuclear transfer	(Flisikowska et al, 2012)*
	APC Knockout	TALENs + nuclear transfer	(Tan et al, 2013)
Osteosarcoma	heterozygous and homozygous *TP53* knockout	gene targeting + nuclear transfer	(Saalfrank et al, 2016)*
	homozygous $TP53^{R167H}$ mutation	gene targeting by AAV + nuclear transfer	(Sieren et al, 2014)*
Other cancers	Cre-induced porcine $KRAS^{G12D}$ $TP53^{R167H}$	random transgene integration + nuclear transfer	(Schook et al, 2015)*
	Cre-inducible $TP53^{R167H}$	gene targeting + nuclear transfer	(Leuchs et al, 2012)
	Cre-inducible $KRAS^{G12D}$	gene targeting + nuclear transfer	(Li et al, 2015b)

To date, random transgene integration, gene targeting and genome editing using TALENs have been utilised to generate a number of porcine cancer models for basal cell carcinoma, breast cancer, colorectal cancer, osteosarcoma and other cancers (Table 1). Even Cre-inducible systems have been introduced to allow a spatio-temporal activation of mutations to better model malignant diseases. However the most successful models with reported phenotypes similar to human disease are those marked with * (Perleberg et al, 2018). The other models were not reported with a phenotype (Tan et al, 2013; Yamakawa et al, 1999), were euthanized early due to bacterial infection (McCalla-Martin et al, 2010), did not survive beyond 18 days (Luo et al, 2011) or in case of the Cre-inducible mutations were not activated yet (Leuchs et al, 2012; Li et al, 2015b). The porcine model for CRC generated by the heterozygous targeted *APC*^{1311} mutation, however showed great potential (Flisikowska et al, 2012).

1.11 The porcine colorectal cancer model *APC*^{1311} pig

The chair of livestock biotechnology of the technical university of Munich, has generated a porcine model for colorectal cancer that carries a translational stop signal at codon 1311 in porcine APC (*APC*^{1311}) (Flisikowska et al, 2012). It is orthologous to the human *APC*^{1309} mutation, which is frequently mutated in CRC and associated with a very severe form of FAP (Crabtree et al, 2002). This stop codon was introduced via homologues recombination-mediated gene targeting of porcine somatic cells and piglets were subsequently generated by nuclear transfer. The animals showed polyp development in the colorectum as early as 4 months of age, like humans during puberty or their twenties (Croner et al, 2005). Closer investigations showed that the polyps exhibit epithelial features such as aberrant crypt foci and adenomatous polyps of low and high grade intraepithelial neoplasia, that are typical for human adenoma-carcinoma sequence pathology (Flisikowska et al, 2012). Molecular analyses of adenomas showed loss of APC heterozygosity, accumulation of β-catenin and high expression of its target c-MYC and MAPK pathway activation by ERK1/2 phosphorylation, all hallmarks of human CRC (Albuquerque et al, 2002; Fearon, 2011). Although no invasive carcinomas were observed in pigs up to 3 years of age, carcinomas *in situ* were identified, a progression stage that has not been seen in mice. As a heterozygous *APC*^{1311} mutation was sufficient to initiate polyposis and those spontaneous molecular changes, that present hallmarks of human FAP and CRC, invasive CRC formation seems to be a function of time. Further validation and analysis of this model, is however required.

1.12 Objective

The two main goals of this work were the analysis of cancer susceptibility and progression of the *APC*^{1311} porcine colorectal cancer model (Flisikowska et al, 2012) on transcriptome level and the acceleration of CRC development of the model by introducing oncogenic mutations or *in vivo* knock-out of tumour suppressors.

The established *APC*^{1311} pigs, consisting of more than five generations by now, replicate hallmarks of human FAP and early CRC development including adenomatous polyps in the colorectum with loss of APC heterozygosity, β-catenin accumulation, upregulation of c-MYC, MAPK pathway activation and progression to carcinoma *in situ*. However, the animals have displayed a phenotypic variation in polyposis severity from high (≥100) polyps (HP) to low (1-10) polyps (LP) in the distal colorectum (last 40 cm).

Therefore, more thorough characterisation of the model is required for use in translational biomedical research. In this work transcriptional analysis of normal mucosa and polyps from the *APC*^{1311} pigs, was aimed to identify elements in the genetic background such as single-nucleotide polymorphisms (SNPs), dysregulated genes, gene sets and miRNAs that may contribute to susceptibility towards severe polyposis and tumour progression, respectively. This process should reveal both similarities to human molecular pathology and novel markers for early detection and drivers. A holistic approach was performed where whole mRNA and whole miRNA were sequenced using next generation sequencing technology and computational analysis pipelines to compare expression and distribution of genes, miRNAs and SNPs between high polyp and low polyp normal mucosa samples and between high grade and low grade intraepithelial neoplasia.

Current data of the *APC*^{1311} animals showing no invasive carcinomas, suggest, that the progression from adenoma to adenocarcinoma in pigs, is a long process just like in humans. Therefore, *APC*^{1311} animals were to be generated carrying a ubiquitously expressed *Cas9* endonuclease in the *ROSA26* locus. This would allow acceleration of polyp progression *in vivo*, by introduction of gRNAs targeting tumour suppressor genes or gRNA combined with HR templates for the introduction of oncogenic mutations via *in vivo* electroporation of the polyps or adeno-associated viral vectors.

2. Materials and Methods

2.1 Material

2.1.1 Laboratory equipment

Table 2 laboratory equipment

2100 Bioanalyzer	Agilent, Santa Clara, USA
3130xl/3100 Genetic Analyzer 16-Capillary Array	Life Technologies, Darmstadt, DEU
ABI 7500 Fast Real-Time PCR System	Applied Biosystems, Warrington, GBR
ABI Prism 3130xl Genetic Analyzer	Life Technologies, Darmstadt, DEU
Analytical semi-micro balance PI-214	Denver Instrument GmbH, Göttingen DEU
AREC.X Digital Ceramic Hot Plate Stirrer	Velp Scientifica, Usmate, ITA
AxioCam HRm Microscope Camera	Carl Zeiss Microscopy GmbH, Göttingen, DEU
Axiovert 200 M	Carl Zeiss Microscopy GmbH, Göttingen, DEU
Bag sealer Vacupack plus	KRUPS, Frankfurt, DEU
Barnstead™ MicroPure™	Thermo Fisher Scientific, Waltham, USA
Centrifuges Sigma 1-15, 1-15K, 3-16 and 4K-15C	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Countess™ automated cell counter	Invitrogen GmbH, Darmstadt, DEU
Digital Graphic Printer UP-D895MD	Syngene, Cambridge, GBR
Dry block for heating and cooling PCH-2	Grant Instruments, Camebridge, GBR
Drying and heating chamber	BINDER GmbH, Tuttlingen DEU
Electronic multi-dispense pipet	Qiagen, Hilden, DEU
Electrophoresis power supply EPS301	Amersham Biosciences, Piscataway, USA
ELISA-Photometer	Thermo Electron Corporation, Shanghai, CHN
Fluorescence light source HXP120C	Carl Zeiss Microscopy GmbH, Göttingen, DEU
Gel documentation system QUANTUM ST5	VILBER LOURMAT Deutschland GmbH, Eberhardzell, DEU
Glassware	Marienfeld GmbH, Lauda-Königshofen, DEU
HiSeq 2500	Illumina, San Diego, USA
Hybridisation oven Shake'n'Stack	Thermo Fisher Scientific, Waltham, USA
iBind Western Device	Thermo Fisher Scientific, Waltham, USA
Incubator Thermo Forma Orbital Shaker	Thermo Fisher Scientific, Waltham, USA
Incubator Thermo Forma Steri-Cycle CO_2	Thermo Fisher Scientific, Waltham, USA
Incubators	BINDER GmbH, Tuttlingen DEU
KGW Dewar-Transportgefäße mit Deckel Typ B	KGW-Isotherm Karlsruher Glastechnisches Werk - Schieder GmbH, Karlsruhe, DEU
Laser Microdissection Systems 6000	Leica Microsystems, Wetzlar, DEU
Magnetic Stand-96	Invitrogen GmbH, Darmstadt, DEU
Mi Seq	Illumina, San Diego, USA
Microm HM 560 Cryostat	Thermo Fisher Scientific, Waltham, USA
Microscope Axiovert 40 CFL	Carl Zeiss Microscopy GmbH, Göttingen, DEU
Mini-PROTEAN 3 Cell system	BioRad, Hercules, USA

Mini-PROTEAN® Comb, 10-well, 0.75 mm, 33 µl, 1653354 and 1.5 mm, 66 µl, 1653365	BioRad, Hercules, USA
Mini-PROTEAN® Short Plates, 1653308	BioRad, Hercules, USA
Mini-PROTEAN® Spacer Plates with Integrated Spacers 0.75 mm, 1653310 and 1.5 mm, 1653312	BioRad, Hercules, USA
Multiporator Eppendorf	Eppendorf, Hamburg, DEU
Nalgene™ Mr. Frosty Freezing containers	Thermo Fisher Scientific, Waltham, USA
Nanodrop Lite	Thermo Fisher Scientific, Waltham, USA
Nucleofector™ 2b Device	Lonza Group Ltd, Basel, CHE
Owl™ EC-105 Compact Power Supply	Thermo Fisher Scientific, Waltham, USA
PeqSTAR 2x Gradient Thermocycler	Peqlab Biotechnologie GmbH, Erlangen, DEU
pH Meter CyberScan PC 510 Meter	Eutech Instruments Europe B.V., Landsmeer, NLD
Pipette controller accu-jet pro	Brand GmbH & Co. KG, Wertheim, DEU
PyroMark Q48 Autoprep Instrument	Qiagen, Hilden, DEU
Qubit® 2.0 Fluorometer	Thermo Fisher Scientific, Waltham, USA
Rainin Pipet-Lite (2, 20, 200, 1000 µl) and Multi Pipette L8-20XLS+, L8-50XLS+	Mettler Toledo GmbH, Giessen, DEU
Scales 440-33N	Kern & Sohn GmbH, Balingen DEU
Shaker Unitwist 3-D	Uniequip, Martinsried, DEU
SpeedMill PLUS	Analytik Jena AG, Jena, DEU
Sterile laminal flow cabinet Herasafe Type HSP	Thermo Fisher Scientific, Waltham, USA
Table centrifuge blue spin mini	SERVA Electrophoresis GmbH, Heidelberg, DEU
Trans-Blot SD Semi-Dry Transfer cell	BioRad, Hercules, USA
Vacuum Centrifuge Savant, SpeedVac, DNA 110	Thermo Fisher Scientific, Waltham, USA
Vacuum Centrifuge Savant, Speed Vac Plus, SC110A	Thermo Fisher Scientific, Waltham, USA
Vortexer VELP Sccientifica $2x^3$	Velp Scientifica, Usmate, ITA
Water bath	Memmert GmbH + Co.KG, Schwabach, DEU
X-ray clip cassette	Rego X-Ray GmbH, Augsburg DEU

2.1.2 Consumables

Table 3 Consumables

Blotting Paper	BioRad, Hercules, USA
Cell counting chamber slides	Invitrogen GmbH, Darmstadt, DEU
Cell scraper	TPP Techno Plastic Products AG, Trasadingen, CHE
Countess™ cell counting chamber slides	Invitrogen GmbH, Darmstadt, DEU
Cryo vials	Corning Incorporated, Corning, USA
Cutfix stainless scalpel 10, 5518059	B. Braun Melsungen AG, Melsungen, DEU
Disposable sterile needles, Sterican, 1.20 x 40 mm	B. Braun Melsungen AG, Melsungen, DEU
Electroporation cuvette 2, 4 mm	Peqlab Biotechnologie GmbH, Erlangen, DEU
Falcon tubes 15 ml, 50 ml	Greiner Bio-One GmbH, Frickenhausen, DEU
iBind cards, Bi15126	Invitrogen GmbH, Darmstadt, DEU
innuSPEED Lysis Tube P	Analytik Jena AG, Jena, DEU
Kimtech Science Precision wipes, 05511 7552	Kimberly-Clark Professional, Roswell, USA
Lysing Matrix D, 2 mlTube	MP Biomedicals, Santa Ana, USA
Membrane, Roti-PVDF (0.45 µm)	Brand GmbH & Co. KG, Wertheim, DEU
MembraneSlide 1.0 PEN (D), 415190-9041-000	Carl Zeiss Microscopy GmbH, Göttingen, DEU
MicroAmp Fast Optical 96-Well Reaction Plate with Barcode, 4346906	Applied Biosystems, Warrington, GBR
MicroAmp™ Optical Adhesive Film, 4360954	Applied Biosystems, Warrington, GBR
Mini Trans-Blot Filter paper 1703932	BioRad, Hercules, USA
MultiScreen HV plates, MAHVN4550	Millipore, Darmstadt, Deutschland
Nylon membrane (positively charged) Amersham Hybond-N+	GE Healthcare Europe GmbH, Freiburg, DEU
PCR tube 0.2 ml 8- strip, I1402-2900	STARLAB International GmbH, Hamburg, DEU
Petri dishes	Brand GmbH & Co. KG, Wertheim, DEU
Rainin pipette tips with and without filter 20, 200, 1000 µl	Mettler Toledo GmbH, Giessen, DEU
Reaction tubes 1.5 ml, 2 ml	Zefa Laborservice, Harthausen, DEU
Serological pipets Costar 1, 2, 5, 10, 25 ml	Corning Incorporated, Corning, USA
Sterile syringes 5 ml, 10 ml, 20 ml	Becton Dickinson GmbH, Sparks, USA
Sterile syringe filters 0.40 µm, 0.22 µm	Berrytec GmbH, Grünwald, DEU
Tissue culture vessels T25, T75, T150, 24 well, 12 well, 6 well plates, 10 and 15 cm dishes	Corning Incorporated, Corning, USA
Tissue-Tek Cryomold® Biopsy, 10 x 10 x 5 mm, 4565	Sakura Finetek Europe B.V., Alphen aan den Rijn, NLD
14 ml tubes	Becton Dickinson GmbH, Sparks, USA
Twin.tec PCR Plate 96 semi skirted, colorless, 951020303	Eppendorf, Hamburg, DEU
X-ray film	Agfa Healthcare, Mortsel, Belgium
PyroMark Q48 Absorber Strips, 974912	Qiagen, Hilden, DEU
PyroMark Q48 Discs, 974901	Qiagen, Hilden, DEU

2.1.3 Chemicals

Table 4 Chemicals

Acetic acid 100 %	Applichem, Darmstadt, DEU
Agarose	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Ammonium acetate (NH_4OAc)	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Ammonium persulphate (APS)	Carl Roth GmbH, Karlsruhe, DEU
Ampicillin	Carl Roth GmbH, Karlsruhe, DEU
Boric acid	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Chloroform	Applichem, Darmstadt, DEU
4',6'-Diamidino-2'-phenylindole-dihydrochloride (DAPI)	Roche Diagnostic GmbH, Mannheim, DEU
Difco Luria Broth Agar, Miller	Becton Dickinson GmbH, Sparks, USA
Difco Luria Broth Base, Miller	Becton Dickinson GmbH, Sparks, USA
Dimethyl sulphoxide (DMSO), A3672	Applichem, Darmstadt, DEU
Dithiotreitol (DTT)	Omnilab, Bremen, DEU
Eosin solution, Conc. Watery 2 %, 2C-140	Waldeck GmbH & Co. KG, Münster, DEU
Ethanol 100 %	Riedel-de-Haen, Seelze, DEU
Ethidiumbromid	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Ethylene diamine tetracetic acid (EDTA)	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Formalin	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Glacial acetic acid	Fluka Laborchemikalien GmbH, Seelzle, DEU
Glycerol 99%	Carl Roth GmbH, Karlsruhe, DEU
Glycine 99%, G7126	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Hydrochloric acid, 37%	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
IGEPAL, CA-630	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Isopropanol	Applichem, Darmstadt, DEU
Isopropyl β-D-thiogalactopyranoside (IPTG)	Bioline, London, GBR
Magnesium chloride ($MgCl_2$)	Merck, Kenilworth, USA
Maleic acid	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Mayer's Hemalaun solution, A0884	Applichem, Darmstadt, DEU
Methanol	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Milk powder	Carl Roth GmbH, Karlsruhe, DEU
N,N-Dimethylformamid (DMF)	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
PeqGOLD Low Melt-Agarose	Peqlab Biotechnologie GmbH, Erlangen, DEU
Phenol-chloroform-isoamyl alcohol	Applichem, Darmstadt, DEU
Ponceau S, 141194	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
POP-7™ Polymer for 3130/3130xl	Life Technologies, Darmstadt, DEU
Potassium chloride	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Sacharose	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Sephadex G5050-100G	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Sodium azide	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Sodium chloride	Applichem, Darmstadt, DEU
Sodium deoxycholate, D6750	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Sodium dodecyl sulphate (SDS)	Omnilab, Bremen, DEU

Sodium hydroxide (NaOH)	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Sodium chloride, A1371	Applichem, Darmstadt, DEU
Sodium citrate, S4641	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Spectinomycin, 85555	Fluka Laborchemikalien GmbH, Seelzle, DEU
Sucrose	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Tetramethylethylenediamine (TEMED)	Carl Roth GmbH, Karlsruhe, DEU
Triton-X 100	Omnilab, Bremen, DEU
Trizma Base/Hydrochloride	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Trizol	Invitrogen GmbH, Darmstadt, DEU
Tween 20	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
X-ray tank developer	Calbe Chemie GmbH, Calbe, DEU
X-ray tank fixer	Calbe Chemie GmbH, Calbe, DEU
β-glycerol phosphate	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
β-mercaptoethanol (14.3 M)	Sigma-Aldrich Chemie GmbH, Steinheim, DEU

2.1.4 Buffers and solutions

Table 5 Buffers and solutions

10x TBE buffer	0.9 M Tris, 20 mM EDTA, 0.9 M boric acid
5x dilution buffer	25 mM $MgCl_2$, 1 mM Tris-HCl pH 8.0
2-Log DNA ladder	New England Biolabs, Frankfurt, DEU
50x TAE buffer	2 M Trisbase, 50 mM EDTA, 5.71 % (v/v) Glacial acetic acid
6x Gel loading dye	New England Biolabs, Frankfurt, DEU
Advanced protein assay reagent	Cytoskeleton Inc., Denver, USA
Chemiluminescent substrate for alkaline phosphatase CDP-Star	Roche Diagnostic GmbH, Mannheim, DEU
cOmplete Protease Inhibitor Cocktail Tablets in EASYpacks, 04693124001	Roche Diagnostic GmbH, Mannheim, DEU
CytoBuster™ Protein Extraction Reagent, 71009	Merck, Kenilworth, USA
DNA/RNA-dye, peqGREEN	Peqlab Biotechnologie GmbH, Erlangen, DEU
dNTPs	New England Biolabs, Frankfurt, DEU
EB Buffer	Qiagen, Hilden, DEU
Gel loading buffer II, AM8546G	Invitrogen GmbH, Darmstadt, DEU
Low Molecular Weight DNA Ladder	New England Biolabs, Frankfurt, DEU
Low TE Buffer	10 mM Tris-HCl pH7.4, 0.1mM EDTA
Lysis buffer for gDNA isolation	100 mM Tris-HCl pH7.4, 0.2 % SDS, 5 mM EDTA, 0.2 M NaCl
β-mercaptoethanol solution for TC	6 ml H2O, 21 µl β-mercaptoethanol
Miniprep solution I	5 mM sucrose, TRIS (pH 8.0), 10 mM EDTA
Miniprep solution II	1 % SDS, 0.2 N NaOH
Miniprep solution III	3 M sodium acetate (pH 5.3)
O.C.T.™ Compound, 4583	Sakura Finetek Europe B.V., Alphen aan den Rijn, NLD

PhosSTOP Phosphatase Inhibitor Cocktail Tablets in EASYpacks, 04906845001	Roche Diagnostic GmbH, Mannheim, DEU
PyroMark Q48 Magnetic Beads, 974203	Qiagen, Hilden, DEU
Quick-Load® 1 kb Extend DNA Ladder	New England Biolabs, Frankfurt, DEU
RiboRuler™ High Range RNA Ladder	Thermo Fisher Scientific, Waltham, USA
RIPA Buffer	0.15 M NaCl, 50 mM Tris-HCl pH 8, 1mM EDTA, 1 % (v/v) Triton X, 0.5 % sodium deoxycholate, 0.1 % (w/v) SDS
Rnase Away	Thermo Fisher Scientific, Waltham, USA
TE Buffer	10 mM Tris-HCl pH 7.4, 1 mM EDTA
Trypan blue solution (0.4 %)	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
X-Gal solution	100 mg X-Gal in N,N-Dimethylformamid (DMF)

Table 6 Buffers and solutions for Southern Blot

Blocking solution	1x Blocking solution in maleic acid buffer
Blocking reagent, 11096176001	Roche Diagnostic GmbH, Mannheim, DEU
Denaturation buffer	0.5 M NaOH, 1.5 M NaCl
Depurination buffer	250 mM HCl
Detection buffer	0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5 (20 °C)
DIG Easy Hyb Granules, 11796895001	Roche Diagnostic GmbH, Mannheim, DEU
Digoxigenin-11-dUTP, 11573152910	Roche Diagnostic GmbH, Mannheim, DEU
DNA Molecular Weight Marker VII, DIG-labeled	Roche Diagnostic GmbH, Mannheim, DEU
High stringency washing buffer	0.1 % SSC, 0.1 % SDS
Low stringency washing buffer	2x SSC, 0.1 % SDS
Maleic acid buffer	0.1 M maleic acid, 0.15 M NaCl, pH 7.5
Neutralisation buffer	0.5 M Tris-HCl, pH 7.5; 1.5 M NaCl
Washing buffer	0.3 % (v/v) Tween 20 in Maleic acid buffer
Ponceau S staining solution	0.5 % (w/v), 1 % glacial acetic acid
20x SSC	3 M NaCl, 0.3M sodium citrate
2x SSC	0.3 M NaCl, 30 mM sodium citrate

Table 7 Buffers and solutions for Western Blot

10 % APS	10 % (w/v) APS
4 x Lämmli buffer + DTT	250 M Tris-HCl, pH 6,8, 4 % (w/v) SDS, 0,1 M Saccharose, traces of bromphenol blue, 26mM DTT (freshly added)
Milk powder blocking solution	5 % (w/v) Milk powder in 1x TBST
Pierce™ ECL Western Blotting Substrate	Thermo Fisher Scientific, Waltham, USA
Precision Plus Protein™ All Blue Standards (10–250 kDa), 161-0373	BioRad, Hercules, USA
10x Running Buffer, pH 8.3	0.25 M Trizma Base, 2 M Glycin, 1 % SDS, pH 8.3
1x Running Buffer + ß-mercapto-Ethanol	25 mM Trizma Base, 0.2 M Glycin, 0.1 % SDS, pH 8.3, 10.64 mM ß-mercaptoethanol

1x Semi Dry Transfer Buffer + 0.1% SDS	25 mM Trizma Base, 0.2 M Glycin, 20 % (v/v) Methanol, 0.1 % (w/v) SDS
10x TBS	0.2 M Trizma Base, 1.4 M NaCl
1x TBST	20 mM Trizma Base, 140 mM NaCl, 0.1 % (v/v) Tween 20

2.1.5 Bacterial media

Table 8 Bacterial media

LB-agar	4 % (w/v) Difco LB Agar, Miller
LB-medium	2.5 % (w/v) Difco LB Base, Miller

2.1.6 Tissue culture media and solutions

Table 9 Solutions for tissue culture

1x Trypsin EDTA, T3924	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Accutase, A6964	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Advanced DMEM, 12491-015	Life Technologies, Darmstadt, DEU
Ala-Gln, G8541	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Amphotericin B solution, A2942	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Blasticidin S	InvivoGen, San Diego, USA
BSA 7.5 %	Life Technologies, Darmstadt, DEU
cell culture water, W3500	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
DMEM, D5671	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
D-PBS, D8537	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Fetal calf serum (FCS)	Biochrom GmbH, Berlin, DEU
G-418, M3118.0050	Genaxxon bioscience GmbH, Ulm, DEU
Hypoosmolar electroporation buffer	Eppendorf, Hamburg, DEU
MEM Non-essential amino acids (NeAA), M7145	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
OptiMEM, 51985-026	Life Technologies, Darmstadt, DEU
Penicillin-Streptomycin, P0781 (100x)	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Puromycin	InvivoGen, San Diego, USA
Sodium pyruvate, S8636	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Trypan blue 0.4 %	Invitrogen GmbH, Darmstadt, DEU

Table 10 Media compositions

cultured cells	**Media components**	**Freezing medium composition**
pADMSCs	500 ml AdvancedDMEM, 50 ml FCS, 5.6 ml Ala-Gln, 5.6 ml NEAA, 560 µl β-mercaptoethanol	10 % DMSO, 10 % medium, 80 % FCS
pKFs, HEK cells	500 ml DMEM, 50 ml FCS, 5.6 ml Ala-Gln, 5.6 ml NEAA, 5.6 ml Sodium Pyruvate, 560 µl β-mercaptoethanol	10 % DMSO, 10 % medium, 80 % FCS
primary cell isolation	medium with 100 U/ml penicillin, 0.1 mg/ml streptomycin, 2.5 mg/ml amphotericin B for first 3 days	10 % DMSO, 40 % medium, 50 % FCS

2.1.7 Kits

Table 11 Kits

Agilent DNA 1000 Kit, 5067-1504	Agilent, Santa Clara, USA
Agilent High Sensitivity DNA Kit, 5067-4626	Agilent, Santa Clara, USA
Agilent RNA 6000 Nano Kit, 5067-1511	Agilent, Santa Clara, USA
Agilent RNA 6000 Pico Kit, 5067-1513	Agilent, Santa Clara, USA
AllPrep DNA/RNA Micro Kit and Mini Kit	Qiagen, Hilden, DEU
Basic Nucleofector Solution Primary Fibroblasts, VPI-1002	Lonza Group Ltd, Basel, CHE
Big Dye Terminator v1.1 Cycle Sequencing Kit, 4337450	Life Technologies, Darmstadt, DEU
Biotool DNA Transfection Reagent	Biotool.com/bimake.com
CloneJET PCR Cloning Kit	Thermo Fisher Scientific, Waltham, USA
Directzol RNA MiniPrep	Zymo Research, Tustin, USA
exoRNeasy Serum/Plasma Maxi Kit	Qiagen, Hilden, DEU
EZ DNA Methylation-Direct™ Kit	Zymo Research, Tustin, USA
Fast SYBR™ Green Master Mix	Applied Biosystems, Warrington, GBR
FirstChoice™ RLM-RACE Kit	Invitrogen GmbH, Darmstadt, DEU
GenElute Mammalian Genomic DNA Purification Kit	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
HiSeq Rapid PE Cluster Kit v2	Illumina, San Diego, USA
HiSeq Rapid SBS Kit v2 (200 cycles), FC-402-4021	Illumina, San Diego, USA
Human MSC Nucleofector® Kit	Lonza Group Ltd, Basel, CHE
iBind™ Solution Kit	Invitrogen GmbH, Darmstadt, DEU
innuPREP RNA Mini Kit	Analytik Jena AG, Jena, DEU
KAPA SYBR FAST Master Mix Universal 2X, KK4602	Kapa Biosystems, Wilmington, USA
Lipofectamine™ 2000 Transfection Reagent	Invitrogen GmbH, Darmstadt, DEU
miScript II RT Kit	Qiagen, Hilden, DEU
miScript SYBR® Green PCR Kit	Qiagen, Hilden, DEU
MiSeq Reagent Kit v2 (50-cycles), MS-102-2001	Illumina, San Diego, USA
pGEM®-T Easy Vector System	Promega Corporation, Madison, USA
Plasmid DNA purification NucleoBond® Xtra Midi	MACHEREY-NAGEL GmbH & Co. KG, Düren, DEU
PyroMark Q48 Advanced CpG Reagents, 974022	Qiagen, Hilden, DEU
QIAamp DNA Micro Kit	Qiagen, Hilden, DEU
QuantiFluor® RNA System	Promega Corporation, Madison, USA
Qubit® RNA BR Assay Kit, Q10210	Thermo Fisher Scientific, Waltham, USA
Qubit™ dsDNA BR Assay Kit, Q32850	Thermo Fisher Scientific, Waltham, USA
Qubit™ dsDNA HS Assay Kit, Q32854/1	Thermo Fisher Scientific, Waltham, USA
Qubit™ RNA HS Assay Kit, Q32852	Thermo Fisher Scientific, Waltham, USA
REPLI-g Mini Kit	Qiagen, Hilden, DEU
RNeasy Mini	Qiagen, Hilden, DEU
TruSeq RNA Library Preparation Kit v2, Set A	Illumina, San Diego, USA
TruSeq Small RNA Library Prep Kit, Set-A	Illumina, San Diego, USA
TURBO DNA-free™ Kit	Invitrogen GmbH, Darmstadt, DEU

VenorGeM Mycoplasma Detection Kit, MP0025	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Wizard® SV Gel and PCR Clean-Up System	Promega Corporation, Madison, USA

2.1.8 Enzymes

Table 12 Enzymes

AccuStart Taq DNA Polymerase HiFi	Quantabio, Beverly, USA
CollagenaseType I-A, C2674	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
DNA Polymerase I, Large Klenow Fragment	New England Biolabs, Frankfurt, DEU
Exonuclease I	New England Biolabs, Frankfurt, DEU
FastGene® Optima HotStart ReadyMix	NIPPON Genetics Europe, Dueren, DEU
PCR Extender System	5Prime GmbH, Hamburg, DEU
Proteinase K	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
PyroMark PCR Kit	Qiagen, Hilden, DEU
Q5® High-Fidelity DNA Polymerase	New England Biolabs, Frankfurt, DEU
Quickextract	Epicentre, Madison, USA
Restriction enzymes	New England Biolabs, Frankfurt, DEU
RiboLock Rnase Inhibitor, EO0381	Thermo Fisher Scientific, Waltham, USA
RNase A	Sigma-Aldrich Chemie GmbH, Steinheim, DEU
Shrimp alkaline phosphatase	New England Biolabs, Frankfurt, DEU
SssI, CpG methyl transferase	New England Biolabs, Frankfurt, DEU
SuperScript™ II, III and IV Reverse Transcriptase	Invitrogen GmbH, Darmstadt, DEU
T4 DNA Ligase	New England Biolabs, Frankfurt, DEU

2.1.9 Oligonucleotide primers

All primers were ordered from Eurofins Genomics, Ebersberg, Germany.

Table 13 Primer list

Primer name	Sequence 5'-3'	Length
3'RACE Adapter	GCGAGCACAGAATTAATACGACTCACTATAGGTNNNNNNNN NNNN	45
5'RACEInnerPrimer	CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG	35
5'RACEOuterPrimer	GCTGATGGCGATGAATGAACACTG	24
BACH1_Ex8_for1	GCGTGTGTGATTAGCTTGGGA	21
BACH1_Ex9_rev1	GCGTTAAATGGCAGTTTCACCT	22
Cas9_3'LR_for1	GCAGATCAGCGAGTTCTCCA	20
Cas9_Probe_for1	AAAGACCGAGGTGCAGACAG	20
Cas9_probe_rev1	CGGTCGATGGTGGTGTCAAA	20
Cas9_RT_for2	AAAGACCGAGGTGCAGACAG	20
Cas9_RT_rev2	AAGCCGCCGTACTTCTTAGG	20
CRISPR_APC_F1	AGCCAGCTCCATCCAAGTTC	20
CRISPR_APC_R1	CTTGGTGGCATGGTTTGTCC	20

CRISPR_PTEN_F1	CTGAGGAGAAGCAGGCCC	18
CRISPR_PTEN_R1	GCTCACTAATCCACTAACACTGT	23
CYP7A1_AS_Ex3_F6	TCCCGAACCAGGTTTGTTTG	20
CYP7A1_AS_Ex4_R6	TGAATCACCTGCAAACTTCCG	21
CYP7A1_CpG1_F1	AGTTGGAATTATAGTTGTTAGTTTATGAT	29
CYP7A1_CpG1_R1_BIO	[BIO]-CCTTCCACACTTTAATTCTATACAC	25
CYP7A1_CpG1_S2	GTTTATGATATAGATA	16
CYP7A1_CpG2_F1	AAGAAGTGATATATGTAGAGGAAAGATAG	29
CYP7A1_CpG2_R1_BIO	[BIO]-ATAACTCCAAAAAAACTTCTTAAATCTTAC	30
CYP7A1_CpG2_S	GTGTGTTTTTTGGGT	15
CYP7A1_Ex4_ASV_F1	TGTTTGCTTTGGTCACTCAAGT	22
CYP7A1_Ex4_ASV_F2	CTGGAGCCTCTGTTGTGACG	20
CYP7A1_Ex4_F7	GGGGATTGTGACAGCAGTG	19
CYP7A1_Ex5_F1	TGTCCACTTCATCACAAATCCCT	23
CYP7A1_Ex5_R7	GCAGTGCACAACCCAGATAG	20
CYP7A1_Ex6_F2	TGACGCAGAGAAAGCCAAGT	20
CYP7A1_Ex6_R1	ATGCTTCTGTGCCCAAATGC	20
CYP7A1_Ex6_R1-2	GGTCAATGCTTCTGTGCCCA	20
CYP7A1_Ex7_F3	AAAACACTGGAGAAGGCGGG	20
CYP7A1_Ex7_R2	TGCTTTCATTGCTTCAGGGC	20
CYP7A1_Ex8_F4	CCACAATTAATGCACCTGGATCC	23
CYP7A1_Ex8_R3	GGAAAGCCTCAGAGACTCCTT	21
CYP7A1_Ex9_R4	CCATGACTGTAGAAGGTGGTCT	22
CYP7A1_SNPSeq_F1	GACCCAGCAAATCCACCCTT	20
CYP7A1_SNPSeq_F10	TTGAAACATGAAGCACAGAAACA	23
CYP7A1_SNPSeq_F2	TAATCCCAACACGACCCCTC	20
CYP7A1_SNPSeq_F3	TCGATACTAAGCCCGGTCCA	20
CYP7A1_SNPSeq_F4	AGGGGTGTGATAGATGCCATG	21
CYP7A1_SNPSeq_F5-2	ACCTACACCACAGCTCACAG	20
CYP7A1_SNPSeq_F6	TCTGTTTAAGGAGGCAAGAATCA	23
CYP7A1_SNPSeq_F7	CGATGGCCAGTTTCGTTGTC	20
CYP7A1_SNPSeq_F8	TCTGCCTGGAGTTCTCTTCCT	21
CYP7A1_SNPSeq_F9	GTTCCACTGCACCACAACG	19
CYP7A1_SNPSeq_R	GGAGGGAAGGCTGGACTTTT	20
CYP7A1_SNPSeq_R1	GTCACACCAGCTGTTTTCTTGA	22
CYP7A1_SNPSeq_R2	GCACGAGGAAGCCAGGAG	18
CYP7A1_SNPSeq_R3	ACACAACGAAAGCCCAGGAA	20
CYP7A1_SNPSeq_R4	CTGCAACTTCCTGTGACTCTATA	23
CYP7A1_SNPSeq_R5	CCTCTTCATTTCTTACGTGTGCA	23
CYP7A1_SNPSeq_R6	GGAAAACGCACAGGAAGCAA	20
CYP7A1_SNPSeq_R7	TGGGAGATGAGAGTGATGGGA	21
CYP7A1_SNPSeq_R8	CCTCCTCTATTACTGCTCACTCA	23
CYP7A1_SNPSeq_R9	CTGTACAGGATCAACATCTCACA	23
DCC_CRISPR_F1	TGAGGGCATTTACAAAGGAGAG	22
DCC_CRISPR_R1	CGGAAGCTATTGTTGAATCAGC	22

ddCas9 F1	AGTTCATCAAGCCCATCCTG	20
ddCas9 R1	TCTTTTCCCGGTTGTCCTTC	20
ddGAPDH F Ya	CCGCGATCTAATGTTCTCTTTC	22
ddGAPDH R Ya	TTCACTCCGACCTTCACCAT	20
ddpoactin Ya F1	TAACCGATCCTTTCAAGCATTT	22
ddpoactin Ya R1	TGGTTTCAAAGCTTGCATCATA	22
ddpoGAPDH F1	CTCAACGACCACTTCGTCAA	20
ddpoGAPDH R1	CCCTGTTGCTGTAGCCAAAT	20
GAPDH_Ex4_for1	GTTGTGGATCTGACCTGCCGC	21
GAPDH_Ex5_rev1	TCAGTGTAGCCCAGGATGCCC	21
GS_MCS_Seq_for1	GGGCAGTTTCGAAGATCG	18
GS_MCS_Seq_rev1	TGGTGTAGTTGTGGGCTGAA	20
let-7a-5p_1	TGAGGTAGTAGGTTGTATAGTT	22
MCS_colonyPCR_F	CCACCGGTGTCGCGATTAATTA	22
MCS_colonyPCR_F	CCACCGGTGTCGCGATTAATTA	22
mGFP_end_F	ACTCTCGGCATGGACGAG	18
mGFP_start_R1	TACCTTCACGTGGCCATTCT	20
miR-16-5p_1	TAGCAGCACGTAAATATTGGCG	22
miR-191-5p_1	CAACGGAATCCCAAAAGCAGCTG	23
miR-25-3p_1	CATTGCACTTGTCTCGGTCTGA	22
miR-26a-5p_1	TTCAAGTAATCCAGGATAGGCT	22
mir-425-5p_1	AATGACACGATCACTCCCGTTG	22
mTom_end_F1	CTGTTCCTGTACGGCATGGA	20
mTom_end_R1	CTTGTACAGCTCGTCCATGC	20
Myco_1F	GGAGCAAACAGGATTAGATACC	22
Myco_1R	CACCATCTGTCATTCTGTTAACC	23
Neo_2pA_Age1_rev1	TAGACCGGTTTACTAGTCCCCAGCATGCC	29
Neo_colony_R1	CATCAGAGCAGCCGATTGTC	20
Neo_HindIII_for1	ATCCGATAAGCTTGATCCGGG	21
qPCR primer 1.1	AATGATACGGCGACCACCGAGAT	23
qPCR primer 2.1	CAAGCAGAAGACGGCATACGA	21
R26endoR	GTTTGCACAGGAAACCCAAG	20
Random hexamer primers	NNNNNN	6
RNU6B_1	ACGCAAATTCGTGAAGCGTT	20
R26_SA_Seq_R1	GGGGCCTAAGGTTTGGAGAT	20
Rosa26 E1 F1	CGCCTAGAGAAGAGGCTGTGC	21
Rosa26 I1 F2	TATGGGCGGGATTCTTTTGC	20
Rosa26 I1 R3	GTTTGCACAGGAAACCCAAG	20
Rosa26 I3 R2	CAGGTGGAAAGCTACCCTAGCC	22
Rosa26 Loc2R	GGAGCGGCGATACCGTAAAG	20
Rosa26 Loc3R	TCCGAAGCCCAACCTTTCAT	20
Rosa26_colony_F1	ATTCCCAGTCCCTCACACAG	20
ROSA26_SAR	GAAAGACCGCGAAGAGTTTG	20
SATB1_Ex6_F3	GCAAATGTCTCAGCGGCAAA	20
SATB1_Ex7_R3	CCTTGCTGGGATAGTTCGGA	20

SFRP5_Ex2_F1	TGCGCCCAGTGTGAGATG	18
SFRP5_Ex3_R1	CTTCCGGTCCCCACTCTCTA	20
ssc-miR-146a-5p	TGAGAACTGAATTCCATGG	19
ssc-miR-194b-5p	TGTAACAGCGACTCCATGTGGA	22
ssc-miR-215	ATGACCTATGAATTGACAGAC	21
ssc-miR-27a-3p	TTCACAGTGGCTAAGTTCCG	20
TP53_CRISPR_F2	ACCCTGGTCCCAAAGTTGAA	20
TP53_CRISPR_R2	GCCCGTAAATTCCCTTCCAC	20
Universal_Primer	CGAATTCTAGAGCTCGAGGCAGG	23

2.1.10 Oligonucleotides for hybridisation

Oligos for hybridisation, to generate tailored short double stranded DNA fragments, were ordered from Eurofins Genomics, Ebersberg, Germany or biomers.net GmbH, Ulm, Germany.

Table 14 CRISPR oligos

Oligo name	Sequence 5'-3'
CRISPR_APC_1A	CACCGCTAATGTGCTGTGGCACTGT
CRISPR_APC_1B	AAACACAGTGCCACAGCACATTAGC
CRISPR_DCC_Ex1_1A	CACCGTTTCTATCTCACCCCTA
CRISPR_DCC_Ex1_1B	AAACTAGGGGTGAGATAGAAAC
CRISPR_TP53_Ex5_1A	CACCGGCAAAACAGCTTATTGA
CRISPR_TP53_Ex5_1B	AAACTCAATAAGCTGTTTTGCC
CRISPR-pX300-PTEN-A	CACCGAGATCGTTAGCAGAAACAAA
CRISPR-pX300-PTEN-B	AAACTTTGTTTCTGCTAACGATCTC
DCC_Ex1-1_A	GTCGGTTTCTATCTCACCCCTACGG
DCC_Ex1-1_B	CGGTCCGTAGGGGTGAGATAGAAAC
PTEN_cCheck_A	GTCGGAGATCGTTAGCAGAAACAAAAGG
PTEN_cCheck_B	CGGTCCTTTTGTTTCTGCTAACGATCTC
TP53_Ex5_A	GTCGGGCAAAACAGCTTATTGAGGG
TP53_Ex5_B	CGGTCCCTCAATAAGCTGTTTTGCC

Table 15 Multiple cloning site oligos

Oligo name	Sequence 5'-3'	Length
MCS_RosaNeoCas_A	TCGAGCTGAGTTTAAACTGCAGGCCGGCCGATCTTCGAAACTGCCCG GGCGATCCTAGGTACGC	64
MCS_RosaNeoCas_B	TCGAGCGTACCTAGGATCGCCCGGGCAGTTTCGAAGATCGGCCGGCC TGCAGTTTAAACTCAGC	64

2.1.11 Cloning vectors

Table 16 Plasmid list

Plasmid	Antibiotic resistance	Source
pGEM-T Easy	Ampicillin	Promega Corporation, Madison, USA
pJET1.2/blunt	Ampicillin	Invitrogen GmbH, Darmstadt, DEU
pSL1180	Ampicillin	Amersham Biosciences, Piscataway, USA
pCAG-Cas9-bpA	Ampicillin	Dr. Oskar Ortiz Sanchez, IDG Helmholtz Center, DEU
pGEMT-Rosa26-Cas9-BS	Ampicillin	Chair of livestock biotechnology, TUM, DEU
pBs-LSL-Neo	Ampicillin	Chair of livestock biotechnology, TUM, DEU
pSL1180 + MR26 MTMG	Ampicillin	Chair of livestock biotechnology, TUM, DEU
p119-c-check	Spectinomycin	Yonglun Luo (Alun), MSc, Ph.D, Associate Professor, Department of Biomedicine, Aarhus University, DNK
pX330-U6-Chimeric_BB-CBh-hSpCas9, 42230	Ampicillin	Addgene, Cambridge, USA
pX330-U6-Chimeric_BB-CBh-hSpCas9_MCS	Ampicillin	Chair of livestock biotechnology, TUM
pSL1180-SA-puro-LA	Ampicillin	Chair of livestock biotechnology, TUM

2.1.12 Antibodies

Table 17 List of primary antibodies

Primary antibodies	Source	Species reactivity	Description	Western blot dilution	iBind dilution
Mouse anti-GAPDH, (monoclonal), G8795	Sigma-Aldrich Chemie GmbH, Steinheim, DEU	rat, chicken, turkey, human, mouse, mink, bovine, canine, monkey, rabbit, hamster	immunized with rabbit GAPDH	1:3000	1:3000
Cas9 (7A9-3A3) Mouse mAb #14697	Cell Signaling Technology, Danvers, USA	-	N- terminus of Cas9 from *S. pyogenes*	1:1000	1:1000
Anti-Digoxigenin	Roche Diagnostic GmbH, Mannheim, DEU	-		Southern blot dilution 1:10000	

Table 18 List of secondary antibodies * when two primary antibodies were used simultaneous

Secondary Antibodies	Source	Western blot dilution	iBind dilution
Goat anti-rabbit IgG, (polyclonal), A9169	Sigma-Aldrich Chemie GmbH, Steinheim DEU	1:5000	
Rabbit Anti-Mouse IgG H&L (HRP), (polyclonal), Ab6728	Abcam, Cambridge, England	1:5000	1:1000/1:500*

2.1.13 Competent bacterial cells

Table 19 Electro competent cells

E. Coli DH10B ElectroMAX	Genotpye: F− mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ− rpsL nupG	Invitrogen GmbH, Darmstadt, DEU

2.1.14 Cultured mammalian cells

Table 20 List of cultured mammalian cells

Cell lines	Organism	Description	Source
pADMSCs 110111	pig	porcine adipose-derived mesenchymal stem cells	11.01.11 isolated by Dr. Benedikt Baumer, Chair of livestock biotechnology, TUM
pKF 270	pig	porcine kidney fibroblasts carrying a mT/mG cassette in the Rosa26 locus	isolated by Dr. PD. Tatiana Flisikowska, Chair of livestock biotechnology, TUM
pKF 73 *APC*^{1311} 240415	pig	porcine kidney fibroblasts carrying a heterozygous *APC*^{1311} mutation	24.04.15 isolated by Dr. PD. Tatiana Flisikowska, Chair of livestock biotechnology, TUM
HEK293	human	1973 transformed embryonic kidney cells using DNA fragments of human Adenovirus 5 (Graham et al. 1977)	Chair of Nutrition and Immunology, TUM

2.1.15 Pigs

Table 21 List of individual pigs analysed

* Animals with more than 100 polyps were classified as high polyp (HP) and animals with 1-10 polyps as low polyp (LP) animals. Pigs with polyp numbers between 10 and 100 were classified as medium polyp (MP).

Animal ID	Born	Sex	Pheno-type*	Animal ID	Born	Sex	Pheno-type*
66	13.04.2012	male	HP	534	18.01.2016	female	LP
71	13.04.2012	female	HP	586	11.03.2016	female	LP
73	13.04.2012	male	HP	588	11.03.2016	female	HP
128	07.03.2013	female	LP	591	15.03.2016	castrated male	LP

145	14.03.2013	male	LP	598	15.03.2016	male	HP
148	14.03.2013	male	MP	722	26.09.2016	male	HP
150	14.03.2013	female	LP	727	26.09.2016	female	LP
152	14.03.2013	female	LP	388	04.04.2015	male	
153	14.03.2013	female	LP	909	18.06.2017	castrated male	HP
155	14.03.2013	male	LP	910	18.06.2017	male	HP
157	14.03.2013	female	LP	911	18.06.2017	castrated male	LP
163	15.03.2013	female	HP	913	18.06.2017	castrated male	LP
168	12.05.2013	male	HP	914	18.06.2017	castrated male	HP
173	12.05.2013	female	LP	916	18.06.2017	female	LP
174	12.05.2013	female	LP	917	18.06.2017	female	LP
251	21.02.2014	castrated male	HP	918	18.06.2017	female	HP
252	21.02.2014	castrated male	HP	919	18.06.2017	male	LP
253	21.02.2014	castrated male	HP	920	18.06.2017	female	LP
300	24.08.2014	female	LP	921	18.06.2017	female	LP
322	18.10.2014	male	HP	929	18.06.2017	male	LP
324	18.10.2014	male	HP	932	18.06.2017	castrated male	LP
326	18.10.2014	female	HP	933	18.06.2017	female	LP
328	18.10.2014	female	LP	937	09.07.2017	male	HP
339	20.10.2014	female	HP	941	09.07.2017	female	HP
471	17.08.2015	female	LP	943	09.07.2017	female	LP
474	17.08.2015	female	HP	944	09.07.2017	female	LP
524	18.01.2016	castrated male	LP	952	09.07.2017	female	LP
525	18.01.2016	castrated male	HP	953	09.07.2017	female	LP
527	18.01.2016	male	HP				

2.1.16 Computer software

Table 22 Software

ELISA-Reader	Ascent Software, Luqa, Malta
PyroMark Q48 Autoprep Software	Qiagen, Hilden, DEU
7500 Software v2.0.5	Applied Biosystems, Warrington, GBR
A Plasmid Editor	M. Wayne Davis
AB Sequencing Analysis Software (v5.2)	Applied Biosystems, Warrington, GBR
Axiovision	Carl Zeiss Microscopy GmbH, Göttingen, DEU
AxioVision Rel. 4.8	Carl Zeiss Microscopy GmbH, Göttingen, DEU
Benchling	https://benchling.com/
CASAVA BCL2FASTQ Conversion Software 1.8.3	Illumina, San Diego, USA
CRISPR design	http://crispr.mit.edu/, Zhang Lab, MIT 2017

DeSeq2	(Love et al, 2014), https://bioconductor.org/-packages/release/bioc/html/DESeq2.html
EdgeR	(McCarthy et al, 2012; Robinson et al, 2010)
Every vector	http://www.everyvector.com/users/login
FASTQC	Andrews S. (2010), http://www.bioinformatics.-babraham.ac.uk/projects/fastqc
FeatureCounts	(Liao et al, 2014), http://bioinf.wehi.edu.au-/featureCounts/
Genecards	http://www.genecards.org/
gNorm	(Vandesompele et al, 2002)
GSEA software	(Mootha et al, 2003; Subramanian et al, 2005)
HiSeq Control Software 2.2.58	Illumina, San Diego, USA
IGV	(Robinson et al, 2011; Thorvaldsdottir et al, 2013), http://software.broadinstitute.org/-software/igv/home
Image J	Nationla Institutes of Health, Bethesda, USA
Kallisto	(Bray et al, 2016)
Leica Application Suite software	Leica Microsystems, Wetzlar, DEU
MatInspector, Genomatix Matrix Library 11.0	https://www.genomatix.de/index.html (Cartharius et al, 2005; Quandt et al, 1995)
mirPath v.3	(Vlachos et al, 2015b)
Normfinder	(Andersen et al, 2004)
Phred/Phrap/Polyphred/Consed-Software	(Ewing & Green, 1998; Ewing et al, 1998) (Gordon et al, 1998; Nickerson et al, 1997)
Picard	http://broadinstitute.github.io/picard
Primer 3	http://primer3.ut.ee/
PyroMark Assay Design2.0	Qiagen, Hilden, DEU
R Studio Version 1.0.153	http://www.rstudio.com/
R version 3.4.1 (2017-06-30)	https://www.R-project.org/, (https://cran.r-project.org/)
Real-Time Analysis (RTA) 1.18.64	Illumina, San Diego, USA
Run 3130xl Data Collection v.3.0	Applied Biosystems, Warrington, UK
Sleuth	(Pimentel et al, 2017), pachterlab.github.io/sleuth/
STAR	(Dobin et al, 2013)
TarBase v7.0	(Vlachos et al, 2015a)
TargetScan	(Lewis et al, 2005)
TIDE: Tracking of Indels by Decomposition	https://tide.deskgen.com/, Brinkman
Vector NTI	Invitrogen GmbH, Darmstadt, DEU
VISION	VILBER LOURMAT Deutschland GmbH, Eberhardzell, DEU

2.2 Molecular biological methods

2.2.1 Isolation of bacterial plasmid DNA

Mini prep

2 ml bacterial D10Hb cultures were pelleted for 5 min at full speed (12300-15493 x g). The supernatant was discarded and the bacterial pellet resuspended in 100 µl solution 1 (5 mM sucrose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0). 200 µl solution 2 (0.2 M NaOH, 1 % SDS) was added for alkaline lysis and inverted 6-8 times for complete mixture. After 3 min incubation at room temperature (RT) 150 µl neutralising solution 3 (3 M sodium acetate pH 4.8) was added and inverted. After 30 min incubation on ice, the solution was centrifuged for 5 min. The supernatant, containing the plasmid DNA was mixed with 1 ml 95 % ethanol for plasmid DNA precipitation by 15 min centrifugation. The pellet was first washed with 500 µl 80 % ethanol, centrifuged for 10 min and subsequently with 500 µl 95 % ethanol. After aspiration of the supernatant the DNA pellet was dried using a vacuum centrifuge for 1-2 min at medium drying stage and resuspended in 50 µl water containing RNase A (40 µg/ml).

Midi and Maxi prep

Plasmid DNA was isolated from 100 ml or 300 ml bacterial cultures with the Plasmid DNA purification NucleoBond® Xtra Midi Kit and Maxi Kit (Macherey-Nagel) respectively, following manufacturer's information. The resulting DNA pellet was eluted in 100 or 200 µl TE buffer for Midi and Maxi preps.

2.2.2 Isolation of mammalian genomic DNA using phenol-chloroform extraction

Isolation from tissue) The tissue piece was cleaned with ethanol and PBS before transferring it to a petri dish and cutting it into small pieces with a sterile scalpel. The petri dish was rinsed with 500 µl lysis buffer (100 mM Tris-HCl pH7.4, 0.2 % SDS, 5 mM EDTA, 0.2 M NaCl) or more (depending on tissue size) to convert the minced tissue to a falcon tube.

Isolation from cells) Cells were pelleted in a 15 ml falcon tube by centrifugation at 300 x g for 5 min. The cell pellet was resuspended in 500 µl lysis buffer or more (depending on pellet size).

At this stage tissue samples and cultured cells were treated in a similar manner by supplementation with proteinase K to a final concentration of 100 µg/ml and incubation at 37 °C over night.

The next day, RNase A was added to a final concentration of 390 µg/ml and incubated for 5 min. The solution was mixed with an equal volume of phenol-chloroform-isoamyl alcohol, vortexed and left at RT for 10 min before centrifugation (4°C, 10 min, 15493 x g). The top aqueous phase, containing DNA, was transferred to a new reaction tube. An equal volume of chloroform was added, mixed, incubated

and centrifuged as before. The top aqueous phase was collected and mixed with 0.7x volume ice cold isopropanol. The DNA was pelleted at 4 °C for 15 min and washed with 200 µl ice cold 70 % ethanol and centrifuged for 5 min. The ethanol was aspirated and the DNA pellet dried for 10-15 min. When dried completely, the pellet was resuspended in 50-100 µl TE Buffer.

2.2.3 Isolation of mammalian genomic DNA using AllPrep Mini Kit

A whole biopsy sample was placed in a lysis tube P with 600 µl RLT Plus buffer (containing β-mercaptoethanol). The samples were homogenised for 30-120 sec in the precooled Speed Mill PLUS, until complete homogenisation. The isolation was performed using the AllPrep Mini Kit (Qiagen) according to manufacturer's information and gDNA was eluted in 100 µl nuclease-free water.

2.2.4 Isolation of mammalian genomic DNA using Quick Extract

Small cell amounts were pelleted and frozen for later, or resuspended right away in 20 µl QuickExtract (Epicentre) per 1/3 12 well. Enzymatic cell lysis was performed at 68 °C for 15 min and 95 °C for 8 min.

2.2.5 Isolation of RNA

During the isolation and purification of RNA, samples were always kept on ice and centrifuged at 4 °C.

2.2.5.1 RNA isolation from cells

Cells were washed with ice cold PBS and detached by pipetting with 350 µl RLT buffer per confluently covered 12 well. Further processing was performed according to protocol step 4 of RNeasy Kit (Qiagen). All centrifugations were run at 10000 x g and RNA was eluted in 30 µl nuclease-free water.

2.2.5.2 RNA isolation from tissues

Biopsy samples were kept on liquid nitrogen and halved before they were placed into a Lysis tube P.

<u>mRNA Isolation from tissues using innuSPEED Tissue RNA Kit</u>

Biopsy samples were transferred to a lysis tube P with 450 µl lysis buffer (supplied) for at least 2x 20 seconds homogenisation using previously cooled Speed Mill PLUS and processed according to manufacturer's information. RNA was eluted in 30 µl nuclease-free water.

<u>Total RNA isolation from tissues using DirectZol RNA MiniPrep</u>

Biopsy samples were placed into a Lysis tube P containing 400 µl Trizol for at least 2x 20 sec homogenisation using previously cooled Speed Mill PLUS. They were spun down for 1 min at 12000 x

g before 350 µl cell lysate was mixed with 350 µl 100 % ethanol. Isolation was performed according to manufacturer's information and RNA was eluted in 30 µl nuclease-free water.

Total RNA isolation from tissues using AllPrep Mini Kit

Biopsy samples were transferred to a lysis tube P with 600 µl RLT Plus buffer (containing β-mercapto-ethanol). for at least 30 sec homogenisation using precooled Speed Mill PLUS. Isolation was performed according to manufacturer's information and total RNA was eluted in 30 µl nuclease-free water.

2.2.5.3 Total RNA isolation from laser microdissected cryo sections

Laser microdissected samples mixed with RLT Plus buffer (containing β-mercaptoethanol) were vortexed for 30 -120 sec. Total RNA was isolated using the Allprep DNA/RNA Micro Kit (Qiagen) according to manufacturer's information and eluted in 14 µl nuclease-free water.

2.2.6 DNase digest

Isolated RNA was treated with TURBO DNA-free™ Kit according to the manufacturer's information except no inactivation reagent was used.

2.2.7 Quantification and Quality control of nucleic acids

2.2.7.1 Determination of nucleic acid concentration using NanoDrop Lite

For dsDNA the NanoDrop Lite multiplied the absorbance at 260 nm by 50 and for RNA by 40. The A260/A280 ratio indicating protein contamination should lie around 1.8 for dsDNA and 2.0 for RNA.

2.2.7.2 Determination of nucleic acid concentration using Qubit 2.0 fluorometer

RNA was fluorometrically quantified using the QuantiFluor® RNA System (Promega) (detection range 0.1–500ng) and DNA using Qubit™ dsDNA BR Assay Kit (Thermo Fisher Scientific) (detection range 2-1000 ng) and Qubit™ dsDNA HS Assay Kit (Thermo Fisher Scientific) (detection range 0.2-100 ng) according to manufacturer's information.

2.2.7.3 Determination of nucleic acid quality using Bioanalyzer

The RNA integrity number (RIN), was determined using Agilent RNA 6000 Nano Kit and Agilent RNA 6000 Pico Kit for laser microdissected RNA isolations according to manufacturer's information.

Size distribution and quality of sequencing libraries were determined using Agilent DNA 1000 Kit and Agilent High Sensitivity DNA Kit according to manufacturer's information.

2.2.7.4 Gel electrophoresis

For analytical gels of RNA or DNA fragments after isolation, digestion or PCR the agarose was dissolved in 1x TBE. For preparative gels (excision of DNA) 1x TAE was used. Gels were always supplemented with a final concentration of 800 x peqGREEN. RNA gels were additionally complemented with 400 µl 37 % formalin per 50 ml gel. Both DNA and RNA samples were mixed with gel loading dye prior to gel electrophoresis, the latter were denatured at 70 °C for 2 min to destroy secondary structures.

2.2.8 Plasmid DNA purification for tissue culture by ethanol precipitation

The volume of plasmid DNA was set to 500 µl with water and 1/10 volume of 3 M sodium-acetate pH 5.2 was added and mixed. A 2-2.5 x volume of cold 100 % ethanol was added, mixed and incubated at -20 °C for at least 20 min before pelleting the DNA (12300-15493 x g, 10 min, 4 °C). From now on, the tube containing the DNA was only opened under a sterile laminal flow cabinet. The supernatant was aspirated and the pellet washed with 1 ml sterile 70 % ethanol. After centrifugation (full speed, 5 min, RT) the supernatant was aspirated and the pellet air dried. Plasmid DNA was eluted in sterile water or sterile low TE buffer, to gain a final concentration of 1-1.5 µg/µl.

2.2.9 Plasmid DNA purification for tissue culture by phenol-chloroform extraction

The volume of the plasmid DNA was set to 500 µl with water and mixed with 500 µl phenol-chloroform-isoamyl alcohol. The procedure was performed according to 2.2.2 except, after chloroform addition the top aqueous phase was mixed with 1/20th volume 3 M Sodium-acetate and a 0.7 volume isopropanol. This mixture was shaken and pelleted at full speed for 10 min. From now on, the DNA tube was only opened under a sterile laminal flow cabinet and proceeded according to 2.2.8.

2.2.10 Column based DNA purification

DNA after restriction digest, blunting, ligation or PCR was purified using the Wizard® SV Gel and PCR Clean-Up System (Promega) according to manufacturer's information with modifications (Appendix).

2.2.11 Restriction enzyme digestion

1 µg plasmid DNA, PCR products and gDNA was digested with 3-5 Units restriction enzyme (NEB), ≤5 % enzyme glycerol concentrations for at least 1 - 4 h at the enzyme's temperature optimum for cloning, Southern blot, digital droplet PCR, and size determination. Digests after ligation were incubated for 20 min with 5-10 units restriction enzyme, water and restriction buffer to gain 1x concentration.

2.2.12 Blunting

DNA sticky ends were blunted using the DNA Polymerase I Large Fragment (Klenow) from NEB. Successful blunting was achieved with 1x concentrated NEB buffer (1.1, 2.1, 3.1 or CutSmart), a final concentration of 60 µM dNTPs and ≤ 1 unit of the Klenow Enzyme (5 U/µl) per 1 µg DNA incubated at 25 °C for 15 min and inactivated with a final concentration of 10 mM EDTA at 75 °C for 20 min.

2.2.13 Oligonucleotide hybridisation

Each single-stranded oligonucleotide was dissolved in a 1 µg/µl concentration in TE buffer. 1 µl of each complementary oligo was added to 98 µl TE buffer, heated for 5 min at 100 °C and slowly cooled to RT.

2.2.14 DNA ligation

Vector DNA ranged from 50 to 200 ng (for DNA > 10 kb) and insert was calculated as below. 30 ng double-stranded oligonucleotides were ligated to 50 ng vector without calculation. To the DNA 1x concentrated T4 Ligase buffer, 400-600 units of T4 DNA Ligase and water was added to a total volume of 20 µl and incubated at RT for 1 h or at 16 °C over night.

$$Insert_{ng} = \frac{Vector_{ng} * size_{Insert}}{size_{Vector}} * 3$$

2.2.15 DNA methylation

All CpG sites in a gDNA sample were methylated by incubating 1 µg DNA with 4.8 units CpG methyltransferase SssI, 160 µM S-adenosylmethionine (SAM) and 1x Buffer 2.1 (NEB) for 4 h at 37 °C. After each hour of incubation another 4.8 units of SssI were added.

2.2.16 Bisulphite conversion

Bisulphite conversion of 200 ng genomic DNA was performed using EZ DNA Methylation-Direct Kit (Zymo Research) according to manufacturer's information.

2.2.17 Whole genome amplification using the REPLI-g Mini Kit

To generate CpG methylation-free DNA, 150 ng gDNA was amplified using the REPLI-g Mini Kit (Qiagen) according to manufacturer's information.

2.2.18 Reverse Transcription

500 ng total RNA were converted to cDNA using SuperScript™ III Reverse Transcriptase (Invitrogen) with 100 pmol random hexamer primers according to manufacturer's information. Entire laser microdissected RNA was converted using Superscript IV Reverse Transcriptase (Invitrogen). For microRNA (miRNA) analysis, 1 µg total RNA was reverse transcribed using the miScript II RT Kit (Qiagen) according to the manufacturer's information.

2.2.19 5' Rapid amplification of cDNA ends (RACE)

5' RACE of 1 µg RNA was performed using the FirstChoice RLM-RACE Kit according to manufacturer's information with small modifications. Phenol:chloroform extraction with 15 µl supplied ammonium acetate solution, 115 µl nuclease free water and 150 µl phenol-chloroform-isoamyl alcohol was mixed and centrifuged following 2.2.2. After adding ice cold isopropanol, RNA was chilled on ice for 10 min. RNA pelleting for 20 min and washing with cold 70 % ethanol for 10 min was performed at full speed at 4 °C. The supernatant was discarded, the pellet air dried and resuspended in 4 µl Tobacco Acid pyrophosphate (TAP) buffer. The entire volume was mixed with 1 µl TAP enzyme and incubated for 1 h at 37 °C. The complete mix was used for adapter ligation according to manufacturer's information over night at 16 °C. The whole reaction was used for reverse transcription using SuperScript III (2.2.18). The resulting cDNA was used for nested PCR.

2.2.20 Polymerase chain reaction

Depending on the purpose of the amplification, different DNA polymerases were used requiring different PCR conditions. The polymerases AccuStart Taq DNA Polymerase HiFi and GoTaq® DNA Polymerase produce 3'A overhangs where Q5® High-Fidelity DNA Polymerase generates blunt ends.

AccuStart Taq DNA Polymerase HiFi PCR was performed with 20-200 ng DNA, 1x HiFi PCR Buffer, 2mM magnesium sulfate, 200 nM of each primer, 200 µM dNTPs and 0.02U/µl AccuStart Taq DNA Polymerase HiFi in 50 µl total volume. Thermal cycling conditions were: 1 min 94 °C, 40 cycles of 20 sec 94 °C, 30 sec 55-65 °C (depending on primer annealing temperature) and 1min/kb 68 °C.

GoTaq® DNA Polymerase PCR was performed with 50-300 ng DNA, 1x Green GoTaq® Reaction Buffer (1.5 mM $MgCl_2$), 200 nM primer each, 200 µM dNTPs and 0.025 U/µl GoTaq® DNA Polymerase in 50 µl total volume. Cycling conditions were: 2 min 95 °C, 40 cycles of 30 sec 95 °C, 30 sec 42-65 °C and 1 min/kb 72 °C and 5 min 72 °C.

Q5® High-Fidelity DNA Polymerase PCR was performed with approx. 200 ng DNA, 1x Q5 Reaction Buffer, 1x Q5 High GC Enhancer, 500 nM of each primer, 200 µM dNTPs and 0.01U/ µl Q5® High-Fidelity DNA Polymerase in 50 µl total volume. Thermal cycling conditions were: 30 sec 98 °C, 35 cycles of 10 sec 98 °C, 30sec 50-72 °C and 1 min/ kb 72 °C and 2 min 72 °C.

PyroMark PCR Kit PCR was performed with ≤ 500 ng DNA or 10-20 ng bisulphite converted DNA (bcDNA), 1x PyroMark PCR Master Mix, 1x Coral Load, 0.5 mM $MgCl_2$, 200 nM of each primer in 25 µl total volume. Cycling conditions were: 15 min 95 °C, 45 cycles of 30 sec 94 °C, 30 sec 60 °C for DNA/ 56 °C for bcDNA and 30 sec 72 °C and a final 10 min 72 °C.

2.2.21 Colony PCR using GoTaq Polymerase

Colony PCR was performed to screen replicate bacterial colonies for correctly cloned DNA constructs. GoTaq® DNA Polymerase PCR was assembled as above, without DNA, as the colonies were picked into the reaction mix and onto an agar plate for culture. Initial denaturation was extended to 5 min 95 °C.

2.2.22 Mycoplasma Test PCR using GoTaq

Medium conditioned by the cells for at least three days, was heated to 95 °C for 5 min. 2 µl was added to GoTaq polymerase PCR with the primers Myco_1F and Myco_1R. Primer concentrations were increased to 500 nM and $MgCl_2$ was added to a final concentration of 1.5 mM in 25 µl total volume. Thermal conditions were: 2 min 94 °C, 40 cycles of 30 sec 94 °C, 30 sec 55 °C and 30 sec 72 °C. Non-infectious DNA-fragments of *Mycoplasma orale* genome from the VenorGeM Mycoplasma Detection Kit (Sigma-Aldrich Chemie GmbH) served as positive control.

2.2.23 Reverse Transcription PCR

The reverse transcription PCR (RT-PCR) is performed using cDNA as a template to confirm transcription. 1 µl undiluted cDNA was used, also in the case of RACE analysis.

2.2.24 Quantitative Real-time PCR

Quantification of sequencing libraries for mRNA sequencing was performed using KAPA SYBR FAST qPCR Master Mix (2X) Universal (Kapa Biosystems) according to Sequencing Library qPCR Quantification Guide (Illumina) with modifications. 5 nM libraries were diluted 1:1000 and vortexed (this and all following dilutions were prepared with 0.1 % (v/v) Tween 20). One previously sequenced 2 nM library served as a positive control and was diluted 1:500 and another was used for standard curve generation by preparing serial dilutions of 20 pM, 2 pM, 0.2 pM, 0.02 pM and 0.002 pM. All dilutions were prepared in triplicates. 4 µl of each dilution, 1x KAPA SYBR FAST qPCR Master Mix, 0.2 µM primer each were mixed in a total volume of 10 µl. Thermal conditions were: 2 min 50 °C, 5 min 95 °C and 30 cycles of 30 sec 95 °C and 45 sec 60 °C.

2.2.25 Reverse transcription quantitative Real-time PCR

Reverse transcription quantitative real-time PCR (RT-qPCR) of mRNA was performed using Fast SYBR™ Green Master Mix (Applied Biosystems) according to the following protocol (Table 23).

Table 23 RT-qPCR reaction set up using Fast SYBR Green Master Mix

Primer Master Mix		**Sample Master Mix**	
Components	Final concentration	Components	Amount [µl]
2xFast SYBR® Green Master Mix	1x (5 µl)	cDNA	3
Forward primer	0.2 µM	Primer Master Mix	30
Reverse primer	0.2 µM		
H_2O	up to 10 µl		

The Sample Master Mix was transferred into 3 wells (technical triplicates), each 10 µl, of a MicroAmp Fast Optical 96-Well Reaction Plate. The plate was sealed with MicroAmp™ Optical Adhesive Film and pulse centrifuged. The reaction was run in the ABI 7500 Fast Real-Time PCR System, with cycling conditions: 20 sec 95 °C, 40-55 cycles of 3 sec 95 °C and 30 sec 60-64 °C. Subsequent high resolution melting (HRM) analysis (60-95 °C) of the products was performed determining melting temperature and number of specific and unspecific products.

RT-qPCR of miRNA was perfomed using miScript SYBR® Green PCR Kit (Qiagen) according to Table 24.

Table 24 MiRNA RT-qPCR reaction set up using miScript SYBR Green PCR Kit

Primer Master Mix		Sample Master Mix	
Components	Final concentration	Components	Amount [µl]
2x QuantiTect SYBR Green PCR Master Mix	1x (6.26 µl)	cDNA	3
Forward primer	0.7 µM	Primer Master Mix	33
miScript Universal Primer	0.7 µM		
H_2O	up to 11.5 µl		

The Sample Master Mix was transferred into 3 wells (technical triplicates), each 11 µl, of a MicroAmp Fast Optical 96-Well Reaction Plate, sealed and pulse centrifuged as above. The reaction was run in the ABI 7500 Fast Real-Time PCR System, using following cycle conditions: 15 min 95 °C, 40-55 cycles of 15 sec 94 °C, 30 sec 55 °C and 30 sec 70 °C. Subsequently, HRM analysis (60-95 °C) was performed.

RT-qPCR Data analysis

For relative quantification analysis of the expression of mRNA and miRNA, the Livak method was used (Livak & Schmittgen, 2001). The fluorescence threshold for determination of the threshold cycle (Ct) was set manually for each gene (Table 25).

Table 25 Fluorescence threshold values for Ct determination

mRNA	Fluorescence threshold	MiRNA	Fluorescence threshold	LMD mRNA	Fluorescence
GAPDH	0.235651	mir-215	0.034205		**threshold**
CYP7A1	0.027038	mir-194b-5p	0.057749	CYP7A1	0.078585
SFRP-5	0.488557	mir-27a-3p	0.023887	GAPDH	0.297384
SATB1	0.542667	mir-146a-5p	0.050586		
		let-7a-5p	0.045831		

1) $dCt = \mu Ct_{Target\ gene} - \mu Ct_{Reference\ gene}$
2) $ddCt = dCt_{test\ sample} - \mu dCt_{calibrator\ sample}$
3) $Foldchange = 2^{(-ddCt)}$

The mean (µ) Ct value of the gene of interest was normalised to the Ct value of the reference gene (1). All samples of the calibrator group were taken together to generate a mean dCt calibrator which was subtracted from the dCt of each test sample, forming the ddCt of each sample (2). DdCt represents the change in expression of the target gene between the test and calibrator group, normalised for any difference in loading between the calibrator and test samples. To obtain the fold change in expression between the test sample and the calibrator the following equation 3) was used. The Foldchange

determines at what fold the expression of the target gene in the test sample differs to the expression of the target gene in the calibrator sample.

2.2.26 Enzymatic PCR purification

For sequencing with SmartSeq from MWG Eurofins and for Sanger sequencing, PCR reactions were enzymatically purified. 10 µl PCR reaction was mixed with 0.4 µl exonuclease I (20U/µl) and 1 µl shrimp alkaline phosphatase (1U) and incubated at 37 °C for 30 min and at 80 °C for 15 min.

2.2.27 Sequencing with SmartSeq from MWG Eurofins

15 µl plasmid DNA (50-100 ng/µl) or 2 µl purified PCR product (2.2.10, 2.2.26) were mixed with 2 µl 10 µM primer and water to 17 µl total volume and transferred to a SmartSeq tube for sequencing.

2.2.28 Sanger Sequencing

Termination reaction

The termination reaction was performed using the Big Dye Terminator v1.1 Cycle Sequencing Kit (Life Technologies) with 2 µl purified PCR product (amplified using GoTaq® DNA Polymerase (2.2.20) and enzymatically purified (2.2.26)), 1x BigDye sequencing buffer, 1x dilution buffer (5 mM $MgCl_2$, 0.2 mM Tris-HCl pH 8.0), 0.25 µM sequencing primer and 1x BigDye Terminator in a total volume of 10 µl. The thermal conditions were: 20 sec 95 °C and 25 cycles of 12 sec 95 °C, 8 sec 51 °C and 4 min 60 °C.

Sephadex gel filtration

The termination reaction mix was purified using a sephadex gel filtrations method. 25 mg sephadex (determined volumetrically) was transferred into a MultiScreen HV plate (Millipore) and soaked in 300 µl water per well for 2 h at RT. After incubation, the plate was centrifuged at 960 x g for 5 min. The sequencing reaction (10 µl) was diluted with 15 µl 0.1 mM EDTA. The total volume of 25 µl was applied to the sephadex plate for gel filtration at 960 x g for 5 min and collected in a fresh 96 well plate.

Capillary gel electrophoresis

The 96 well plate was transferred to the ABI Prism 3130xl Genetic Analyzer for capillary gel electrophoresis in a 36 cm capillary using the POP-7™ Polymer (Life Technologies). The procedure ran with standard settings and the data was collected using the software Run 3130xl Data Collection v.3.0.

Analysis of polymorphisms

Quality assessment and base calling was performed by the AB Sequencing Analysis Software (v5.2). The polymorphisms in the sequences were detected using the Phred/Phrap/Polyphred/Consed-Software (Ewing & Green, 1998; Ewing et al, 1998; Gordon et al, 1998; Nickerson et al, 1997).

2.2.29 Pyrosequencing

A sequence of interest was amplified from bcDNA (2.2.16) with one biotinylated primer and the PyroMark PCR Kit (Qiagen) (2.2.20). The samples were sequenced on the PyroMark Q48 Autoprep Instrument using the PyroMark Q48 Advanced CpG Reagents (Qiagen) according to manufacturer's information with small modifications. When all reagents except the enzyme and substrate were brought to RT, 4 µl PyroMark Q48 Magnetic Beads (Qiagen) and 12 µl PCR reaction were loaded. 2 µl sequencing primer (10 µM primers diluted in PyroMark Annealing Buffer) was added automatically.

2.2.30 Next Generation Sequencing using Illumina technology

mRNA Sequencing

400 ng total RNA was used for the library preparation with the TruSeq RNA Library Preparation Kit v2 (Illumina) according to "TruSeq RNA Sample Preparation v2 Guide" with one small modifications. Elution2-Frag-Prime program was reduced to 4 min at 94 °C.

Quality and Quantity control was performed according to the "TruSeq RNA Sample Preparation v2 Guide" using the Bioanalyzer (2.2.7.3) and Qubit (2.2.7.2) respectively. Obtained average fragment/library size multiplied with the average molecular weight of a single DNA base pair (660 g/mol) and concentration resulted in the molarity of each library.

$$\frac{Concentration\ [ng/\mu l]}{average\ library\ size\ [bp] x\ 660\ g/mol}\ x\ 10^6 = molarity\ [\text{nM}]$$

The libraries were set to a 5 nM with EB buffer and quantified with qPCR according to "Sequencing Library qPCR Quantification Guide" from Illumina (2.2.24). The molarities of the libraries were corrected according to qPCR results and set to 2nM with EB buffer. Twelve 2 nM libraries carrying 12 different adapters were pooled and processed according to the "HiSeq and GAIIx Systems Denature and Dilute Libraries Guide: Denature and Dilute Libraries for HiSeq Clustering Standard Normalization Method" (Illumina) with small modifications. 10 µl pooled 2 nM libraries was mixed with 10 µl 0.1 M NaOH and incubated for 5 min at RT. The resulting 1 nM library pool was diluted with prechilled HT1 buffer to 20 pM in a total volume of 1000 µl. The 20 pM solution was set to a chosen molarity with HT1

buffer reaching a total volume of 420 µl and supplemented with 5 µl 12.5 pM PhiX library (Table 26). The PhiX library was denatured and diluted according to "HiSeq and GAIIx Systems Denature and Dilute Libraries Guide: Denature and Dilute PhiX for HiSeq Clustering" (Illumina).

Table 26 Calculations for the molar adjustment of library pools for Illumina Sequencing of mRNA

Final concentration	12 pM	13 pM	14 pM	18 pM
20 pM denatured library pool	252	273	294	378
Prechilled HT1 buffer	168	147	126	42
12.5 pM PhiX	5	5	5	5
Total volume	425	425	425	425

The HiSeq was prepared and operated according to the "HiSeq® 2500 System Guide, Chapter 5 Sequencing in Rapid Run Mode" (Illumina). The HiSeq Rapid v2 flow cell was rinsed with laboratory grade water, dried with low lint Kimtech Science Precision wipes (Kimberly-Clark Professional) and together with the library pool loaded into the HiSeq2500. Clustering and sequencing was performed using the HiSeq Rapid PE Cluster Kit v2 (Illumina) and HiSeq Rapid SBS Kit v2 (Illumina) at 2 x 100 bp read configuration to generate Fastq files.

Small RNA Sequencing

1 µg total RNA was used for small RNA library preparation using the TruSeq Small RNA Library Prep Kit (Illumina) according to manufacturer's information. Quality and quantity of libraries was analysed using the Bioanalyzer Agilent High Sensitivity DNA Kit (2.2.7.3) and Qubit™ dsDNA HS Assay Kit (2.2.7.2).

400 ng of 6 samples each, all carrying different adapters were pooled and run on a gel. After purification quality and molarity of the library pools were analysed using the Bioanalyzer Agilent High Sensitivity DNA Kit (2.2.7.3). Library pools were set to 4 nM with 10 mM Tris-HCl (pH 8.5) and processed according to the "MiSeq System Denature and Dilute Libraries Guide: Standard Normalization Method". 5 µl 4 nM library pool was mixed with 5 µl 0.2 N NaOH and incubated for 5 min at RT. 990 µl prechilled HT1 buffer was added resulting in 1 ml 20 pM library pool. 360 µl 20 pM library pool was mixed with 240µl prechilled HT1 buffer for a 12 pM dilution. 594 µl 12 pM library pool was supplemented with 6 µl 12.5 pM PhiX library that was equally denatured. The flow cell was rinsed with laboratory-grade water and dried with low lint Kimtech Science Precision wipes (Kimberly-Clark Professional). The flow cell and library pool were loaded into the MiSeq and sequencing was performed using the MiSeq Reagent Kits v2 (Illumina) at 1×50 bp read configuration to generate Fastq files. The MiSeq was prepared and operated according to the "MiSeq® System Guide" (Illumina).

2.2.31 Southern blot analysis

2.2.31.1 Preparing Digoxygenin labelled probes

Digoxigenin (DIG) labelled DNA probes of 600-750 bp and 50-60 % GC content were generated via GoTaq DNA Polymerase PCR (2.2.20) with supplementation of 60 mM Digoxigenin-11-dUTP (Roche Diagnostic GmbH). Successful labelling (slower gel migration than control PCR without digoxigenin addition) was assessed by gel electrophoresis and extracted from the gel (2.2.10).

2.2.31.2 Dot blot

Dot blot was performed to determine binding capacity and optimal hybridisation temperature of a probe. DNA was applied directly to the positively charged Nylon membrane. The membrane was baked for 30 min at 120 °C and prepared according to Southern blot protocol (2.2.31.3).

2.2.31.3 Southern blot

10 µg gDNA were digested for 4 h at 37 °C with 4 Units Enzyme per µg DNA. The DNA was separated on a 1x TAE 0.8 % agarose gel free of peqGREEN with 2-Log DNA Ladder (New England Biolabs) and the DIG-labeled DNA Molecular Weight Marker VII (Roche Diagnostic GmbH). The Gel ran at 120 V for 10 min and 30 V overnight. The next day, half the 2-Log DNA Ladder lane was cut off and stained in a peqGREEN bath for 10-30 min. The stained ladder was photographed with a ruler to visualise the run and cut the gel correspondingly. All following incubations or washings were performed at RT shaking. For hybridisation of targets larger than 5 kb, the gel was incubated in depurination solution (250 mM HCl) for 10 min maximum. The gel was rinsed in demineralised water and incubated in denaturation solution (0.5 M NaOH, 1.5 M NaCl) twice for 15 min. The gel was rinsed and incubated twice in neutralisation solution (0.5 M Tris-HCl, pH 7.5; 1.5 M NaCl) for 15 min. The neutralised gel was equilibrated in 20x SSC (3 M NaCl, 0.3M sodium citrate) for at least 10 min.

The capillary transfer blot was assembled from bottom to top as follows: pan filled with two litres 20x SSC, a bridge of blotting paper, glad wrap with a window the size of gel and membrane, one blotting paper soaked in 20x SSC, the gel facing down, the dry Nylon membrane (positively charged, Amersham Hybond-N+, GE Healthcare Europe GmbH), a dry blotting paper and a big stack of paper towels. The construct was set under pressure over night. The next day, blot was disassembled to wash the membrane in 2x SSC (0.3 M NaCl, 30 mM sodium citrate) and bake it for 30 min at 120 °C. The dry membrane was placed into a hybridisation bottle with 3 ml DIG Easy Hyb Buffer (DIG Easy Hyb Granules, Roche Diagnostic GmbH) for blocking for 30 min to 3 h at the hybridisation temperature of the probe (T_{Hyb}), not exceeding 43°C.

$$T_{Hyb} = T_m - (20°C - 25°C)$$

$$T_m = 49.82 + 0.41 * GC_{probe} - {}^{600}/_{length_{probe}}$$

During incubation, the DIG labelled hybridisation probe was diluted in 50 µl water, denatured for 5 min at 95 °C and chilled on ice. When the probe had been used before, it was denatured for 5 min at 68 °C. DIG Easy Hyb Buffer was removed from the membrane and replaced with fresh DIG Easy Hyb Buffer containing 40 ng/ml probe. The membrane was hybridised in motion over night at T_{Hyb}. The next day, the probe in DIG Easy Hyb buffer was stored at -20 °C for reuse. The membrane was shaken in low stringency buffer (2x SSC, 0.1 % SDS) for 15 and 10 min. High stringency buffer was heated (Table 27).

Table 27 Incubation conditions with high stringency buffer, dependent on target homology and GC content of the probe

Target homology	GC content	Buffer	temperature
80-100 %	Average (40 %)	0.5 % SSC, 0.1 % SDS	65 °C for probes > 100bp, <65 °C for probes ≤ 100 bp
<80 %	Average (40 %)	0.5 % SSC, 0.1 % SDS	approx. 60 °C
80-100 %	High (≥50 %)	0.1 % SSC, 0.1 % SDS	68 °C

The membrane was shaken in high stringency buffer (0.1 % SSC, 0.1 % SDS) for 15 min twice. The membrane was washed for 2 min with wash buffer (0.3 % (v/v) Tween 20 in maleic acid buffer (0.1 M maleic acid, 0.15 M NaCl, pH 7.5)) and incubated at RT for 1-1 ½ h in blocking solution (1 % (w/v) Blocking Reagent 11096176001 (Roche Diagnostic GmbH) in maleic acid buffer). Next, the membrane was shaken in Anti-Digoxigenin-AP sheep antibody solution (1:10000 in blocking solution) for 30 min. Two washing steps of 15 min each with washing buffer followed. The membrane was equilibrated for 3 min in detection buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5 (20 °C)) while the substrate for the phosphate conjugated antibody CDP-star solution (Roche Diagnostic GmbH) was diluted 1:100 in detection buffer. The membrane was placed on a plastic envelope and the CDP-star dilution was applied. The envelope was sealed and incubated at 37 °C for 5 min in the dark. In a dark room, the membrane was placed on top of an x-ray film in an x-ray clip cassette. The closed cassette was incubated at 37 °C for 1-2 h. After sufficient exposure the film was developed in the dark room rinsing it in developer (until the length marker was visible), water, fixer solution and water.

2.3 Microbiological methods

2.3.1 Bacterial culture

A mini culture of 3 ml, a midi culture of 100 ml or a maxi culture of 300 ml LB medium (2.5 % (w/v) Difco LB Base, Miller) with ampicillin (Carl Roth GmbH) (100 µg/ml) or spectinomycin (Fluka Laborchemikalien GmbH) (50 µg/ml) was inoculated with a glycerol stock, a colony from an agar plate or a fluid bacterial culture. The culture was shaken over night at 37 °C.

2.3.2 Cryoconservation of bacterial cultures

200 µl 100 % glycerol were mixed with 200 µl bacterial culture and stored at -80 °C.

2.3.3 Transformation of bacteria

50 µl competent DH10b *E. coli* cells (Invitrogen) were thawed on ice and mixed with 1-3 µl ligation (3 µl purified ligation or 1:10000 diluted plasmid). The mix was transferred to a cold electroporation cuvette (2 mm gap, Peqlab Biotechnologie GmbH), put into the Multiporator® (Eppendorf) and shocked at 2500 V for 5 msec. The shocked cells were submerged in 500 µl LB medium and incubated for 30-45 min at 37 °C shaking. The recovered cells were plated onto agar (4 % (w/v) Difco LB Agar, Miller) plates with 100 µg/ml ampicillin or 50 µg/ml spectinomycin and incubated over night at 37 °C.

2.3.4 Blue white screening of bacterial colonies

For cloning PCR fragments into the pGEM®-T Easy Vector System (Promega Corporation), ampicillin agar plates were coated 30 min prior to bacteria culture with 20 µl 100 mg/ml X-gal solution (solved in N,N-dimethylformamid) and 4 µl 1 M IPTG. After overnight culture at 37°C white bacterial colonies carried pGEM®-T Easy Vector where the lacZ gene was disrupted by incorporated PCR fragment.

2.4 Tissue culture methods

Mammalian cells were cultured in a humified Steri-Cycle CO2 incubator at 37 °C with 5 % CO_2. and handled only with sterile equipment in a sterile laminal flow cabinet. The medium was exchanged every 2-3 days. Unsterile solutions were sterile filtered with 0.22 µm filters.

2.4.1 Passaging cells

Cells of 80-100 % confluence, were washed with D-PBS (Sigma-Aldrich Chemie GmbH) and incubated with accutase (Sigma-Aldrich Chemie GmbH) for 4-10 min in the incubator. When cells were fully detached, the reaction was stopped by adding ≥ volume of medium. The cell suspension was distributed to a new vessel, counted (2.4.2), frozen (2.4.4) or pelleted for DNA isolation (2.2.4).

2.4.2 Counting cells

10 µl cell suspension was mixed with 10 µl trypan blue 0.4 % (Invitrogen GmbH). 10 µl were transferred to a cell counting chamber slide (Invitrogen GmbH) and inserted into the Countess (Invitrogen GmbH). The Countess calculated the concentration of the total, life and dead cells and the viability.

2.4.3 Isolation and culture of primary porcine kidney fibroblasts

Kidneys of euthanized animals were cleared of their fine skin and rinsed in 80 % ethanol for six times before transfer to a sterile laminal flow cabinet. A piece of ca. 1 x 1 cm was cut from the inside of the kidney, avoiding fat, vascular tube or outside tissue, with a scalpel (B. Braun Melsungen AG). The piece was washed in three different tubes of each 80 % ethanol and D-PBS containing 100 U/ml penicillin and 0.1 mg/ml streptomycin (Sigma-Aldrich Chemie GmbH) and 2.5 mg/ml amphotericin B (Sigma-Aldrich Chemie GmbH) (D-PBS-P/S/A). The cleaned tissue was minced in a petri dish with 2 ml D-PBS-P/S/A and incubated in an erlenmeyer flask with 1mg/ml Collagenase Type I-A (Sigma-Aldrich Chemie GmbH) in D-PBS-P/S/A for 20-30 min at 37 °C stirring. If many tissue pieces remained intact, the suspension was filtered through a mesh and transferred to a 50 ml falcon. 13 ml medium containing 100 U/ml penicillin and 0.1 mg/ml streptomycin and 2.5 mg/ml amphotericin B (medium-P/S/A) were added to stop collagenase reaction and centrifuged at 300 x g for 5 min. The supernatant was aspirated and the pellet resuspended in 14 ml medium-P/S/A for centrifugation. This was repeated once more. The pellet was resuspended in medium-P/S/A and depending on its size distributed to 2-8 T150 flasks.

The medium was exchanged daily for 3 days. On the fourth day, antibiotic and antimycotic-free medium was applied and conditioned for 3 days for mycoplasma testing 2.2.22.

2.4.4 Cryoconservation of mammalian cells

Cells were detached (2.4.1), pelleted at 300 x g for 5 min and resuspended in freezing medium (Table 10). The suspension was transferred to cryo vials and put into Nalgene Mr. Frosty Freezing containers (Thermo Fisher Scientific), filled with isopropanol.

Thawing was conducted as quickly as possible. Cryo vials were thawed in the 37 °C water bath and transferred directly to fresh medium, diluting the cytotoxic DMSO. The cells were pelleted at 300 x g for 5 min and cleared of the supernatant. The pellet was resuspended in DMSO-free medium and transferred to a fresh culturing vessel.

2.4.5 Transfection of mammalian cells

2.4.5.1 Electroporation

Cells were detached and counted. 1 x 10^6 cells were pelleted at 300 x g for 5 min. The supernatant was aspirated and the cells were resuspended in 800 µl hypoosmolar electroporation buffer (Eppendorf). 10 µg sterile DNA (in sterile water or low TE buffer) was added, mixed and incubated at RT for 5 min. The suspension was transferred to an electroporation cuvette (4 mm gap, Peqlab Biotechnologie GmbH) avoiding bubbles and shocked with 1200 V for 85 µsec. The cells were incubated at RT for 5 min and transferred to two T25 flasks with fresh medium. The next day, dead cells were washed away with D-PBS and fresh medium was added.

2.4.5.2 Nucleofection

Cells were detached and counted. Different cell types required different nucleofector kits according to manufacturer's information and small modifications (Table 28). The required cell amount was pelleted at 300 x g for 5 min. The pellet was resuspended in 100 µl nucleofector solution and mixed with 2-5 µg sterile DNA (in sterile water or low TE buffer). The suspension was transferred to a nucleofection cuvette and treated with the corresponding program. The nucleofected cells were submerged with fresh medium and transferred to a T25 flask. The next day, dead cells were washed away with D-PBS and fresh medium was added.

Table 28 Nucleofection conditions for different cell types

Cell type	Kit	Cell number	Nucleofector solution	DNA	Program
pADMSCs	Human MSC Nucleofector® Kit	5 x 10^5	100 µl	2-5 µg	C-17, U-23
pKFs	Basic Nucleofector Solution Primary Fibroblasts	1 x 10^6	100 µl	2-3 µg 4 µg vector + 400 ng CRISPR	N-24

2.4.5.3 Lipofection

One day prior lipofection, cells were plated on 10 cm culture dishes to reach 30-50% confluence at the point of transfection. The cells were washed with D-PBS twice and covered with 4 ml OptiMEM (Life

Technologies). For each 10 cm dish 6 µl Lipofectamine™ 2000 Transfection Reagent (Invitrogen GmbH) was mixed with 294 µl OptiMEM in one reaction tube and 4-10 µg DNA was mixed with OptiMEM (total volume 300 µl) in a second tube. Each tube was vortexed and incubated at RT for 5 min. The Lipofectamine 2000 mix was carefully dropped into the DNA mix, vortexed and incubated for 25-30 min at RT. The Lipofectamine-DNA mix was dropped directly onto the cells. The cells were put into the incubator and after 4 h supplemented with 6 ml medium.

2.4.6 Killing curve experiment

To determine optimal selection concentrations, 1 and 0.5 x 10^4 cells were plated onto a 12 well and cultured with different concentrations of the antibiotics (Table 29). The optimal concentration, was the one, where cells were dead after 7 days of cultivation.

Table 29 Concentration ranges for antibiotics when performing a killing curve experiment

Antibiotic	Range
G-418	0-1200 µg/ml
Blasticidin S	0-10 µg/ml
Puromycin	0-1.5 µg/ml

2.4.7 Selection

24-48 h (depending on confluence) after lipofection cells were set under selection. 48 h after nucleofection and electroporation, cells were detached, counted, less than 1 x 10^5 expanded onto a 15 cm culture dish and set under selection with G-418 (Genaxxon bioscience GmbH) or Blasticidin S (InvivoGen) for 10-14 days until single-cell clones had reached a size of about 100 cells. The antibiotic containing media were exchanged every 2-3 days. Puromycin (InvivoGen) selection was performed 24 h after transfection for 48 h. The puromycin medium was exchanged every day. 24 h after the end of puromycin selection, the cells were expanded and selected on G-418 as described above.

Table 30 optimal antibiotic concentrations for selection of different cell isolations

Cell isolate	G-418	Blasticidin-S	Puromycin
pADMSCs 110111	600 µg/ml	8 µg/ml	1 µg/ml
pKFs 73	1000 µg/ml	-	1.5 µg/ml
pKF 270	800 µg/ml	-	0.5 µg/ml

2.4.8 Clone picking

Single-cell clones of ~100 cells were marked on the outer wall of the culturing vessel. Cells were washed with D-PBS. After D-PBS aspiration, small sterilised cloning rings were dipped into sterile silicon

fat and tightly placed around the marked single-cell clones. 50 µl accutase was applied into each ring and incubated (2.4.1). When cells were fully detached 100 µl medium was added into each ring. The cell suspension in each ring was transferred into a 24 well with 1 ml fresh medium. If cells had been cultured in selection medium before, selection was reduced (eg. from 1000 -800 µg/ml) but not taken off.

2.4.9 Clone expansion and screening

Confluent single-cell clones were detached (2.4.1) with 200 µl accutase and supplemented with 400 µl medium after incubation. 400 µl cell suspension was transferred to a 12 well and 200 µl were transferred into a PCR reaction tube and pelleted for gDNA isolation (2.2.4) and screening PCR (2.2.20).

2.4.10 Cell preparation for somatic cell nuclear transfer

Correctly targeted clones (nuclear donors) were pooled or separately plated on a 12 well to reach 70-80 % confluence. 48 h before somatic cell nuclear transfer (SCNT) cells were washed 2x with D-PBS and synchronised in G0/G1 phase by culture in starvation medium (0.5 % FCS). SCNT and embryo transfer was done by the Chair for Molecular Animal Breeding and Biotechnology (LMU, Munich, Germany).

2.5 Biochemical methods

2.5.1 Protein extraction from cultured cells

Cells were washed 2x with ice cold D-PBS. Per 15 cm dish cells were mechanically detached on ice with cells scrapers and 500 µl CytoBuster™ Protein Extraction Reagent (Merck KGaA) or RIPA buffer containing 1x cOmplete Protease Inhibitor Cocktail and PhosSTOP Phosphatase Inhibitor Cocktail (Roche Diagnostic GmbH). After 2-5 min mechanical cell dissociation, the cells suspension was frozen at -80 °C for at least 30 min. The suspension was thawed on ice and centrifuged at 4 °C at 15493 x g for 30 min to pellet dead cells and cell debris. The supernatant containing proteins was converted into a new reaction tube and stored at -80 °C.

2.5.2 Determination of protein concentration

Protein concentration was determined using the Advanced Protein Assay Reagent (Cytoskeleton Inc.). 5 µl protein sample was mixed with 995 µl 1x Advanced protein assay Reagent (1:200 dilution) and distributed to three 96 wells, 300 µl each. Blanks of Advanced Protein Assay Reagent alone and a 1:200

dilution of the protein extraction buffer were applied to three wells each. The plate was inserted into the ELISA -Photometer. Together with the ELISA-Reader software (Ascent Software) the absorbance at 595 nm was measured. Protein concentration was calculated using the following formula.

1.0 $OD_{570\ to\ 615\ nm}$ = 37.5 µg protein per ml reagent per 0.8 cm

2.5.3 Western blot Analysis

2.5.3.1 Sodium Dodecyl Sulfate Polyacrylamide Gel electrophoresis (SDS-PAGE)

The SDS polyacrylamide gel was prepared with different percentages depending on the molecular weight of the protein of interest (Table 31) using the Mini-PROTEAN 3 Cell system (BioRad). The separation gel was prepared first. TEMED and APS were applied last. Right after, the gel was mixed and 3-3.5 ml (leaving room for the collection gel) applied into the gel pouring chambers. To remove bubbles, a mix of water and isopropanol (1:1) was added on top. After full polymerisation for 30 min the water-isopropanol mix was removed. The collection gel was prepared (Table 31) and applied on top. The Mini-PROTEAN® Comb (BioRad) was inserted and left to dry for 30 min. The fully polymerised gel was stored in a moist plastic bag at 4 °C for later use, or loaded into the running chamber. The running chamber was filled with running buffer (25 mM Trizma Base, 0.2 M Glycin, 0.1 % SDS, pH 8.3, 10 mM ß-mercaptoethanol). The protein samples were thawed on ice and 20-40 µg protein was mixed with 4x Laemmli buffer (including DTT) (250 M Tris-HCl, pH 6.8, 4 % (w/v) SDS, 0,1 M saccharose, traces of bromphenol blue, 26mM DTT (freshly added)) to reach a 1x buffer concentration. The same was done with the Precision Plus Protein™ All Blue Standard (BioRad) before both were denatured at 95 °C for 5 min. They were kept on ice until loaded into the pockets of the SDS-polyacrylamide gel. The electrophoresis ran for 20-30 min at 60 V, 100 V for 40 min and 140 V for 75-90 min.

Table 31 Preparation and composition of 0.75 mm SDS-polyacrylamide gels

0.75 mm gels	**Separation gel**		**Collection gel**	
	<100 kDa	>100 kDa	<100 kDa	>100 kDa
Reagents	12%	10%	5%	4%
0.5 M Tris-HCl, pH 6.8	-	-	666 µl	666 µl
1 M Tris-HCl, pH 8.8	1.5 ml	1.5 ml	-	-
Water	1.22 ml	1.42 ml	1.61 ml	1.68 ml
SDS (10 %)	40 µl	40 µl	26.6 µl	26.6 µl
Polyacrylamide (40 %)	1.2 ml	1 ml	333 µl	266.6 µl
Temed	1.6 µl	1.6 µl	2.6 µl	2.6 µl
APS (10 %)	40 µl	40 µl	26.6 µl	26.6 µl
Total	4 ml	4 ml	2.6	2.6

2.5.3.2 Western blot

The proteins separated by SDS-PAGE were transferred onto the Roti-PVDF membrane (0.45 µm) by semidry transfer blot using the Trans-blot SD Semi-Dry Transfer cell. Two pieces blotting paper and the SDS-polyacrylamide gel were equilibrated in semi dry transfer buffer (25 mM Trizma Base, 0.2 M glycin, 20 % (v/v) methanol, 0.1 % (w/v) SDS). The Roti-PVDF membrane (0.45 µm) (Brand GmbH & Co. KG) was activated in methanol for 1 min and soaked in semi dry transfer buffer. The blot was assembled on the Trans-Blot SD Semi-Dry Transfer cell as follows from bottom to top: semi dry transfer buffer, blotting paper, the activated membrane, the gel and another blotting paper. Air bubbles were rolled out using a serological pipette. Semi dry transfer buffer was poured on top of the blot, and the Trans-Blot SD Semi-Dry Transfer cell was closed. The blot was run at 15 V for 70 min to blot proteins of up to 160 kDa from a 10 % 1.5 mm SDS-polyacrylamide gel. When two SDS-polyacrylamide gels were run at the same time, a voltage of 25 V and a time of 2.5 h was not exceeded. After blotting, the coloured bands of the Precision Plus Protein™ All Blue Standard ladder were retraced on the membrane.

2.5.3.3 Ponceau S staining

The membrane was soaked for 2-3 min in Ponceau S solution (0.5 % (w/v), 1 % glacial acetic acid), shaking at RT. The membrane was washed with demineralised water, until the background turned white and only protein bands were stained red. A photo was taken and the dye was fully washed off.

2.5.3.4 Conventional Antibody application

The membrane was blocked in blocking solution (5 % (w/v) milk powder (Carl Roth GmbH) in 1xTBST (20 mM Trizma Base, 140 mM NaCl, 0.1 % (v/v) Tween 20) for at least 1 h shaking at RT. The membrane was washed three times with 1x TBST and incubated with primary antibody (diluted in blocking solution Table 17) over night at 4 °C shaking. The next day the primary antibody dilution was supplemented with a trace of sodium azide and stored at 4 °C for reuse. The membrane was rinsed three times and incubated 3 times with 1x TBST for 15 min shaking. The secondary antibody (horse-radish peroxidase (HRP) conjugated), diluted in blocking solution (Table 18), was added to the membrane and incubated for 1 h at RT. The membrane was again washed 3 times by rinsing and incubating for 15 min in 1x TBST.

For visualisation of the antibody bound proteins of interest, the membrane was covered with Pierce™ ECL Western Blotting Substrate (Thermo Fisher Scientific), sealed in a plastic foil and incubated in the dark for 1 min. In the dark room, the membrane was placed onto an x-ray film in an x-ray clip cassette. The film was exposed to the membrane for 1-20 min and developed in the dark room by rinsing it in developer, water, fixer solution and water.

2.5.3.5 iBind antibody application

The membrane was blocked for at least 1 h in blocking solution and subsequently rinsed in 1x TBST three times. The iBind™ Solution Kit (Invitrogen GmbH) was prepared according to manufacturer's information for HRP detection. The membrane was submerged in 1x iBind solution and the antibodies were diluted with 1x iBind solution according to Table 17 and Table 18. When two different primary antibodies were applied simultaneously the amount of secondary antibody was doubled (Table 18). The membrane was incubated for 2.5-6 h and washed 3 times by rinsing and incubating for 15 min in 1x TBST. Visualisation was performed as above (2.5.3.4).

2.5.4 Colonoscopy of pigs

All experiments on animals were approved by the Government of Upper Bavaria (permit number 55.2-1-54-2532-6-13) and performed according to the German Animal Welfare Act and European Union Normative for Care and Use of Experimental Animals.

Colonoscopies were performed by Professor Dr. Dieter Saur (Klinikum Rechts der Isar II, Technische Universität München, Munich, Germany) ca. every 6 months, starting at 3 months. A STORZ colonoscopy system allowing macroscopic images of the colorectum and colonic lesions was used. Normal mucosa (at 40 cm depth) and polyp biopsies were collected and snap frozen for RNA, DNA and protein isolation or embedded into Tissue-Tek Cryomolds (Sakura Finetek Europe B.V.) with O.C.T.™ Compound (Sakura Finetek Europe B.V.) on dry ice for cryosectioning and stored at -80°C.

2.5.5 Cryosectioning

Cryosectioning was performed using the Microm HM 560 Cryostat. During sectioning all samples were kept at -20 °C or on dry ice. 3 – 4 sections of 4 µm thickness were mounted on a MembraneSlide 1.0 PEN (D) (Carl Zeiss Jena GmbH) and stored at -80 °C until staining and laser microdissection was performed (never exceeding more than 2 days). In general, cryosections were cut, stained, the areas of interest laser microdissected and the nucleic acids isolated, within one day.

2.5.6 Haematoxylin-Eosin staining of cryosections

To guarantee an RNase free environment, all staining solutions were made using autoclaved 0.2 µm filtered demineralised water and stored at RT.

Before staining the slides were thawed at RT for 2 min. Staining was performed in 50 ml falcons, with fresh solution aliquots every 8 slides, as follows: 1 min 70 % ethanol, 30 sec water, 1 min Hematoxylin (Mayer's Hemalaun solution, Applichem), brief dip into two water aliquots, 1 min Eosin (Eosin solution, 2C-140, Waldeck GmbH & Co. KG), brief dip into two 96 % ethanol aliquots, 30 sec 100 % ethanol followed by drying for 2-5 min.

2.5.7 Microscopy

Morphology, viability, density and fluorescence microscopy of cultured cells was visually assessed using the microscopes Axiovert 40CFL and Axiovert 200M (Carl Zeiss Microscopy GmbH). Red fluorescent proteins were excited at 554 nm with emission at 581 nm. Green fluorescent proteins were excited at 484 nm with emission at 510 nm. Photographs were acquired using the AxioCam HRm and AxioCam MRc cameras (Carl Zeiss Microscopy GmbH) and the Axiovision and Axiovision Rel. 4.8 (Carl Zeiss Microscopy GmbH) software respectively.

2.5.8 Laser microdissection

Laser microdissection was performed immediately after the haematoxylin-eosin staining using the UV-laser cutting system Laser Microdissection Systems 6000 (Leica Microsystems). Cryosection were scanned and areas of interest were marked on the Leica Application Suite software (Leica Microsystems). In total 15 crypts or 120 000 µm² of stromal tissue was marked and cut by the laser (DM6000 B9, Leica Microsystems). The excised areas were collected in a reaction tube cap containing 40 µl RLT Plus buffer (including 1:100 β-mercaptoethanol) supplied by the AllPrep® DNA/RNA Micro Kit (Qiagen). After complete dissection, additional 60 µl RLT Plus buffer was added. To prevent degradation of RNA the LMD procedure did not exceed 2 h per slide. Samples were stored at - 80°C until DNA and RNA isolation.

2.6 Data analysis

2.6.1 Statistical Analysis

Heat maps, statistical analysis and graph generation of sequencing data, RT-qPCR data, polymorphism analysis and the CpG methylation analysis was generated using the open source tool R (https://www.R-project.org/).

2.6.2 *In silico* miRNA target prediction using Diana tools

To identify potential pathways influenced by differential miRNA expression, *in silico* target prediction was performed using mirPath v.3 (Vlachos et al, 2015b) in combination with TarBase v7.0 (Vlachos et al, 2015a) and TargetScan (Lewis et al, 2005).

2.6.3 Gene set enrichment analysis

Gene set enrichment analysis was performed using the GSEA software (version 2.2.4) (Mootha et al, 2003; Subramanian et al, 2005).The log 2 fold change, adjusted p-Value and the Human Genome Organisation (HUGO) gene symbols were used to generate a preranked file as input for the GSEAPreranked tool. The enrichment analysis was performed under the following specifications: classic enrichment statistics, 1000 permutations and hallmark gene sets from Molecular Signatures Database (MSigDB) (version 6.1).

3. Results

3.1 Attemp to identify modifier genes in the porcine model for colorectal cancer

The generation of a porcine model for colorectal cancer (CRC) carrying a translational stop signal at codon 1311 of the endogenous APC gene, orthologous to human mutation 1309, was published in 2012 (Flisikowska et al, 2012). Four generations of *APC*^{1311} pigs were regularly analysed by colonoscopy. The analyses showed, just like in humans, a wide variation in the severity of polyposis (Crabtree et al, 2002), ranging from ≥100 (high polyp animals (HP)) to only 1-10 polyps (low polyp animals (LP)) in the distal colorectum (last 40 cm). Also like in humans, a correlation between severe polyposis and the amount of high grade neoplasia was observed, indicating that severe polyposis might be a susceptibility factor towards CRC also in pigs (Debinski et al, 1996; Shussman & Wexner, 2014). Variability in polyposis severity in humans is believed to be mediated by genetic loci (Crabtree et al, 2002; Houlston et al, 2001). In humans some of these so called modifier loci have been identified using microarrays for genome wide association studies comparing healthy individuals and CRC patients (Broderick et al, 2007; Tomlinson et al, 2007; Tomlinson et al, 2008; Whiffin et al, 2014). A study focussed on FAP patients showed, that two of the single-nucleotide polymorphisms (SNPs) associated with sporadic CRC risk (rs16892766 at 8q23.3 and rs3802842 at 11q23.1) were also associated with severe polyposis in familial adenomatous polyposis (FAP) (Ghorbanoghli et al, 2016). Thus the identification of modifier genes or loci mediating severe polyposis in FAP may also mediate severe polyposis and thus higher risk of CRC in sporadic CRC with APC mutations. The *APC*^{1311} pigs, offer the possibility to analyse modifiers directly in the normal mucosa of the distal colorectum (last 40 cm) between HP and LP animals. Changes promoting CRC are not only due to genomic but also epigenetic alterations (1.2.1, 1.2.3, 1.2.4) and dysregulated microRNA (miRNA) (1.2.2, 1.2.3, 1.2.4) that influence the amount of functional proteins on pre and posttranscriptional level. Gene expression was compared between HP and LP to to investigate this in the porcine model. Messenger RNA (mRNA) sequencing analysis does not only allow detection of single-nucleotide polymorphisms (SNPs) and mutations in protein coding regions, but also epigenetic differences and miRNA differences that influence gene expression. Furthermore, protein coding regions are generally conserved between species and therefore better annotated in the porcine genome than non-coding regions. The power of epigenetic mechanisms and miRNAs on CRC development is similar if not equal to gene mutations (1.2.1, 1.2.2, 1.2.3, 1.2.4). Additional sequencing of miRNAs from normal mucosa and comparative miRNA expression analysis between HP and LP animals allowed identification of miRNA modifiers that influence cellular processes.

3.1.1 Attempt to identify modifier genes on mRNA level

MRNA of 35 normal mucosa samples at 40 cm colorectum depth of animals aged 3-9 months was isolated. The RNA was DNase treated, and the quality and quantity determined. 400 ng RNA was used for library preparation. samples marked with * were enriched for 12 and all other for 15 cycles (Table 32). The quality and quantity of the resulting libraries were determined to calculate the molarity of each library. After QPCR quantification and molarity corrections (2.2.24), the libraries were pooled (12 libraries/ flow cell), clustered and sequenced.

Table 32 Animals sequenced for analysis of modifier genes on mRNA level
A, innuPREP RNA Mini Kit (Analytik Jena); Z, Direct-zol™ RNA Miniprep Kit (Zymo Research).

Animal ID	Born	Collection	Age	Sex	Pheno-type	Kit	RIN	Sequenced
128*	07.03.2013	16.07.2013	0y 4m 9d	female	LP	Z	9.1	15.10.2014
145	14.03.2013	04.12.2013	0y 8m 20d	male	LP	Z	6.5	04.02.2015
145*	14.03.2013	25.06.2013	0y 3m 11d	male	LP	Z	8.1	15.10.2014
150	14.03.2013	16.07.2013	0y 4m 2d	female	LP	A	8.4	16.03.2016
152*	14.03.2013	22.10.2013	0y 7m 8d	female	LP	Z	8.9	15.10.2014
153*	14.03.2013	23.10.2013	0y 7m 9d	female	LP	Z	7.4	15.10.2014
155*	14.03.2013	26.06.2013	0y 3m 12d	male	LP	Z	9.2	15.10.2014
157*	14.03.2013	17.07.2013	0y 4m 3d	female	LP	Z	8.8	15.10.2014
163*	15.03.2013	19.09.2013	0y 6m 4d	female	HP	Z	9.1	15.10.2014
163	15.03.2013	03.12.2013	0y 8m 18d	female	HP	Z	7.5	04.02.2015
168*	12.05.2013	18.09.2013	0y 4m 6d	male	HP	Z	8.7	15.10.2014
173*	12.05.2013	23.10.2013	0y 5m 11d	female	LP	A	8.9	15.10.2014
173	12.05.2013	02.12.2013	0y 6m 20d	female	LP	Z	7.3	04.02.2015
174*	12.05.2013	23.10.2013	0y 5m 11d	female	LP	A	8.8	15.10.2014
251	21.02.2014	24.06.2014	0y 4m 3d	castrated male	HP	Z	6.6	10.03.2016
252	21.02.2014	25.06.2014	0y 4m 4d	castrated male	HP	Z	7.5	10.03.2016
253	21.02.2014	24.06.2014	0y 4m 3d	castrated male	HP	A	7.9	10.03.2016
300	24.08.2014	16.03.2015	0y 6m 20d	female	LP	A	7.8	10.03.2016
322	18.10.2014	16.03.2015	0y 4m 26d	male	HP	A	8.1	16.03.2016
324	18.10.2014	16.03.2015	0y 4m 26d	male	HP	A	8.3	16.03.2016
326	18.10.2014	17.03.2015	0y 4m 27d	female	HP	Z	6.6	10.03.2016
328	18.10.2014	17.03.2015	0y 4m 27d	female	LP	Z	7	10.03.2016
339	20.10.2014	17.03.2015	0y 4m 25d	female	HP	A	8.6	16.03.2016
471	17.08.2015	04.12.2015	0y 3m 17d	female	LP	A	9.1	12.05.2017
474	17.08.2015	09.12.2015	0y 3m 22d	female	HP	A	8.6	12.05.2017
524	18.01.2016	18.04.2016	0y 3m 0d	castrated male	LP	A	9.3	12.05.2017

525	18.01.2016	18.04.2016	0y 3m 0d	castrated male	HP	A	9	12.05.2017
527	18.01.2016	18.04.2016	0y 3m 0d	male	HP	A	10	12.05.2017
534	18.01.2016	19.04.2016	0y 3m 1d	female	LP	A	9.1	12.05.2017
586	11.03.2016	13.09.2016	0y 6m 2d	female	LP	A	9	12.05.2017
588	11.03.2016	12.09.2016	0y 6m 1d	female	HP	A	10	12.05.2017
591	15.03.2016	19.09.2016	0y 6m 4d	castrated male	LP	A	7.8	12.05.2017
598	15.03.2016	12.09.2016	0y 5m 28d	male	HP	A	8.6	12.05.2017
722	26.09.2016	17.01.2017	0y 3m 22d	male	HP	A	9.1	12.05.2017
727	26.09.2016	16.01.2017	0y 3m 21d	female	LP	A	8.3	12.05.2017

The sequencing data of all 35 samples, were analysed for differential gene expression between HP and LP group using two different methods and computer algorithms for sequencing analysis. Both analyses were performed using the porcine reference genome assembly Sscrofa10.2.

3.1.1.1 Differential expression analysis

There is no gold standard for the computational analysis of gene expression data obtained by next generation sequencing. Therefore, two independent analysis pipelines using the same porcine genome annotation but different software for the analysis were utilised to increase specificity and true positive rates.

Data analysis pipeline 1

The analysis was performed in collaboration with Prof. Dr. Fries and later with Dr. Hongen Xu. The DNA fragments sequenced, called reads, from each sample were aligned to the porcine reference genome Sscrofa10.2 using STAR aligner (Dobin et al, 2013). Quality assessment using FASTQC, revealed an average of 30 x 10^6 reads per sample sequenced, showing a good average coverage with about 80% reads mapping to the reference genome. Reads that sequenced the exact same fragment more than once (called duplicates), were marked using MarkDuplicates tool of Picard (http://broadinstitute.github.io/picard). All reads that aligned to the reference genome (except duplicates) were assigned to annotated gene sequences as defined in the 10.2.77 porcine gene set and counted for each sample using featureCounts (Liao et al, 2014). The resulting files containing the amount of reads sequenced for each annotated gene of all the samples were then used for gene expression analysis using DESeq.2 (Love et al, 2014). The algorithm normalised all reads of each gene to the total number of reads of the sample, to allow accurate comparison also between samples of different total read numbers. The software then presented a table of genes that were found to be differentially expressed between the two groups sorted by p-value, to show those most significant at

the top (Table 33). Here the LP group was used as calibrator group, therefore the positive Log2FoldChange values signify higher expression in HP and values below 0 lower expression in HP.

Table 33 Differential expression results table of pipeline 1
adjusted p-value, the p-value multiplied by the number of comparisons in this case number of genes

Ensemble gene id	Log2Fold-Change	Fold-Change HP	P-value	Adjusted p-value	External gene name
ENSSSCG00000006238	**2.317689528**	**4.99**	**6.63x10^{-11}**	**1.45 x10^{-06}**	***CYP7A1***
ENSSSCG00000026852	1.830310386	3.56	1.10x10^{-06}	0.012098162	*NPPC*
ENSSSCG00000004114	0.82133311	1.77	2.06x10^{-05}	0.150472607	*ADGB*
ENSSSCG00000010529	**-1.07942183**	**0.47**	**4.89x10^{-05}**	**0.234395797**	***SFRP5***
ENSSSCG00000004578	-0.466422918	0.72	5.35 x10^{-05}	0.234395797	*ANXA2*
ENSSSCG00000002780	-1.253950352	0.42	7.32 x10^{-05}	0.26725109	*TPPP3*
ENSSSCG00000000398	-0.572581988	0.67	8.56 x10^{-05}	0.267797495	
ENSSSCG00000004968	0.476938546	1.39	0.000106493	0.287989837	*PAQR5*
ENSSSCG00000007727	0.44963094	1.37	0.000120242	0.287989837	*AUTS2*
ENSSSCG00000011201	**0.607588903**	**1.52**	**0.00013143**	**0.287989837**	***SATB1***
ENSSSCG00000001560	1.015986068	2.02	0.000177572	0.295143079	*C6orf222*
ENSSSCG00000029359	-0.725696293	0.6	0.000205919	0.295143079	*PHLDA3*
ENSSSCG00000029714	1.183631763	2.27	0.000212948	0.295143079	*BPIFB2*
ENSSSCG00000013919	-0.424007561	0.75	0.00022637	0.295143079	*HOMER3*
ENSSSCG00000007084	1.798886616	3.48	0.000242664	0.295143079	*BFSP1*
ENSSSCG00000014169	**0.666055769**	**1.59**	**0.000259109**	**0.295143079**	***PCSK1***
ENSSSCG00000029346	0.934230111	1.91	0.00026284	0.295143079	
ENSSSCG00000007038	2.093818278	4.27	0.000276619	0.295143079	
ENSSSCG00000014157	-0.586390819	0.67	0.000284305	0.295143079	*NR2F1*
ENSSSCG00000028518	0.781283303	1.72	0.000299205	0.295143079	

Data analysis pipeline 2

The analysis was performed in collaboration with Prof. Dr. Hubert Pausch. Here the reads were not aligned but pseudoaligned and assigned to 27,370 porcine transcripts obtained from ensembl version 89 (ftp://ftp.ensembl.org/pub/release-89/fasta/sus_scrofa/cdna/) (compatible to reference genome assembly Sscrofa10.2) using kallisto software (Bray et al, 2016). Kallisto also quantified the abundance of reads assigned per transcript. Differential transcript expression between HP and LP was estimated with normalised reads using sleuth (Pimentel et al, 2017). WALD test was performed to obtain a regression coefficient that approximates Log2FoldChange and p-value and adjusted p-value.

That resulted in a table of genes that were differentially expressed, sorted by WALD p-value. As HP animals were taken as calibrator group here, positive WALD Log2FoldCahnge values signify higher expression in LP animals.

Table 34 Differential expression results table of pipeline 2

Transcript id	WALD Log2Fold-Change	Fold Change LP	p-value	adjusted p-value	External gene name
ENSSSCT00000006836.2	**-1.5846592**	**0.33**	**$1.15x10^{-08}$**	**0.00021018**	***CYP7A1***
ENSSSCT00000012267.2	**-0.4078281**	**0.75**	**0.00025069**	**0.55458713**	***SATB1***
ENSSSCT00000028912.1	-0.628939	0.65	0.00025155	0.55458713	*UPP2*
ENSSSCT00000011519.2	**0.65411265**	**1.57**	**0.00025608**	**0.55458713**	***SFRP5***
ENSSSCT00000025876.1	-1.6445094	0.32	0.00033099	0.55458713	*RAP2C-AS1*
ENSSSCT00000004901.1	-0.9991806	0.5	0.00043884	0.55458713	*COL10A1*
ENSSSCT00000007343.2	-0.4858069	0.71	0.000441	0.55458713	*NBPF6*
ENSSSCT00000034564.1	1.53329781	2.89	0.00044682	0.55458713	*HSD17B10*
ENSSSCT00000005400.2	0.3478905	1.27	0.00049678	0.55458713	*SERPINB5*
ENSSSCT00000015626.2	-0.2996278	0.81	0.00054336	0.55458713	*JADE2*
ENSSSCT00000024538.1	0.44512355	1.36	0.00071622	0.55458713	*PHLDA3*
ENSSSCT00000030802.1	-2.2810175	0.21	0.00076314	0.55458713	*KIF16B*
ENSSSCT00000026609.1	-1.0035712	0.5	0.00076783	0.55458713	*STK32A*
ENSSSCT00000009725.2	1.60534411	3.04	0.00080014	0.55458713	*RPL15*
ENSSSCT00000015478.2	**-0.4153519**	**0.75**	**0.00082432**	**0.55458713**	***PCSK1***
ENSSSCT00000023378.1	-0.9929958	0.5	0.00084108	0.55458713	*BCL2L11*
ENSSSCT00000024120.1	-0.747594	0.6	0.00086092	0.55458713	*GABRD*
ENSSSCT00000029652.1	-1.2132739	0.43	0.00087033	0.55458713	*CYP2B6*
ENSSSCT00000028303.1	-1.1019999	0.47	0.00088974	0.55458713	*CYP2B7*
ENSSSCT00000004132.2	-0.7025889	0.61	0.00095227	0.55458713	*MEP1B*

The two independent data analysis approaches, presented different results among their top 20 differntially expressed genes. Only 4 were identical in both pipelines and only three of them were among the top ten in both pipelines: *CYP7A1*, *SFRP5* and *SATB1*. Only *CYP7A1* was significantly higher expressed in HP animals in both analyses according to p-value and adjusted p-value (multiple comparison adjustment) below 0.05. This gene was specifically interesting because high expression of this gene in the liver has been found to associate with bile acid-mediated CRC promotion in humans (Gadaleta et al, 2017; Hagiwara et al, 2005).The gene *SATB1* was also higher expressed in the HP group, significantly according to p-value but not significantly according to the adjusted p-value. This gene was also of interest as high *SATB1* expression was found in human CRC, where it promotes tumorigenesis and tumour progression(Al-Sohaily et al, 2014; Brocato & Costa, 2015; Lv et al, 2016; Mir et al, 2016; Zhang et al, 2014c).The gene *SFRP5*, a WNT antagonist that has been found methylation-silenced in human cancers including CRC (Samaei et al, 2014; Takagi et al, 2008; Veeck et al, 2008) showed significantly lower expression in the HP group according to p-value, however not significant on multiple comparison level (adjusted p-value) in both analyses. The detected differential expression was in all three cases in accordance with literature.

Quantitative reverse transcription PCR (RT-qPCR) validation of these three genes in the samples sequenced, using the primers SFRP5_Ex2_F1, SFRP5_Ex3_R1, SATB1_Ex6_F3, SATB1_Ex7_R3, Cyp7a1_Ex5_F1 and Cyp7a1_Ex6_R1-2, confirmed reduced expression of *SFRP5* and higher expression of *SATB1* and *CYP7A1* in HP animals. However, like in the sequencing analyses, where the adjusted p-values were above 0.05, the differential expressions of *SATB1* (1.5-fold higher expression in HP) and *SFRP5* (0.6-fold lower expression in HP) were not significant. The differential expression of *CYP7A1* (5.6 fold higher in HP) was in accordance with the sequencing analyses highly significant. As expression levels of *CYP7A1* were rather low, the results were confirmed with five-fold higher template concentration, showing even clearer expression differences (7.6-fold higher expression in HP).

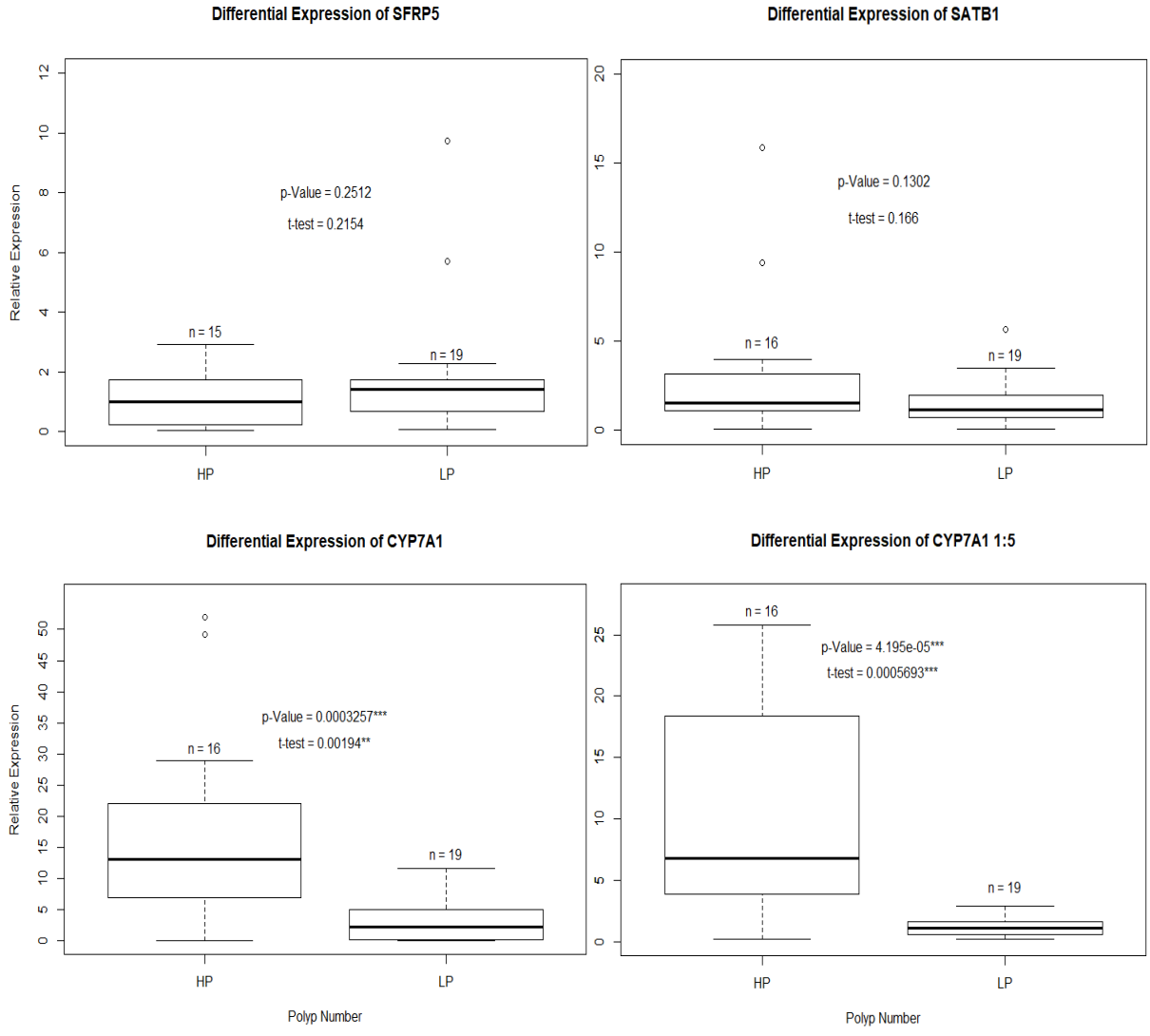


Figure 5 Differential expression validation of SFRP5, SATB1 and CYP7A1 using RT-qPCR with the primers SFRP5_Ex2_F1, SFRP5_Ex3_R1, SATB1_Ex6_F3, SATB1_Ex7_R3, Cyp7a1_Ex5_F1 and Cyp7a1_Ex6_R1-2.

Gene set enrichment analysis

Sequencing analysis showed only one gene, *CYP7A1*, significantly differentially expressed after multiple comparison adjustment. It is possible that not only one gene mediates susceptibility to severe polyposis but that a number of genes collectively contribute to significant differential pathway regulation. Therefore, gene set enrichment analysis was performed using differential expression data generated using both pipelines. For both data sets, differentially expressed genes with a p-Value below 0.05 and a Human Genome Organisation (HUGO) gene symbols were used. Resulting in 1779 genes as input for pipeline1 (file:///C:/Users/js_ca/gsea_home/output/mai29/Classic.GseaPreranked.-1527607258798/index.html). The pipeline2 data was further reduced by excluding genes with multiple transcripts to an input of 1062 genes (file:///C:/Users/js_ca/gsea_home/output/mai29/my_analysis.-GseaPreranked.1527614804154/index.html). Hallmark gene sets that represent and summarise clearly-defined biological processes or states were used. The analysis of both datasets showed similar results (Table 35 and Table 36). Few gene sets were enriched in the HP group compared to a large amount of enriched gene sets enriched in LP. The gene set of oestrogen response was significantly enriched in the HP group analysed using pipeline 2 (false discovery rate (FDR) < 0.25). Gene sets significantly enriched in LP animals included DNA repair, UV response, p53 and apoptosis components, which are guarding the integrity and functioning of the genome and cell. However, gene sets associated with oncogenic pathways such as MYC targets and mesenchymal transition were also enriched significantly.

Table 35 Gene sets enriched in HP animals
size, number of genes in the gene set; ES, enrichment score; NES, normalised enrichment score across analysed sets; FDR, false discovery rate; Rank at max, position in the ranked list at which the maximum enrichment score occurred

Pipeline 1	**SIZE**	**ES**	**NES**	**p-value**	**FDR**	**RANK AT MAX**
Hallmark oestrogen response early	23	0.22	1.28	0.19	0.69	685
Hallmark unfolded protein response	17	0.18	0.91	0.57	1.00	1444
Hallmark mitotic spindle	18	0.17	0.84	0.64	0.87	1179
Hallmark oestrogen response late	31	0.09	0.56	0.96	0.96	21
Pipeline 2						
Hallmark oestrogen response early	16	-0.29	-1.40	0.11	0.21	163
Hallmark heme metabolism	16	-0.13	-0.62	0.93	0.92	594

Table 36 Gene sets enriched in LP animals

size, number of genes in the gene set; ES, enrichment score; NES, normalised enrichment score across analysed sets; FDR, false discovery rate; Rank at max, position in the ranked list at which the maximum enrichment score occurred

Pipeline 1	SIZE	ES	NES	p-value	FDR	RANK AT MAX
Hallmark oxidative phosphorylation	76	-0.32	-3.23	0.00	0.00	1225
Hallmark MYC targets v1	67	-0.35	-3.22	0.00	0.00	1122
Hallmark E2F targets	57	-0.34	-2.97	0.00	0.00	1029
Hallmark DNA repair	34	-0.37	-2.51	0.00	0.00	807
Hallmark G2M checkpoint	31	-0.35	-2.36	0.00	0.00	1093
Hallmark adipogenesis	38	-0.30	-2.20	0.01	0.00	1199
Hallmark MTORC1 signalling	40	-0.28	-2.12	0.00	0.01	1012
Hallmark hypoxia	30	-0.29	-1.90	0.01	0.02	376
Hallmark epithelial mesenchymal transition	21	-0.34	-1.85	0.01	0.03	337
Hallmark UV response up	23	-0.31	-1.84	0.02	0.03	836
Hallmark fatty acid metabolism	37	-0.25	-1.78	0.02	0.04	1098
Hallmark glycolysis	30	-0.27	-1.77	0.02	0.03	359
Hallmark apoptosis	26	-0.29	-1.77	0.02	0.03	854
Hallmark myogenesis	18	-0.30	-1.51	0.07	0.11	1043
Hallmark p53 pathway	34	-0.20	-1.37	0.12	0.18	853
Hallmark IL2 STAT5 signalling	17	-0.26	-1.32	0.14	0.21	1096
Hallmark xenobiotic metabolism	23	-0.23	-1.32	0.16	0.20	900
Hallmark PI3K AKT MTOR signalling	15	-0.25	-1.19	0.23	0.30	1203
Hallmark heme metabolism	22	-0.20	-1.11	0.29	0.36	1098
Hallmark allograft rejection	17	-0.18	-0.91	0.55	0.62	1446
Hallmark complement	18	-0.15	-0.75	0.77	0.82	1409
Hallmark apical junction	19	-0.11	-0.61	0.94	0.93	267
Pipeline 2						
Hallmark MYC targets v1	49	0.48	3.98	0.00	0.00	569
Hallmark oxidative phosphorylation	53	0.45	3.73	0.00	0.00	610
Hallmark E2F targets	39	0.45	3.36	0.00	0.00	573
Hallmark MTORC1 signalling	25	0.52	3.04	0.00	0.00	441
Hallmark fatty acid metabolism	26	0.44	2.72	0.00	0.00	563
Hallmark adipogenesis	28	0.41	2.60	0.00	0.00	556
Hallmark DNA repair	26	0.41	2.43	0.00	0.00	514
Hallmark G2M checkpoint	24	0.36	2.09	0.00	0.01	550
Hallmark glycolysis	19	0.31	1.58	0.04	0.09	573
Hallmark UV response up	16	0.32	1.56	0.06	0.09	392
Hallmark hypoxia	24	0.25	1.44	0.07	0.13	360
Hallmark epithelial mesenchymal transition	16	0.26	1.24	0.21	0.27	390
Hallmark apoptosis	24	0.19	1.13	0.29	0.36	462
Hallmark p53 pathway	22	0.19	1.07	0.34	0.41	234
Hallmark xenobiotic metabolism	20	0.17	0.89	0.56	0.62	567
Hallmark oestrogen response late	22	0.14	0.79	0.72	0.72	382

3.1.1.2 SNP identification and allele-specific expression analysis

SNPs associated with severe polyposis and sporadic CRC risk have been identified in humans (Ghorbanoghli et al, 2016). Although those SNPs identified are located outside human exonic regions, the analysis of the RNA sequencing data for differentially expressed SNPs may reveal porcine equivalents. To identify SNPs and determine differential association and allele-specific expression differences, sequencing reads were aligned to the Sscrofa11.1 reference genome using STAR aligner (Dobin et al, 2013) by Prof. Dr. Hubert Pausch. Duplicates were marked using Picard tools (https://broadinstitute.github.io/-picard/). The reads were assigned to exons using SplitNCigarReads tool from the GATK software suite (DePristo et al, 2011). GATK's Haplotypecaller was used to identify SNPs and analyse them for differential association between HP and LP. The resulting genotypes were tested for non-random association between LP and HP using Fisher exact tests of allelic association using Plink (version 1.9) (Chang et al, 2015). The top 30 are shown in the table below (Table 37). The SNPs identified showed significance on p-value level. Porcine homologous regions of the known human SNPs associated with severe polyposis in FAP and sporadic CRC risk (rs16892766 at 8q23.3 and rs3802842 at 11q23.1) are located on chromosomes 9 and 4 in pigs (Ghorbanoghli et al, 2016). Although one SNP differentially associated between LP and HP is located on chromosome 4 it is far away from the region homologous to the human SNP location.

Table 37 The top 30 SNPs detected.

Chromosome	Base pair position	P-value	Chromosome	Base pair position	P-value
13	192206011	$1.08x10^{-08}$	1	268990705	$1.36x10^{-06}$
AEMK02000452.1	762390	$5.76x10^{-08}$	39	79426	$1.36x10^{-06}$
17	31868703	$1.65x10^{-07}$	6	71775092	$1.55x10^{-06}$
12	2263873	$1.66x10^{-07}$	18	8263672	$1.55x10^{-06}$
1	44967406	$3.25x10^{-07}$	13	157511545	$1.58x10^{-06}$
7	36791633	$3.39x10^{-07}$	1	166171033	$1.79x10^{-06}$
2	142947459	$5.10x10^{-07}$	1	161756743	$1.90x10^{-06}$
14	106776396	$5.10x10^{-07}$	17	8451491	$2.29x10^{-06}$
16	71219623	$5.10x10^{-07}$	17	30781986	$2.29x10^{-06}$
AEMK02000452.1	1042678	$6.04x10^{-07}$	13	135442206	$2.31x10^{-06}$
1	253094682	$7.48x10^{-07}$	2	151688112	$2.69x10^{-06}$
1	253094686	$7.48x10^{-07}$	4	111634710	$2.98x10^{-06}$
1	251101477	$1.05x10^{-06}$	17	30781990	$3.13x10^{-06}$
12	52303005	$1.05x10^{-06}$	1	268256223	$3.22x10^{-06}$
17	32654033	$1.06x10^{-06}$	10	46791255	$3.22x10^{-06}$

Allele-specific expression analysis

SNPs can not only be distributed differentially between two groups, but they can also show imbalanced expression associating with HP or LP animals. Allele-specific expression analysis was performed to identify SNPs that are expressed more abundantly in HP or LP and may modulate or promote severe polyposis. SNPs identified with the GATK's Haplotypecaller were filtered for strand unbiased SNPs of certain confidence that were heterozygous in at least 4 LP and HP animals with at least 30 reads. The probability of allelic imbalance for each SNP was calculated based on the number of reference and alternate allele reads in heterozygous animals using a two-sided binomial test that was implemented with the binom function in R. The SNPs that showed differential allele-specific expression between HP and LP were plotted according to their decadic logarithmic p-value versus chromosomal location (Figure 6).

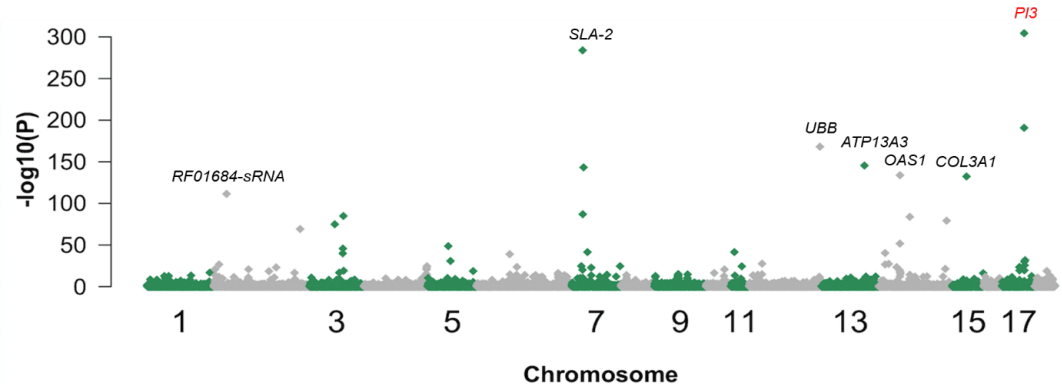


Figure 6 Display of the allele-specific SNPs expressed differentially, plotted by p-value and chromosome.

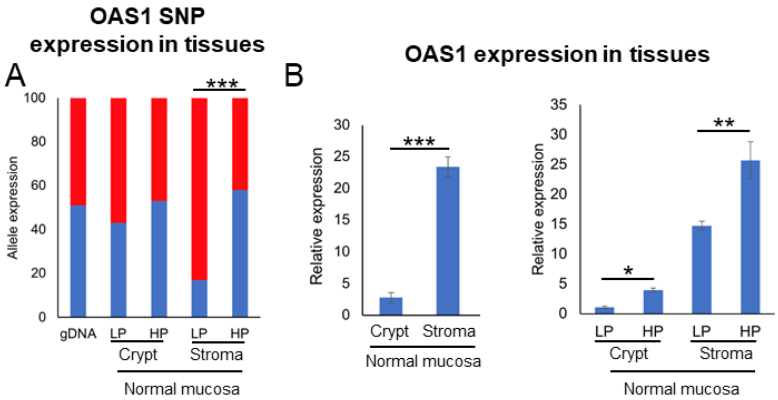


Figure 7 Allele-specific expression of the SNP located in *OAS1*.
gDNA, genomic DNA; Red, T variant; Blue, C variant.

SNPs in the genes *PI3*, *SLA-2*, *UBB*, *ATP13A3*, *OAS1*, *COL3A1* and *RF01684-sRNA* showed among others high significance in differential allele-specific expression. From these selected SNPs, the SNP in the gene *OAS1* on chromosome 14, position 38856577 bp C/T was validated using pyrosequencing (Figure

7). *OAS1* is an interferone-induced enzyme that is involved in cellular innate antiviral response (http://www.genecards.org/). Therefore, pyrosequencing of laser microdissected epithelium from normal mucosa samples was performed to identify whether the imbalanced SNP expression originates from the epithelium or the stroma. The SNP variant T (in red) is significantly higher expressed in the stroma of LP animals compared to HP animals and expression analysis showed that the SNP variant T is associated with reduced *OAS1* expression(Figure 6).

In summary mRNA Sequencing analysis helped to identify one potential modifier gene CYP7A1 which showed significantly higher expression in HP animals in both sequencing-based and PCR-based methods. Gene set enrichment analysis of the mRNA sequencing data identified estrogene response pathway enriched in HP. The differential distribution and allele specific expression between HP and LP animals showed the SNP in the *OAS1* gene on chromosome 14 position 38856577 bp C/T has higher expression of the cytosine allele, associated with higher overall expression in the HP animals.

3.1.1.3 CYP7A1

3.1.1.3.1 Elucidating gene structure of *CYP7A1*

The gene structure of *CYP7A1* was analysed to identify the source of differential expression. As many proteins are highly conserved between human and pig, the known DNA and amino acid sequences of porcine and human CYP7A1 were aligned. Protein alignment showed 80% identity between human and pig (UniprotKB). Alignment of the genomic CYP7A1 sequences revealed, different to the human sequence, two more untranslated exons more than 10 kb 5' of the ATG in the porcine *CYP7A1* gene annotation (Sscrofa11.1, ensemble genome browser 93) (Figure 8)

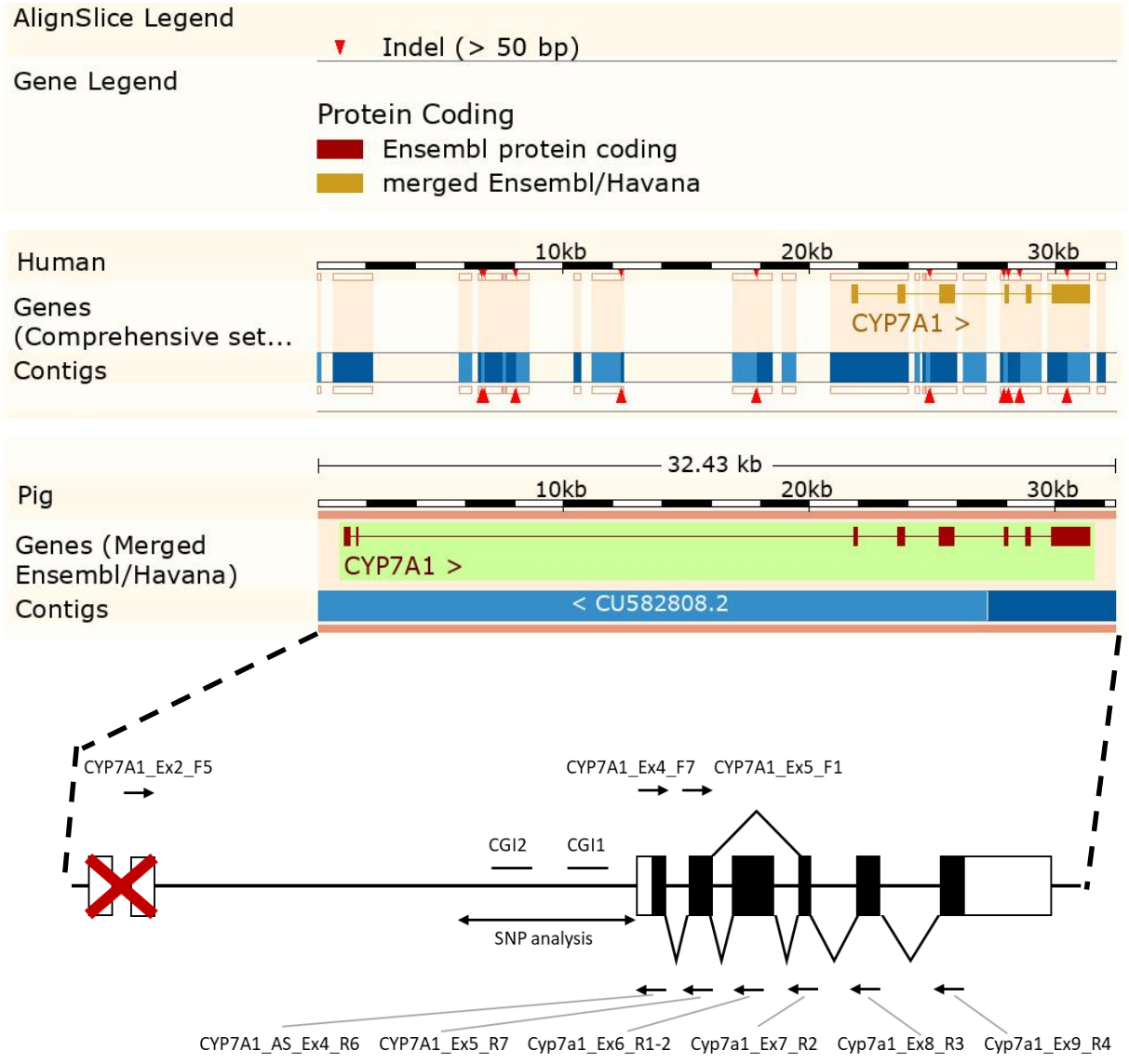
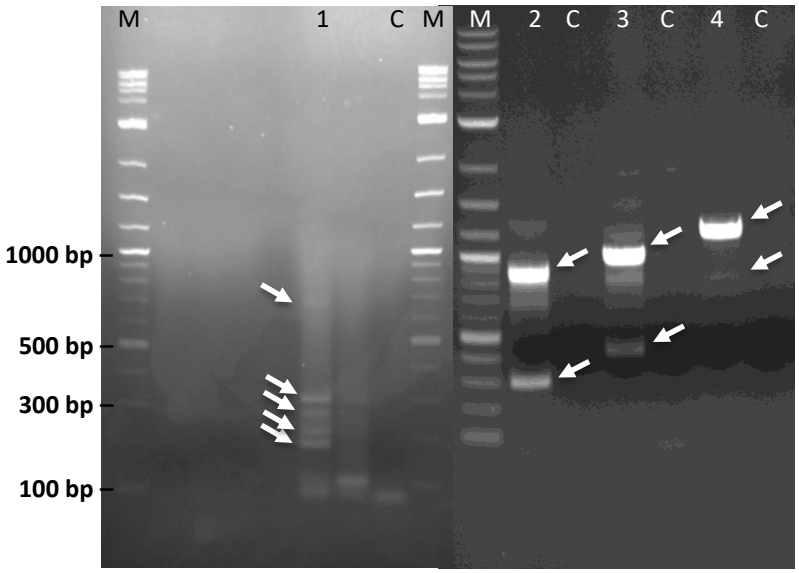


Figure 8 Ensembl genomic alignment of porcine CYP7A1 with human CYP7A1 (genome assemblies Sscrofa11.1 and GRCh38.p12) and depiction of *CYP7A1* structure in colon mucosa of *APC*^{1311} pigs
The first two exons annotated according to ensemble Sscrofa11.1 were not found to be expressed. The primers used for RT-PCR, RACE sand RT-qPCR are marked. CpG islands (CGI) 1 and 2 are indicated. The region of SNP analysis 5742 bp 5' of the ATG is indicated by an arrow.

The DNA alignment showed no association of porcine exon 1 and 2 with the human CYP7A1 gene.

Therefore, RACE together with RT-PCR was performed to investigate the presence or absence of these two exons (Figure 8 and Figure 9). The resulting bands were gel extracted, subcloned into the pGEMT vector system (Promega) and sequenced. 5'RACE from the third exon (lane 1) and sequencing showed a lack of exon 1 and 2 (Figure 9). Other RT PCRs from exon 2 (not shown here) did not produce products. Furthermore, RT-PCRs from exon 3 to exon 6, 3-7 and 3-8 confirmed the presence of 6 exons of *CYP7A1* in the porcine colon mucosa of the analysed animals and revealed the presence of two transcript variants visible by two bands in lane 2, 3 and 4 respectively. The higher molecular weigth bands showed the expected sizes of 898 bp, 1039 bp and 1248 bp. Sequencing of both the higher and lower molecular weight bands revealed one transcript variant with all 6 exons transcribed (higher molecular weigth band) and one variant that skips transcription of exon 5 (lower molecular weigth band). Taken together, the pigs analysed in this study express *CYP7A1* composed of 6 exons, with two transcription variants, one of which skips exon 5 (Figure 8 and Figure 9). It should be noted however, that the differential expression determined via RT-qPCR was performed with primers positioned in exon 4 and 5. Thus the transcription variant containing exon 5 was found to be differentially expressed.



	Exon5 included	Exon5 skipped
Ex3-6	898 bp	320 bp
Ex3-7	1039 bb	461 bp
Ex3-8	1248 bp	670 bp

Figure 9 Gel photographs of the 5' RACE and the RT-PCR from exon 4-9.
Lane 1: 5'RACE from exon 4 to the 5' end of the mRNA using primer CYP7A1_Ex5_R7. Lane 2, 3,4: RT-PCR from exon 3-6, 3-7 and 3-8 respectively using the primers CYP7A1_Ex4_F7, CYP7A1_Ex7_R2, CYP7A1_Ex8_R3 and CYP7A1_Ex9_R4. M: marker, C: water control.

3.1.1.3.2 Analysis of *cis*-regulation of CYP7A1

Once the structure of the *CYP7A1* gene in the normal mucosa samples was analysed, the source of the differential expression was investigated on the level of cis-regulatory elements (CREs). CREs include promoters and transcription factors (TFs) that exhibit enhancing and silencing effects on gene expression. Therefore, the potential promoter region was analysed for altered TF binding sites. First, SNP analysis 5742 bp 5' of the ATG was performed (Figure 8) using Sanger sequencing with the primers CYP7A1_AS_Ex3_F6, CYP7A1_AS_Ex4_R6 and all CYP7A1_SNPSeq primers (Table 13). 78 SNPs were identified and analysed. 20 showed a different allelic distribution between HP and LP of at least 10 % difference. However, none showed significant association with either LP or HP. As differential expression was validated in the transcript variant with exon 5, intron 4 was sequenced to identify the cause of exon skipping using the primers CYP7A1_Ex5_F1, CYP7A1_Ex6_R1-2, CYP7A1_SNP_I5_F, CYP7A1_SNP_I5_F, CYP7A1_SNP_I5_R, CYP7A1_SNP_I5_R2. Severe heterozygous deletions and insertions in this region allowed no clear analysis.

Also epigenetic changes in the CpG methylation of the potential promoter region can alter TF binding sites, that in turn influences gene expression. Therefore, two CpG islands (CGI) up to 3459 bp 5' of the ATG, in the potential promoter region were analysed (Figure 8). 4 CpG sites were examined in CGI1 with the primers CYP7A1_CpG1_F1, CYP7A1_CpG1_R1_BIO and CYP7A1_CpG1_S2 and 3 CpG sites in CGI2 with the primers CYP7A1_CpG2_F1, CYP7A1_CpG2_R1_BIO and CYP7A1_CpG2_S using pyrosequencing (Figure 10).

No methylation differences between HP and LP were detected at CGI1 for any of the four CpG sites. At CGI2 significant methylation differences were detected at CpG site 1, that showed significantly higher methylation in the LP group compared to HP. Reduced methylation of the CpG site in HP animals may cause increased binding of both enhancing but also silencing TFs.

Therefore, the sequence 45 bp upstream and 45 bp downstream of the CpG site was analysed for TF binding sites using MatInspector by Genomatix Matrix Library 11.0 with standard parameters (Cartharius et al, 2005; Quandt et al, 1995) to determine whether the reduced CpG methylation in HP animals induces higher *CYP7A1* expression via improved binding of activating TFs. This resulted in a list of 21 TFs binding to this sequence. Those that included the CpG in their binding site, but not their core sequence, were TF families signal transducer and activator of transcription 3 (STAT3), GLIS family zinc finger 1 (GLIS1), Ccaat/enhancer binding protein beta (CEBPB) and GA binding protein TF alpha (GABPA) (Figure 11).

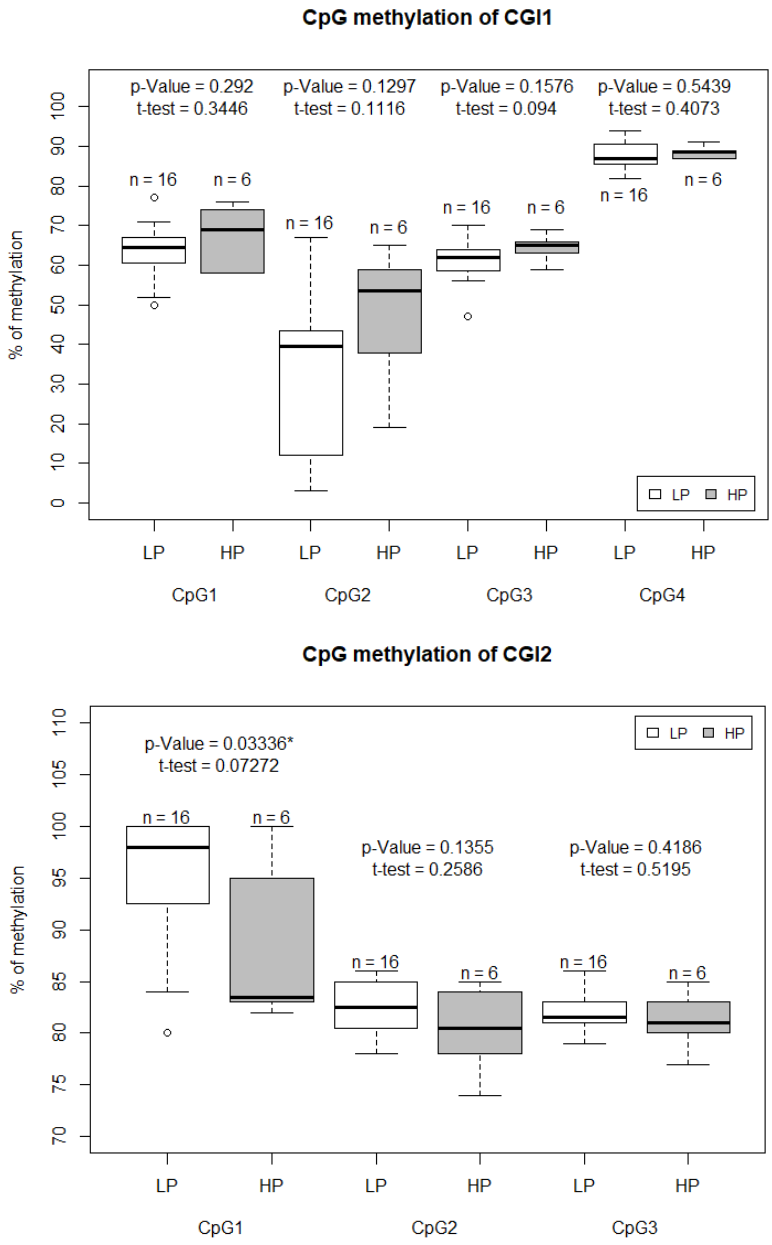


Figure 10 CGI methylation analysis of CGI1 and CGI2.

Genecards search showed expression of all 4 TFs in the colon and components of the immune system (http://www.genecards.org/). STAT3 is a transcription activator that responds to cell stimuli through cytokines and growth factors and is known to have oncogenic effects on proliferation, invasion and metastasis (Yu et al, 2014). GLIS1 functions are still widely unknown, however it is associated with reprogramming efficiency (Jetten, 2018). CEBPB regulates genes involved in immune and inflammatory response with tumour-promoting effects via NF-κB signalling (http://www.genecards.org/, NCBI) (Yang et al, 2017). GABPA functions are unknown.

GABPA **cpg**ttgtgtGGAAgagggacgg

CEBPB **cpg**ttgTGTGgaagag

STAT3 gtgcTTCCtgggcttt**cpg**t

CTCAGCCCACCAAGGAGAACTGGGTGTGTGTGCTTCCTGGGCTTT**CpG**TTGTGTGGAAGAGGGACGGTGTGGCCAAAGCGGGAAGACCAGC

GLIS1 a**cpg**aaagccCAGGaagc

Figure 11 Display of the four transcription factors (TFs) with their binding sequence.
The core sequence was capitalised and the CpG site marked by bold writing.

CYP7A1 in the normal mucosa of the pigs analysed is transcribed in two different variants in the colonic mucosa. The variant including all 6 exons (not skipping exon 5) is higher expressed in HP compared to LP animals. The cause of differential expression on *cis*-regulatory level could not be associated with differentially distributed SNPs in the potential promoter region but with reduced CpG methylation in the CGI2 of the potential promoter. The reduced methylation may influence the binding of TFs GLIS1, CEBPB, GABPA and STAT3, which is associated with oncogenic effects.

3.1.1.3.3 Function and source of high CYP7A1 expression in high polyp animals

CYP7A1 is an endoplasmic reticulum membranous monooxigenase that contributes to drug metabolism and is the rate limiting enzyme of bile acid synthesis from cholesterol. High *CYP7A1* expression in the liver, is associated with a high amount of bile acids in the colon, which correlates with a high risk of CRC development. The function of CYP7A1 in the colon however, is not yet elucidated. There were reports of *CYP7A1* expression in macrophages (Bao et al, 2015). RT-qPCR validation of laser microdissected normal mucosa samples was aimed to reveal which cell type exhibits increased *CYP7A1* expression in HP animals. Crypts and surrounding tissue, called stroma, was isolated from HE stained cryo sections vial laser microdissection. RNA was isolated and cDNA was generated followed by RT-qPCR.. No significant difference between *CYP7A1* expression in crypts and stroma was detected. Comparison of the expression in LP crypts and HP crypts and LP stroma and HP stroma, showed a significantly higher expression of *CYP7A1* in HP stroma. This suggests that the CYP7A1 expression originates from cells located in the stroma.

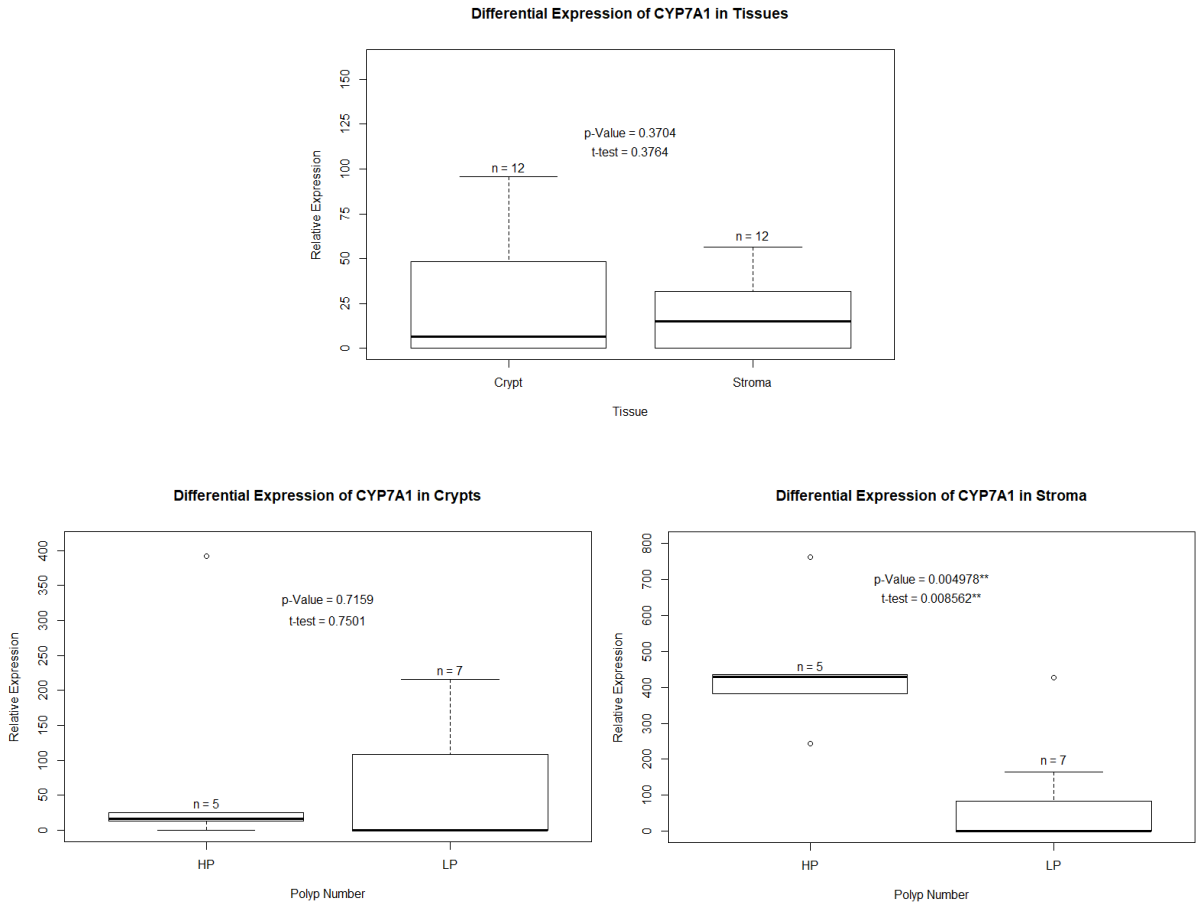


Figure 12 Differential expression analysis of CYP7A1 in crypts and stroma using RT-qPCR.

3.1.2 Attempt to identify modifier genes on miRNA level

MiRNA from 19 normal mucosa samples taken at 40 cm colorectum depth of animals aged 3-9 months (Table 38) were sequenced to identify those that might cause or modulate susceptibility to severe polyposis. After DNase treatment, quality and quantity assessment 1 µg total RNA was used for the library preparation. The resulting libraries were again checked for quality and quantity and 6 libraries were pooled (6 libraries/flow cell), clustered and sequenced on the MiSeq. The resulting sequencing data was used to identify miRNA differentially expressed between HP and LP animals.

Table 38 Animal samples sequenced for the analysis of modifier genes on miRNA level

Animal ID	Born	Collection	Age	Sex	Pheno-type	RIN	Sequenced
128	07.03.2013	16.07.2013	0y 4m 9d	female	LP	9.1	16.02.2016
145	14.03.2013	25.06.2013	0y 3m 11d	male	LP	8.1	16.02.2016
150	14.03.2013	16.07.2013	0y 4m 2d	female	LP	7.9	19.02.2016
152	14.03.2013	22.10.2013	0y 7m 8d	female	LP	8.9	11.02.2016
153	14.03.2013	23.10.2013	0y 7m 9d	female	LP	7.4	11.02.2016
155	14.03.2013	26.06.2013	0y 3m 12d	male	LP	9.2	11.02.2016
157	14.03.2013	17.07.2013	0y 4m 3d	female	LP	8.8	16.02.2016
163	15.03.2013	19.09.2013	0y 6m 4d	female	HP	9.1	11.02.2016
168	12.05.2013	18.09.2013	0y 4m 6d	male	HP	8.7	16.02.2016
173	12.05.2013	02.12.2013	0y 6m 20d	female	LP	7.3	16.02.2016
251	21.02.2014	24.06.2014	0y 4m 3d	castrated male	HP	6.6	18.02.2016
252	21.02.2014	25.06.2014	0y 4m 4d	castrated male	HP	7.5	18.02.2016
253	21.02.2014	24.06.2014	0y 4m 3d	castrated male	HP	6.6	19.02.2016
300	24.08.2014	16.03.2015	0y 6m 20d	female	LP	4.3	18.02.2016
322	18.10.2014	16.03.2015	0y 4m 26d	male	HP	5.7	19.02.2016
324	18.10.2014	16.03.2015	0y 4m 26d	male	HP	6.1	19.02.2016
326	18.10.2014	17.03.2015	0y 4m 27d	female	HP	6.6	18.02.2016
328	18.10.2014	17.03.2015	0y 4m 27d	female	LP	7	18.02.2016
339	20.10.2014	17.03.2015	0y 4m 25d	female	HP	5.3	19.02.2016

3.1.2.1 Differential expression analysis

The processing of miRNA sequencing data for differential expression analysis was performed by Dr. Stefan Bauersachs. Adapters (added in the process of library preparation) were removed from the reads and the read quality was assessed before and after this process using FastQC (v0.11.2) and multiqc (Galaxy Version 0.6). Sequences with read counts lower than 300 counts in sum of all samples were filtered out resulting in approx. 9000 sequences that were compared to all transcripts of Sus scrofa including non-coding RNAs and human and bovine sequences with NCBI BLAST+ (Cock et al, 2015) blastn-short. Duplicates were removed and sequences assigned to miRNAs were used for analysis of differential expression of miRNAs using EdgeR (Robinson et al, 2010). EdgeR presented a table of all sequences and stem sequences of mature miRNAs differentially expressed between HP and LP. Among the top 20 stem sequences of mature miRNA, 19 miRNAs were significantly differential expressed between LP and HP according to their p-value (Table 39). Only one of those, miR-215 was significantly higher expressed in HP animals according to both p-value and adjusted p-value (multiple comparison adjustment), here called FDR (false discovery rate). Here the LP group was used as calibrator group, therefore the positive Log2FoldChange values signify higher expression in HP and

values below 0 lower expression in HP. The miRNAs miR-215 (higher expressed in HP), miR-194b-5p (higher expressed in HP), miR-27a-3p (lower expressed in HP) and miR-146a-5p (lower expressed in HP) were chosen for RT-qPCR validation.

Table 39 Top 20 differentially expressed miRNAs

MiRNA	Iso-miRs	BLAST hits	Log2Fold-Change	Fold-Change	P-value	Adjusted p-value
mir-215	**4**	**ssc-miR-215**	**2.14**	**4.41**	**0.0002**	**0.0476**
mir-194b-5p	**3**	**ssc-miR-194b-5p**	**1.86**	**3.63**	**0.0009**	**0.072**
mir-27a-3p	**5**	**bta-miR-27a-3p,ssc-miR-27a-3p**	**-0.55**	**0.68**	**0.001**	**0.072**
mir-23a	4	ssc-miR-23a	-0.44	0.74	0.0088	0.4553
mir-192-5p-v1	6	miR-192-5p-v1	0.73	1.66	0.0139	0.4956
mir-146a-5p	**7**	**ssc-miR-146a-5p**	**-0.72**	**0.61**	**0.0173**	**0.4956**
let-7d-5p	2	ssc-let-7d-5p	-0.38	0.77	0.0247	0.4956
pre-mir-192-5p	12	pre-ssc-miR-192-5p	0.55	1.46	0.0247	0.4956
mir-375	13	bta-mir-375,hsa-miR-375	0.49	1.40	0.0292	0.4956
mir-192-5p	43	bta-miR-192-5p,ssc-miR-192-5p	0.39	1.31	0.0307	0.4956
mir-192-5p-v2	1	miR-192-5p-v2	0.67	1.59	0.0339	0.4956
mir-139-5p	1	ssc-miR-139-5p	0.55	1.46	0.0356	0.4956
mir-182-5p	9	bta-mir-182-5p,ssc-miR-182-5p	0.46	1.38	0.0367	0.4956
novel_mir_2	3	novel_miR_2	1.07	2.10	0.037	0.4956
mir-214	2	ssc-miR-214	-0.57	0.67	0.0384	0.4956
mir-6529a	1	bta-miR-6529a	-0.49	0.71	0.0394	0.4956
mir-192-5p-v3	1	miR-192-5p-v3	0.59	1.51	0.0431	0.4956
mir-582-3p	2	ssc-miR-582-3p	0.53	1.44	0.0445	0.4956
mir-92b-3p	9	bta-mir-92b-3p,ssc-miR-92b-3p	-0.39	0.76	0.046	0.4956
mir-340	2	ssc-miR-340	0.37	1.29	0.0501	0.4956

Establishment of a normalising miRNA for the analysed normal mucosa samples was essential to perform RT-qPCR validation of the top 3 differentially expressed miRNAs miR-215, 194b-5p, 27a-3p and the migration and invasion promoting miRNA 146a-5p (Lu et al, 2017). Therefore, literature search for human normaliser miRNA for normal colon and cancer samples lead to testing equivalent porcine sequences for RNU6B_1, miR-191-5p_1, miR-25-3p_1, miR-16-5p_1, miR-26a-5p_1, mir-425-5p_1, let-7a-5p_1 (Chang et al, 2010; Peltier & Latham, 2008; Schmitz et al, 2009). MiR-191-5p_1, miR-25-3p_1, miR-16-5p_1, and let-7a-5p were tested for their stability on a subset of the sequenced normal mucosa samples using NormFinder (Andersen et al, 2004). Ct values were transformed into a linear form by

generating 2^{-Ct} values for NormFinder analysis. The resulting stability value was calculated from intergroup and intragroup variation of comparing miRNA expression of MiR-191-5p_1, miR-25-3p_1, miR-16-5p_1, and let-7a-5p between LP animals 152, 155, 173, 328 and HP animals 163, 168, 253 and 339 (Table 40). The smallest stability value indicated the lowest variation. Let-7a-5p showed the best intra- and inter-group stability value of 0.065 and was therefore taken as a normaliser in the validation of the sequencing results.

Table 40 NormFinder analysis.

miR name	Without group identifiers		With group identifiers
	Stability value	Standard error	Stability value
miR-25-3p	**0.140**	**0.062**	0.073
miR-16-5p	0.211	0.068	0.098
miR-191-5p	0.192	0.066	0.093
let7a-5p	0.160	0.063	**0.065**

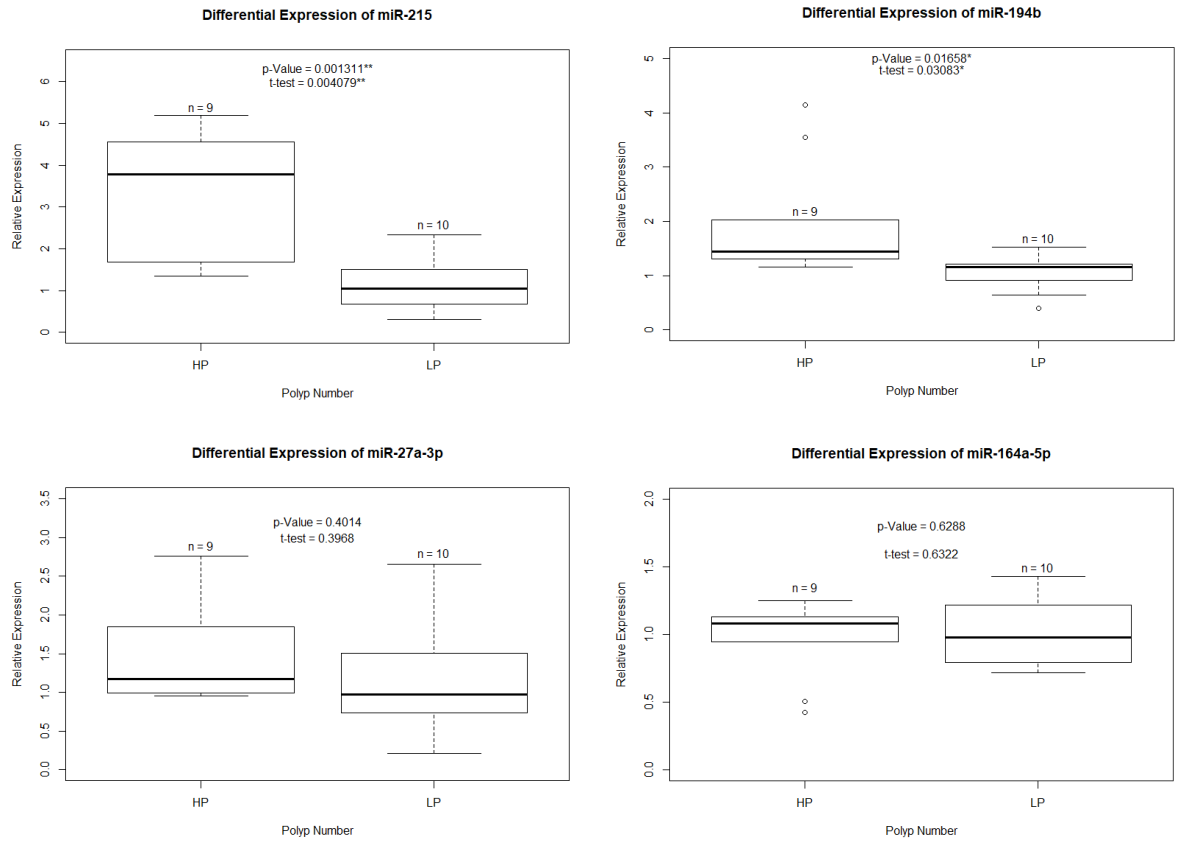


Figure 13 Differential expression analysis of miR-215, 194b, 27a-3p and 146a-5p in the sequenced samples using RT-qPCR.

RT-qPCR Validation of miR-215, 194b, 27a-3p and 146a-5p expression confirmed significantly increased miR-215 expression in HP animals in accordance with sequencing results (Table 39). Higher expression of miR-194b in HP animals was significant in RT-qPCR (Figure 13). The reduced expression of miR-27a-3p and 146a-5p detected in the sequencing analysis could not be confirmed (Figure 13).

The RT-qPCR results of miR-215 and 194b confirmed the sequencing results. Therefore we analysed a number of additional samples (909-953, 2.1.15), that had not been sequenced (Figure 14). Higher MiR-215 expression was significant when the additional samples were analysed together with the sequenced samples and alone, highlighting the strength of this biological effect (Figure 14 A, C). Increased miR-194b expression in HP could only be confirmed in comibination with the sequenced samples but not in the additional samples alone (Figure 14 B, D).

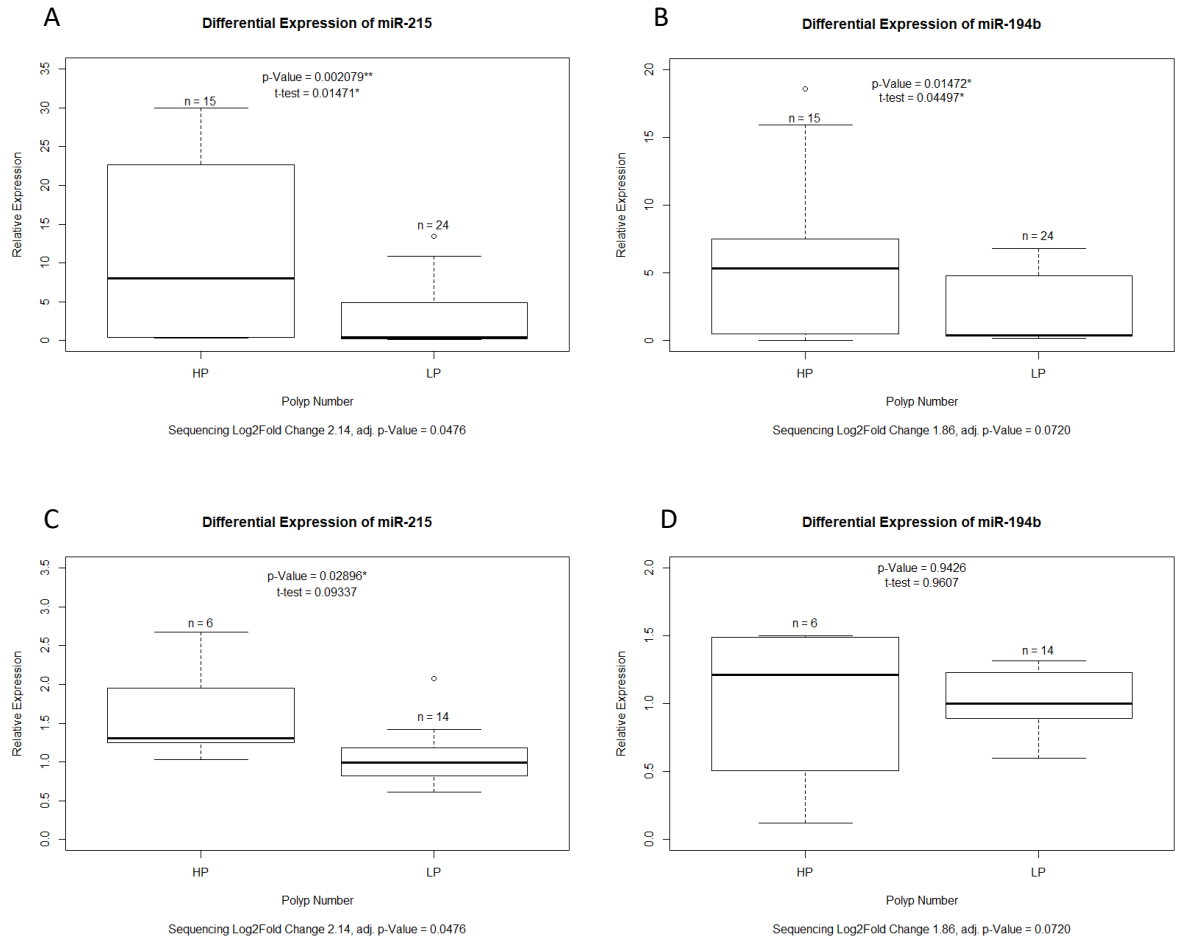


Figure 14 Differential expression analysis of miR-215 and 194b in the sequenced and additional samples (A, B) and in the additional samples alone(C,D).

3.1.2.2 *In silico* miRNA target analysis

In silico target analysis was performed with all differentially expressed miRNAs with p-values < 0.05 and a human equivalent using Diana tools (2.6.2). The KEGG pathway analysis showed that the differentially expressed miRNAs miR-215-5p, 194-5p, 27a-3p, 23a-3p, 192-5p, 146a-5p, let-7d-5p, miR-375-3p, 139-5p, 182-5p, 214-3p, 582-3p and 92b-3p influence many cancer associated pathways (Table 41).

Table 41 *In silico* miRNA target analysis of differentially expressed miRNAs between HP and LP animals.

KEGG pathway	**P-value**	**Number of genes targeted**	**Number of miRNAs involved in pathway**
Proteoglycans in cancer	$1.77x10^{-17}$	130	13
Adherens junction	$2.71x10^{-12}$	57	12
Protein processing in endoplasmic reticulum	$1.80x10^{-11}$	115	12
Cell cycle	$9.47x10^{-11}$	87	14
Hippo signalling pathway	$2.21x10^{-10}$	93	13
TGF-beta signalling pathway	$9.53x10^{-10}$	56	12
Pathways in cancer	$2.59x10^{-09}$	231	14
Viral carcinogenesis	$1.24x10^{-08}$	122	13
Renal cell carcinoma	$1.69x10^{-08}$	48	12
Glioma	$1.69x10^{-08}$	46	12
Chronic myeloid leukemia	$2.68x10^{-08}$	54	12
Prion diseases	$3.00x10^{-08}$	20	9
Hepatitis B	$8.92x10^{-08}$	90	12
Signalling pathways regulating pluripotency of stem cells	$3.39x10^{-07}$	90	12
Colorectal cancer	$4.52x10^{-07}$	46	13
Prostate cancer	$5.06x10^{-07}$	63	12
Oestrogen signalling pathway	$7.05x10^{-07}$	61	13
Oocyte meiosis	$1.45x10^{-06}$	71	13
Ubiquitin mediated proteolysis	$5.00x10^{-06}$	86	12
mTOR signalling pathway	$9.60x10^{-06}$	45	12
Neurotrophin signalling pathway	$9.60x10^{-06}$	76	13
Pancreatic cancer	$1.05x10^{-05}$	46	12
Mucin type O-Glycan biosynthesis	$1.46x10^{-05}$	17	6
Bacterial invasion of epithelial cells	$2.06x10^{-05}$	50	12
Bladder cancer	$2.40x10^{-05}$	30	12
Endocytosis	$3.50x10^{-05}$	118	12
AMPK signalling pathway	$3.66x10^{-05}$	79	14
Shigellosis	$3.75x10^{-05}$	45	12
Thyroid cancer	$3.86x10^{-05}$	22	12
FoxO signalling pathway	$4.09x10^{-05}$	84	12

Endometrial cancer	$4.59x10^{-05}$	37	12
ErbB signalling pathway	$5.38x10^{-05}$	57	14
Glycosaminoglycan biosynthesis - heparan sulphate / heparin	0.00012515	15	9
Non-small cell lung cancer	0.00013151	38	12
Transcriptional misregulation in cancer	0.00014138	96	12
HIF-1 signalling pathway	0.0001588	67	12
Spliceosome	0.00022485	78	13
p53 signalling pathway	0.00040416	45	13
Fatty acid biosynthesis	0.00057782	6	7
mRNA surveillance pathway	0.00057782	59	12
Focal adhesion	0.00063506	117	13
Glycosaminoglycan biosynthesis - keratan sulphate	0.00064327	10	7
Thyroid hormone signalling pathway	0.00133893	72	14
Central carbon metabolism in cancer	0.00179693	41	12
Acute myeloid leukaemia	0.00235642	36	12
Insulin signalling pathway	0.00298371	81	13
Gap junction	0.00328537	52	13
Small cell lung cancer	0.00387188	52	13
HTLV-I infection	0.00435804	138	13
Wnt signalling pathway	0.00473261	78	13
Sphingolipid signalling pathway	0.00483015	66	12
Melanoma	0.00659562	41	12
Glycosaminoglycan biosynthesis - chondroitin sulphate/ dermatan sulphate	0.00661598	11	7
Lysine degradation	0.00800587	25	13
ECM-receptor interaction	0.0080235	39	11
Pathogenic Escherichia coli infection	0.00894165	35	12
TNF signalling pathway	0.00979755	63	14
Vibrio cholerae infection	0.01295563	34	12
RNA transport	0.01320152	90	13
Progesterone-mediated oocyte maturation	0.01778249	51	13
Prolactin signalling pathway	0.02360585	42	13
Toxoplasmosis	0.02606497	65	12
Salmonella infection	0.03123455	48	13
Adrenergic signalling in cardiomyocytes	0.03993007	68	13
GnRH signalling pathway	0.03993007	51	13
Epstein-Barr virus infection	0.047274	105	13
Notch signalling pathway	0.04904792	29	11
Fc gamma R-mediated phagocytosis	0.04904792	50	12
MAPK signalling pathway	0.04904792	127	13
Regulation of actin cytoskeleton	0.04904792	106	13

Targeted analysis of the genes targeted by the miRNAs (Appendix) showed that none of the 13 miRNAs target *CYP7A1* or *SFRP5*. The miR-194-5p (higher expressed in HP), 27a-3p and 23a-3p (both lower expressed in HP) have SATB1, which was found to be expressed in higher levels in the HP animals and has been associated with human CRC (Al-Sohaily et al, 2014; Brocato & Costa, 2015; Lv et al, 2016; Mir et al, 2016; Zhang et al, 2014c), among their targets. MiR-192-5p and 182-5p, both highly expressed in HP, target *APC*. MiR-92b-3p, lower expressed in HP has *KRAS* among their targets and the miR-23a-3p, 214-3p (both lower expressed in HP) and 582-3p (higher expressed in HP) target PTEN. The miR -27a-3p, 214-3p and let-7d-5p (lower expressed in HP) and miR-182-5p (higher expressed in HP) target TP53. As miRNAs can influence protein expression also by reducing translation without reducing the transcript abundance significantly, these results suggest early modulation of the crucial 5 pathways of CRC in the APC^{1311} animals.

Taken together, the analysis of differentially expressed miRNAs between HP and LP animals revealed that miR-215 and 194b were significantly higher expressed in HP animals in both sequencing and PCR-based methods and miR-215 qualified as potential modifier of polyposis severity in pigs. In silico miRNA target analysis of all miRNAs significantly differentially expressed according to p-value revealed a multitude of targets, among which CYP7A1 and SFRP5 could not be identified, however SATB1, APC, KRAS, PTEN and TP53 were targets. KEGG pathway analysis showed, that the targeted mRNAs were playing a role in many cancer-associated pathways.

3.2 Analysis of genes mediating tumour progression in the porcine model for colorectal cancer

Regular endoscopy and molecular analysis of the *APC*^{1311} pigs revealed that the porcine model recapitulates key aspects of human FAP and CRC including adenomatous polyps in the colorectum with low-grade intraepithelial neoplasia (LG-IEN), high-grade intraepithelial neoplasia (HG-IEN), loss of *APC* heterozygosity, β-catenin accumulation, upregulation of c-MYC, MAPK pathway activation and progression to carcinoma *in situ*.

The search for novel unknown drivers of human CRC to fully understand disease pathology is a continuous focus of CRC research (Cancer Genome Atlas, 2012). Repeated analysis of one and the same polyp in humans is not possible and the analysis of early polyps is also difficult as the most part are diagnosed late in disease progression (US national cancer institute) (Guinney et al, 2015), therefore late stage adenomas or even carcinomas are mostly analysed in humans. The FAP pigs provide the opportunity to analyse early adenomas LG-IEN and to follow and analyse molecular changes from LG-IEN to HG-IEN over time *in vivo*. Thus the model can not only help investigate the differences between adenomas and normal mucosa, but also between LG-IEN and HG-IEN.

Frequent mutations driving human CRC progression have been identified (Fearon & Vogelstein, 1990), but CRC is very complex and very few cases of CRC carry all key mutations collectively (Guinney et al, 2015). Other mechanisms such as epigenetic modifications and mi-RNA dysregulation have been found to have similar if not equal CRC-promoting power (1.2.1, 1.2.2, 1.2.3, 1.2.4). Therefore, processes driving the progression from LG-IEN to HG-IEN were analysed in LG-IEN and HG-IEN adenomas of the colorectum on mRNA and miRNA level to detect gene and miRNA expression changes caused by genomic mutations but also by epigenetic mechanisms and miRNA dysregulation. The sequencing data was used to perform gene and miRNA expression analysis, the identification of SNPs differentially distributed and expressed and gene set enrichment analysis between LG-IEN and HG-IEN. The analysis was aimed to show molecular replication of human FAP and CRC but also to identify novel drivers through unique experimental setup that is not possible in humans.

3.2.1 Analysis of tumour progression on mRNA level

To compare porcine CRC pathology to human and to identify novel drivers of CRC, porcine tumour progression was analysed by comparing RNA sequencing results from 5 HG-IEN with 5 LG-IEN adenomas of 1 cm diameter of *APC*^{1311} animals. Additional to bulk samples, laser microdissected LG-

IEN and HG-IEN were sequenced and compared. These following results have been published in Scientific Reports (Flisikowska et al, 2017).

3.2.1.1 Differential expression analysis

Differential expression analysis resulted in 52 genes that were significantly (adjusted p-Value < 0.05) differentially expressed and showed a clear distinction between HG and LG-IEN (Figure 15). The top differentially expressed gene was AHNAK, which was lower expressed in HG-IEN. AHNAK is a known tumour suppressor that negatively regulates cell growth via the TGFβ signalling pathway (Lee et al, 2014). Among the differentially expressed genes, where most showed a reduced expression in HG-IEN, were genes involved in metabolic processes (*MAPK4*, *S100A9*), intracellular transport (*SLC46A1*) and in immune response (*IL7*, *CD40*). Among highly expressed genes were genes associated with stress response (*HSPA1L*), WNT (*WISP1*) and TNF (*SLC12A6*) signalling. Selected genes, especially immune related genes (Table 42) were validated using RT-qPCR. LG samples served as calibrator group, therefore positive values signify higher regulation in HG samples.

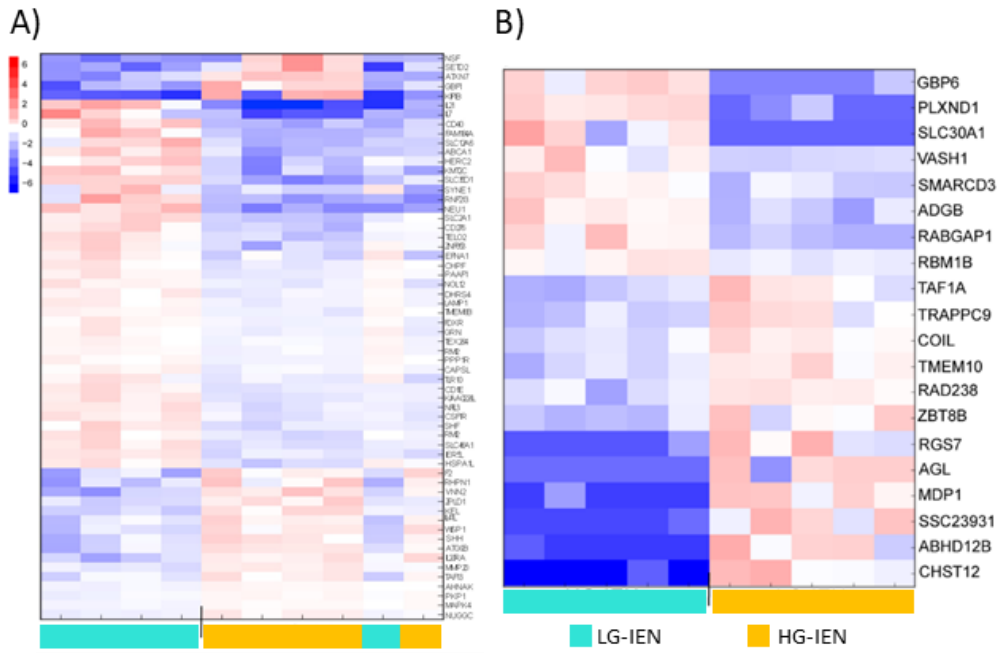


Figure 15 A) cluster analysis of the expression of the top 52 differentially expressed genes between HG-IEN and LG-IEN. B) cluster analysis of the expression of the top 20 differentially expressed genes between laser microdissected HG-IEN and LG-IEN crypts.
Blue signifies low and red high expression.

Table 42 RT-qPCR validation results of genes differentially expressed between whole HG and LG-IEN.

Gene symbol	Gene description	RNA seq Log2Fold-Change HG	P-value	RT-qPCR Log2Fold-Change HG	P-value
IL7	Interleukin 7	-1.09	$6.88x10^{-6}$	-5.18	$4.02x10^{-7}$
EBF1	Early B-cell factor 1	-1.69	$7.33x10^{-10}$	-4.03	0.008
CD40-1	CD40 molecule isoform 1	-1.10	$2.09x10^{-12}$	-1.96	0.013
CD40-2	CD40 molecule isoform 2	-1.22	$2.88x10^{-5}$	-2.04	0.085
CD40-LG	CD40 ligand	-1.28	$4.56x10^{-6}$	-2.11	0.009
TRAF5	TNF receptor associated factor 5	-1.16	$1.20x10^{-9}$	-2.21	0.007
IL21	Interleukin 21	-2.30	$2.06x10^{-17}$	-8.97	0.006
CD101	CD101 molecule	-0.95	0.0003	-3.08	0.044
S100A8	S100 calcium-binding protein A8	2.79	$3.16x10^{-12}$	3.79	0.021
S100A9	S100 calcium-binding protein A9	-2.34	$5.27x10^{-9}$	-2.66	0.048
IL20RA	Interleukin 20 receptor subunit alpha	2.18	$6.08x10^{-11}$	3.40	$6.7x10^{-6}$

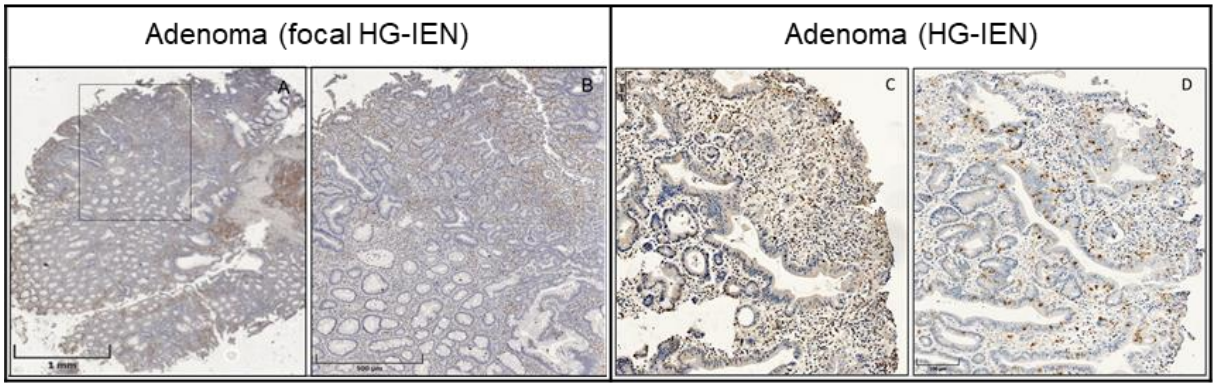


Figure 16 Immunohistochemical staining of HG-IEN.
A) and B) show brown staining of CD3+ infiltrating T-cells (B is a magnification of A). C) displays brown CD4+ and D) CD8+ infiltrating T-cells.

To specifically analyse aberrant epithelial tissue and eliminate signals from surrounding stroma, sequencing and differential expression analysis of laser microdissected LG and HG-IEN crypts was performed. Differential expression analysis of laser microdissected samples sequenced showed no differential expression of genes such as *IL7*, *S100A8*, and *S100A9*, assigning their expression to other cell types contained in the stroma, such as infiltrating immune cells (e.g. T-cells), observed in the polyps analysed (Figure 16). Differential expression visualisation showed clear distinction between LG and HG-IEN (Figure 15).

High expression of the selected genes *PLXND1*, *SLC30A1*, *GBP6*, *VASH1*, and *SMARCD3* in HG-IEN was confirmed using RT-qPCR (Table 43) suggesting novel drivers of CRC. *PLXND1* has been found highly

expressed in a number of human tumours (Roodink et al, 2009) and associated with epithelial mesenchymal transition and invasiveness and metastasis (Casazza et al, 2010; Tseng et al, 2011). The other genes were associated with interferon-γ signalling (*GBP6*) oxidative stress and anti-inflammatory activity (*SLC30A1*), p53 regulation and cell proliferation (*VASH1*) and WNT pathway regulation (*SMARCD3*).

Table 43 RT-qPCR validation results of genes differentially expressed between laser microdissected HG and LG-IEN.

Gene symbol	Gene description	RNA seq Log2Fold-Change HG	P-value	RT-qPCR Log2Fold-Change HG	P-value
PLXND1	plexin D1	2.26	$5.40x10^{-8}$	4.23	$1.40x10^{-3}$
SLC30A1	solute carrier family 30 member 1	2.21	$9.74x10^{-8}$	3.95	0.0009
GBP6	guanylate binding protein family member 6	1.91	$3.07x10^{-6}$	5.40	$2.10x10^{-3}$
VASH1	vasohibin 1	1.60	$7.65x10^{-5}$	3.10	0.0091
SMARCD3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3	1.32	$6.79x10^{-4}$	2.88	$1.99x10^{-3}$

Gene set enrichment analysis

Gene set enrichment analysis of porcine whole biopsy HG-IEN versus LG-IEN and laser microdissected HG-IEN versus LG-IEN were compared to human microsatellite stable (MSS) T1 polyps versus normal mucosa (data from TCGA database) to compare the porcine CRC carcinogenesis in a global molecular approach (Figure 17Figure 15). Of 26 significantly differentially enriched gene sets in human and 25 in pig, 19 sets were commonly enriched in both species. 13 genes sets were similarly significantly differentially enriched in the laser microdissected samples compared to the human and porcine bulk samples including HG-IEN enriched gene sets essential for CRC progression as MYC targets and cell cycle related E2F targets and G2M checkpoint components (Figure 17). This though broad comparison highlights the evident similarity between porcine and human CRC progression.

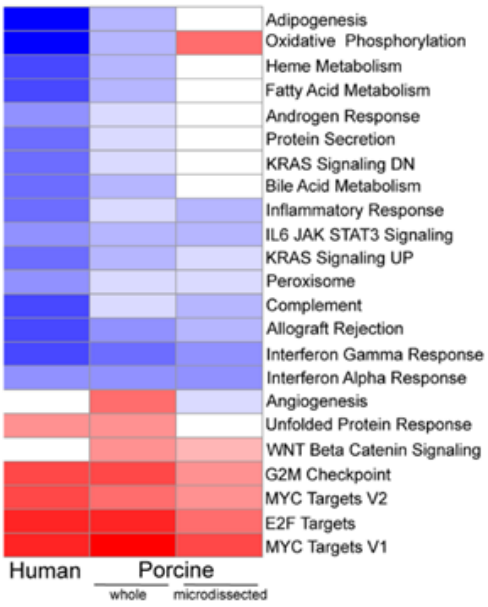


Figure 17 Gene set enrichment analysis of the pathways up or downregulated in porcine whole and microdissected polyps compared to human T1 microsatellite stable polyps.

3.2.1.2 Allele-specific expression analysis

Molecular changes in parts of the cancer over time have been observed in CRC by allele-specific expression (ASE) analysis (Tuupanen et al, 2008). *In silico* analysis of raw RNA sequencing data showed increased expression of the *APC*^{1311} allele in HG-IEN compare to LG_IEN (Figure 18).

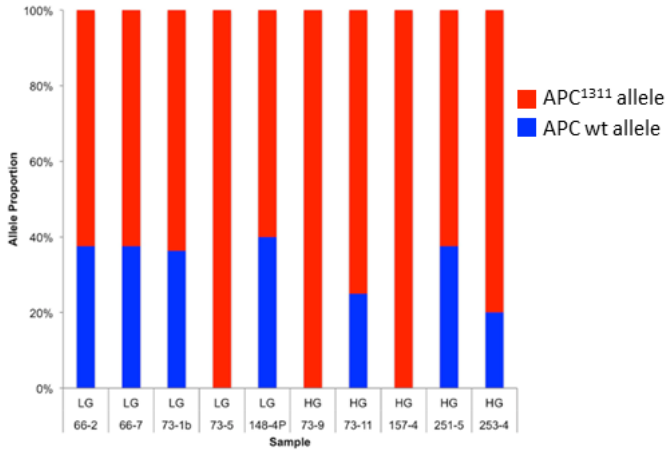


Figure 18 Allele-specific expression of *APC*, analysed using raw sequencing data

ASE analysis using the RNA sequencing data from the bulk samples identified 48 000 SNPs that showed significant allelic imbalance in at least one sample. SNPs in known CRC related genes with epithelial function *MMP9*, *CEACAM7*, *LMAN2* and an immune related gene *SLA2* were validated in bulk and laser

microdissected samples using pyrosequencing (Figure 19). The allelic imbalance that were well visible in laser microdissected samples were masked in bulk samples.

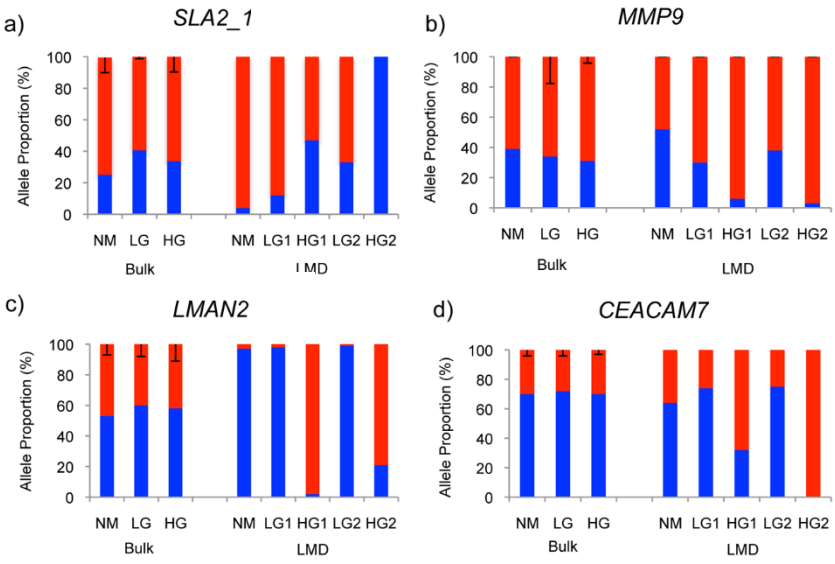


Figure 19 Allele-specific expression analysis of *SLA2_1.0*, *MMP9*, *LMAN2* and *CEACAM7* using pyrosequencing.

3.2.2 Analysis of tumour progression on miRNA level

MiRNA from 5 LG-IEN and 5 HG-IEN (same samples as used for mRNA analysis) were sequenced and compared. These results have been published in Oncotarget (Stachowiak et al, 2017). 44 differentially expressed miRNAs with a p-value below 0.05 were identified. Six of those had an adjusted p-value below 0.05, rendering them significant even in the face of multiple comparison testing. Ten selected miRNAs including the top differentially expressed miR-126-3p between HG-IEN and normal mucosa, were selected for RT-qPCR validation (Table 44). HG was used as calibrator in this comparison, therefore, negative Log2FoldChange values signify increased expression in HG. All RT-qPCRs showed the same trend as the sequencing results, but the increased expression of let-7e, miR-98, 146a-5p, 146b, 183 an 196a and the reduced expression of miR-126-3p in HG-IEN were significant (p-value < 0.05) or even highly significant (p-value < 0.01) (Figure 20). Further all 7 miRNAs had been found to be associated with human CRC (Bandres et al, 2006; Chen et al, 2016; Ge et al, 2014; Lu et al, 2017; Mosakhani et al, 2012; Motoyama et al, 2009; Schimanski et al, 2009; Zhu et al, 2017a; Zhu et al, 2017b).

Table 44 Selected differentially expressed miRNAs between LG and HG-IEN and HG-IEN and normal mucosa.

MiRNA / IsomiR	Log2Fold-Change LG	P-value	Adjusted p-value
ssc-miR-196a	-1.72	0.0000	0.0012
ssc-miR-98	-0.95	0.0001	0.0118
ssc-miR-146a-5p	-1.67	0.0003	0.0224
ssc-miR-146b|isomiR|20_42|	-1.43	0.0005	0.0265
ssc-let-7e	-1.25	0.0014	0.0482
ssc-miR-181b|isomiR|12_33|	-1.34	0.0020	0.0543
ssc-miR-191|isomiR|9_30|	-0.95	0.0035	0.0764
ssc-miR-155-5p|isomiR|10_33|	-1.19	0.0050	0.0945
ssc-miR-183|isomiR|6_27|	-1.11	0.0052	0.0945
MiRNA / IsomiR	**Log2FC HG/N**	**p-value**	**Adjusted p-value**
ssc-miR-126-3p|isomiR|45_66|	-2.45	0.0000	0.0137

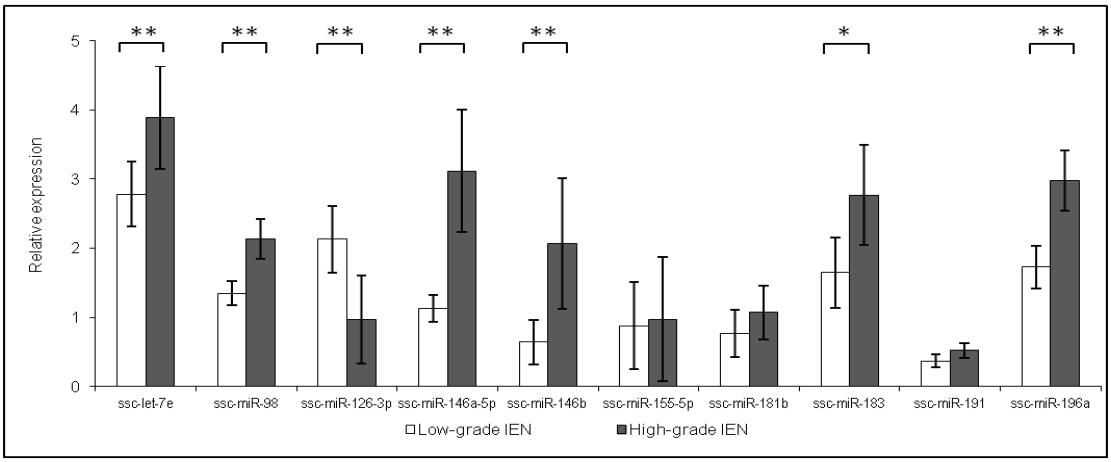


Figure 20 RT-qPCR validation of the differentially expressed miRNAs selected.

In silico miRNA target analysis

In silico target analysis revealed a large number of cancer related pathways influenced by the differentially expressed miRNAs including miRNAs in cancer, proteoglycans in cancer, p53 signalling pathway, viral carcinogenesis and colorectal cancer (Table 45).

Table 45 *In silico* miRNA target analysis of differentially expressed miRNAs between HG and LG-IEN.

KEGG pathway	P-value	Number of genes targeted	Number of miRNAs involved in pathway
MicroRNAs in cancer	$9.53x10^{-74}$	131	23
Proteoglycans in cancer	$1.08x10^{-14}$	146	23
Hepatitis B	$5.68x10^{-12}$	105	22
Protein processing in endoplasmic reticulum	$1.21x10^{-09}$	124	23
Renal cell carcinoma	$1.25x10^{-08}$	56	23
Cell cycle	$1.44x10^{-08}$	97	23
Lysine degradation	$2.75x10^{-08}$	37	23
Ubiquitin mediated proteolysis	$2.75x10^{-08}$	103	23
Adherens junction	$3.67x10^{-08}$	60	23
Hippo signalling pathway	$1.69x10^{-07}$	101	23
Prion diseases	$3.21x10^{-07}$	24	22
Pathways in cancer	$6.92x10^{-07}$	255	23
Oestrogen signalling pathway	$1.33x10^{-06}$	69	23
p53 signalling pathway	$1.88x10^{-06}$	56	23
Viral carcinogenesis	$4.50x10^{-06}$	146	23
Prostate cancer	$7.91x10^{-06}$	68	23
Transcriptional misregulation in cancer	$1.68x10^{-05}$	117	23
Pancreatic cancer	$1.88x10^{-05}$	52	23
Glioma	$4.22x10^{-05}$	47	23
Spliceosome	$4.97x10^{-05}$	94	23
Glycosaminoglycan biosynthesis - keratan sulphate	$8.15x10^{-05}$	11	18
Signalling pathways regulating pluripotency of stem cells	$8.15x10^{-05}$	95	23
Chronic myeloid leukaemia	$8.73x10^{-05}$	56	22
TGF-beta signalling pathway	0.000114499	55	22
Shigellosis	0.000114499	48	23
Bladder cancer	0.000114499	32	23
Sphingolipid signalling pathway	0.00011831	82	23
Oocyte meiosis	0.000125888	76	23
Acute myeloid leukaemia	0.00017571	42	23
Endocytosis	0.000195403	137	23
Thyroid hormone signalling pathway	0.000385396	84	23
Central carbon metabolism in cancer	0.000385396	46	23
Regulation of actin cytoskeleton	0.000455559	135	23
Fatty acid metabolism	0.000456395	29	20
FoxO signalling pathway	0.000456395	91	23
Focal adhesion	0.000587346	135	23
Colorectal cancer	0.000635242	45	23
Non-small cell lung cancer	0.000689357	41	22

RNA transport	0.000905502	110	23
TNF signalling pathway	0.000905502	77	23
ErbB signalling pathway	0.001752074	58	23
Bacterial invasion of epithelial cells	0.002263929	50	23
HIF-1 signalling pathway	0.002639591	76	23
mTOR signalling pathway	0.002897181	45	23
Steroid biosynthesis	0.003093938	13	15
Neurotrophin signalling pathway	0.003536926	79	23
RNA degradation	0.004208333	53	23
Thyroid cancer	0.006358412	21	22
AMPK signalling pathway	0.008043309	83	23
N-Glycan biosynthesis	0.008687542	32	20
Small cell lung cancer	0.009531305	58	23
NF-kappa B signalling pathway	0.009585347	56	23
Fc gamma R-mediated phagocytosis	0.010505971	59	23
Endometrial cancer	0.010954543	36	22
Axon guidance	0.012636065	75	23
Wnt signalling pathway	0.014365498	86	23
Chagas disease (American trypanosomiasis)	0.016876026	66	22
Huntington's disease	0.016876026	113	23
Legionellosis	0.016928086	38	22
Ribosome	0.021827887	86	23
HTLV-I infection	0.021827887	168	23
Salmonella infection	0.022887739	56	23
Notch signalling pathway	0.024647157	34	23
Fatty acid biosynthesis	0.025038322	6	18
Melanoma	0.030733908	45	23
Insulin signalling pathway	0.03093753	88	23
MAPK signalling pathway	0.032704866	151	23
Gap junction	0.040246432	55	23
PI3K-Akt signalling pathway	0.040246432	192	23
Progesterone-mediated oocyte maturation	0.040999811	57	23
Prolactin signalling pathway	0.049040505	47	22

The characterisation of the *APC*^{1311} animals on mRNA and miRNA level revealed that only few genes and miRNAs are significantly differentially expressed between HP and LP. Higher gene expression of CYP7A1, the SNP in the OAS1 gene (chromosome 14, position 38856577 bp C/T) and the miR-215 in HP animals were validated with methods other than mRNA and miRNA sequencing and associated with a more severe polyposis (HP). Their potential function as modifier genes requires further confirmation in a larger set of samples. MiR-215 could already be validated in animals apart from the sequenced samples. The mechanism of differential CYP7A1 expression was analysed on *cis*-regulatory level and revealed reduced CpG methylation in the potential promoter region of HP animals where TFs STAT3, GLIS1, CEBPB and GABPA bind. MRNA targets of differentially expressed miRNAs were determined and clustered and showed involvement in many cancer-associated pathways.

Further analysis of tumour progression on mRNA level between low grade (LG) and high grade (HG) intraepithelial neoplasia (IEN) showed a clear distinction between the two on expression in both bulk and laser microdissected samples. Genes involved in metabolic processes and immune response were lower expressed in HG-IEN while genes associated with stress response were higher expressed. Analysis of laser microdissected IENs eliminated these signals originating from stroma and revealed high *PLXD1* and *GBP6* expression in HG-IEN. So both the stroma and the epithelium itself change during tumour progression. Further, SNPs in known CRC related genes with epithelial function *MMP9*, *CEACAM7*, *LMAN2* and one immune related gene *SLA2* were detected. Gene set enrichment analysis of differential gene expression data showed cancer-associated gene sets enriched and further presented homology to human CRC data. MiRNAs analysis in bulk samples identified miRNAs let-7e, miR-146a-5p, 146b, 183, 196a higher expressed in HG-IEN that have not only been reported to exhibit tumour promoting functions, but that also target genes involved in cancer relevant pathways.

3.3 Optimisation of the CRC model

The *APC*^{1311} pig model for FAP and CRC has replicated key aspects of human FAP such as the variation in polyposis severity and hallmarks of human CRC development and progression to carcinoma *in situ*. However, no progression to invasive colorectal carcinoma was so far observed in more than 100 animals of 4 generations of *APC*^{1311} pigs up to age of 2-3 years. This suggests, that just like in humans, where 10-15 % of all adenomas progress to carcinoma over decades (Fearon, 2011), porcine CRC progression takes equivalent time. With the premise that development of an invasive carcinoma is a matter of time, it was decided to accelerate the CRC development by causing frequently diagnosed oncogenic mutations (Guinney et al, 2015) in polyps of the pigs *in vivo* via genome editing using CRISPR/Cas9. It was planned to deliver the guide RNAs to target tumour suppressor genes and oncogenes together with single-stranded oligodeoxynucleotides (ssODNs) for the introduction of oncogenic mutations to the polyps by *in vivo* electroporation or adeno-associated viral vectors. The endonuclease Cas9 is very large. Therefore, ubiquitous *Cas9* endonuclease isolated from *Streptococcus pyogenes* (*SpCas9*) was to be introduced into the *ROSA26* locus of porcine kidney fibroblasts (pKFs) carrying heterozygous *APC*^{1311} mutation followed by generation of piglets via nuclear transfer.

3.3.1 Generating targeting vectors for *Cas9* placement into the porcine *ROSA26* locus

Targeted introduction of numerous transgenes via homologous recombination (HR) in mice have been directed to the murine *Rosa26* locus because transgene introduction into intron 1 of *Rosa26* resulted in viable fertile animals that expressed the transgene stably without silencing (Zambrowicz et al, 1997). Therefore, a targeting vector allowing *SpCas9* placement into the porcine homologue *ROSA26* locus (Kong et al, 2014; Li et al, 2014) via HR was performed (Figure 21).

Promoter trap strategy was applied (Friedel et al, 2005) to increase the efficiency of HR-mediated gene targeting. The *APC*^{1311} animals carried a targeting cassette with a blasticidin S resistance (*bsr*) gene in the endogenous *APC* gene. Therefore, the *bsr* of the vector GEMT-Rosa26-BS-Cas9 (previously generated by Dr. Judy Ng) was replaced with a neomycin resistance (*neo*) gene from the pBs-LSL-Neo plasmid (Figure 21). This resulted in the targeting vector GEMT-Rosa26-Neo-Cas9 with a splice acceptor (SA), a promoterless *neo* gene and a CAG-driven *SpCas9* gene between the homologous arms targeting *ROSA26*.

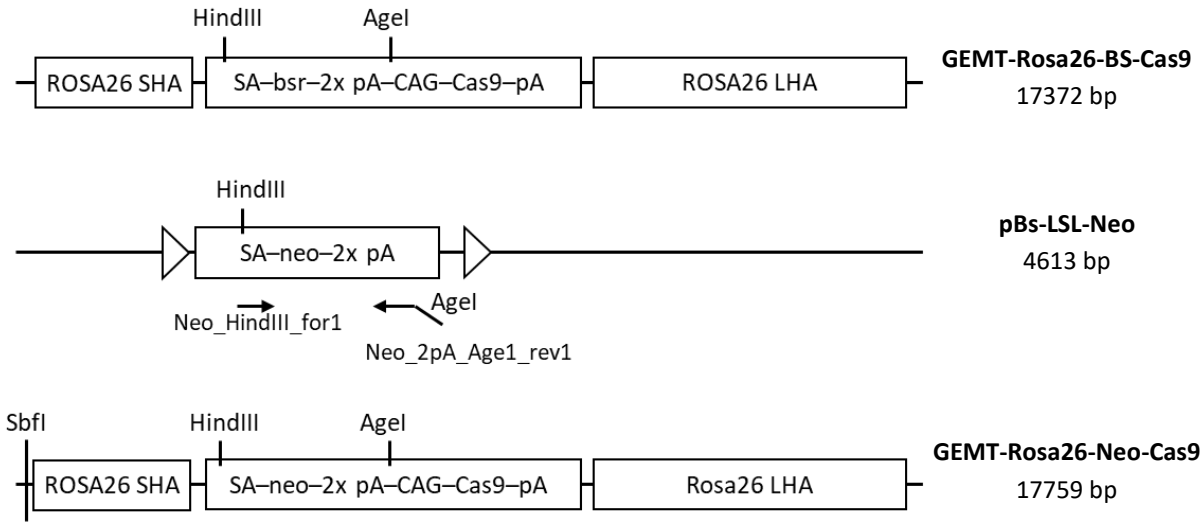


Figure 21 Cloning strategy for the targeting vector GEMT-Rosa26-Neo-Cas9.
bsr, blasticidin S resistance gene; CAG, Chicken beta-actin promoter and cytomegalovirus enhancer element; LHA, long homology arm; neo, neomycin resistance gene; pA, polyadenylation signal; SA, splice acceptor; SHA, short homology arm

As in the future both the Cre-loxP system together with the CRISPR/Cas9 system shall be used together to model human cancers, the coupling of the *SpCas9* gene and a Cre reporter cassette in the *ROSA26* locus is an essential preparation. The Cre reporter system allows visualisation of the location of Cre recombination *in vivo* and represents a powerful tool to monitor Cre specificity when working with conditional and tissue-specific Cre inducible oncogenic mutations (Li et al, 2014).

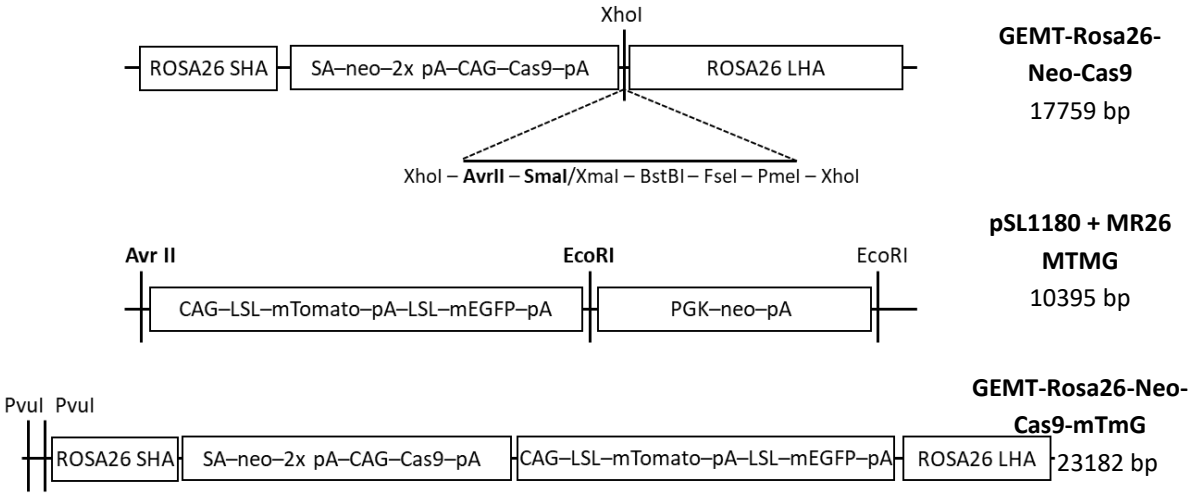


Figure 22 Cloning strategy for targeting vector GEMT-Rosa26-Neo-Cas9-mTmG
CAG, Chicken beta-actin promoter and cytomegalovirus enhancer element; LHA, long homology arm; LSL, loxP-Stop-loxP cassette; neo, neomycin resistance gene; mEGFP, membrane-targeted enhanced green fluorescent protein gene; mTomato, membrane-targeted tdTomato red fluorescent protein gene; pA, polyadenylation signal; PGK, phosphoglycerate kinase promoter; SA, splice acceptor; SHA, short homology arm

Therefore, the GEMT-Rosa26-Neo-Cas9 was expanded by introduction of a Cre reporter cassette mTmG consisting of a CAG promoter and the two fluorescent protein genes membrane-targeted tdTomato (*mTomato*) flanked by two loxP sites and membrane-targeted EGFP (*mEGFP*) (LSL-mTomato-pA-LSL-mEGPF-pA) (Figure 22). Both vectors were linearised to increase the rate of homologous recombination (Kucherlapati et al, 1984) and purified for nucleofection of pKFs of *APC*^{1311} pig 73.

3.3.2 Generation and analysis of *Cas9*-targeted clones

Clones were generated by transfection, selection and expansion of primary pKFs. The single-cell clones were screened for correct targeting of the *ROSA26* allele via PCR amplification across the 5' junction of the vector and the target site using the primer Rosa26 I1 F2 and Rosa26 Loc2R and sequence analysis (Appendix) of the amplified products (Figure 23). Targeting PCR identified 9 targeted clones (Figure 24). A PCR to detect the endogenous, unmodified *ROSA26* allele with the primer Rosa26 I1 F2 and Rosa26 I1 R3 showed that only 1 allele of *ROSA26* was targeted (Figure 24).

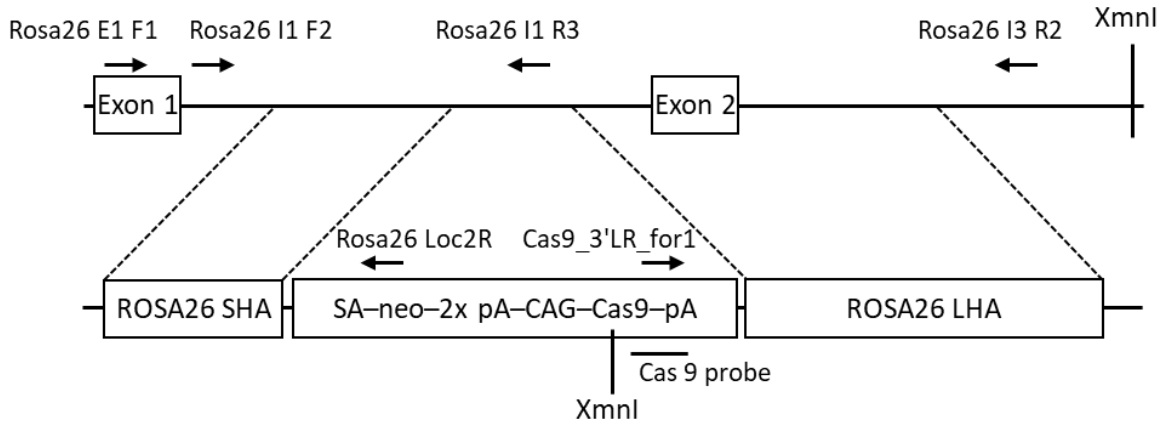


Figure 23 Targeting of the *Rosa26* locus with the GEMT-Rosa26-Neo-Cas9 targeting vector
CAG, Chicken beta-actin promoter and cytomegalovirus enhancer element; LHA, long homology arm; neo, neomycin resistance gene; pA, polyadenylation signal; SA, splice acceptor; SHA, short homology arm.

The clones, 4, 5, 9, 19, 22,35 and 45 were used as donors for nuclear transfer without additional analysis as they had ceased to proliferate. Clones 89, 92 and 93 could be expanded for further analysis and correct *ROAS26* targeting was confirmed by PCR amplification across the 3'junction of the vector and target locus (Figure 25) and sequence analysis (Appendix) of the amplified products.

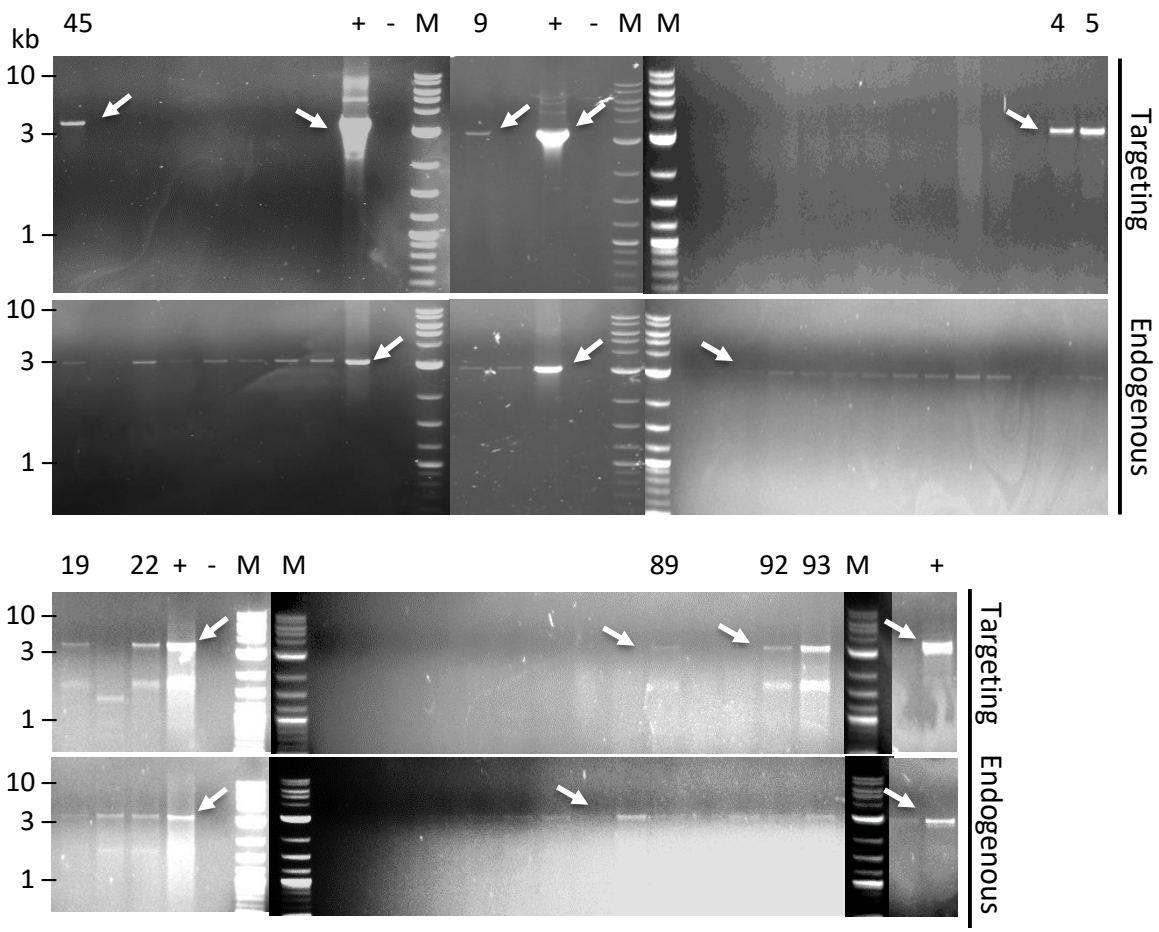


Figure 24 Gel electrophoresis of the 5' screening PCR of *Cas9*-targeted clones.
Targeting PCR of the clones was performed using primer Rosa26 I1 F2 binding the *ROSA26* locus outside the homologous arm and Rosa26 Loc2R that binds the neo of the targeting vector to generate a 3313 bp. Endogenous PCR of 3105 bp was amplified using the primer Rosa26 I1 F2 and Rosa26 I1 R3 that binds only the untargeted wildtype allele of ROSA26. +, positive control; -, water control; M, marker.

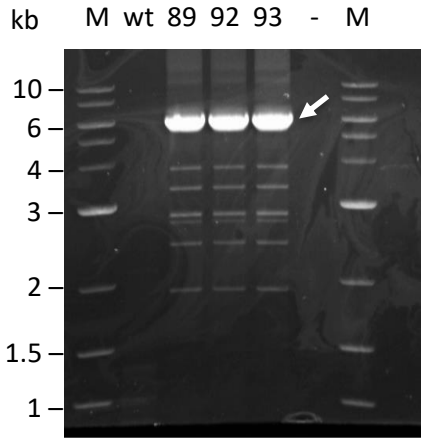


Figure 25 Gel picture of the 3'LR PCR of clones 89, 92 and 93.
The primer Cas9_3'LR_for1 and Rosa26 I3 R2 (Figure 23) were used to generate a product of 6118 bp. -, water control; M, marker; wt, wildtype control.

3.3.2.1 Southern blot analysis

Southern blot analysis was performed to validate correct targeting of the *ROSA26* locus and to exclude additional random integration. Genomic DNA of clones 89, 92, 93, 13 (a cell clone generated by Beate Rieblinger and Nina Simm where *SpCas9* was introduced into the *ROSA26* locus of wildtype pKFs) and wildtype cells was isolated and digested with XmnI.

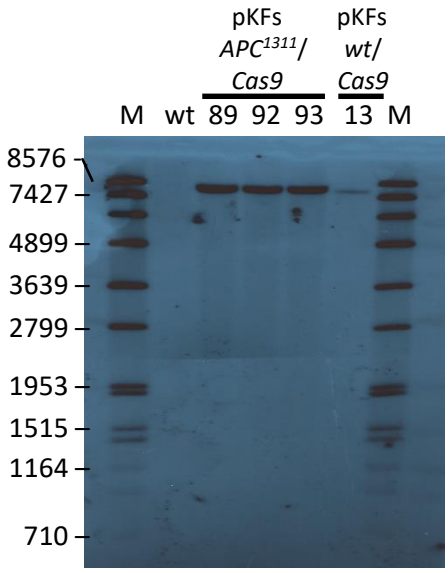


Figure 26 Southern Blot of the clones 89, 92, 93 and 13.
M, marker; wt, wildtype control.

The clones showed the expected bands of 7777 bp (clones 89, 92 and 93) and 7698 bp (clone 13) (Figure 26) using the Cas9 probe (Figure 23). No random bands were visible, confirming that the integration was not only targeted but also unique. However, although the DNA concentration of all clones was measured fluorometrically using the Qubit and based on this, 10 µg DNA was used for Southern blot, clearly clone 13 showed a significantly weaker band, indicative of less DNA.

3.3.2.2 Expression analysis

Correct splicing from exon 1 of the *ROSA26* locus into *neo* was detected by reverse transcription (RT) PCR of cDNA of the clones 89, 92 and 93 (Figure 23). The expected bands of 877 bp and 990 bp were well visible in all three clones and the positive control (Figure 27). Additionally, the correct sequence of the PCR products was confirmed by sequencing analysis (Appendix).

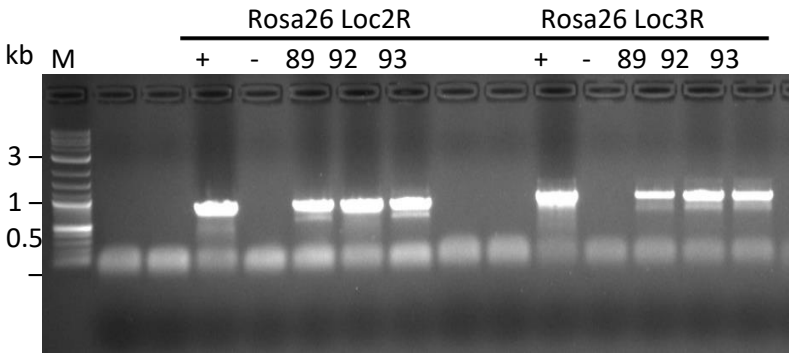


Figure 27 Gel electrophoresis of RT PCR of the clones 89, 92 and 93.
The PCR from exon 1 of the Rosa26 locus, using the primer Rosa26 E1 F1 to the neomycin cassette, using Rosa26 Loc2R and Rosa26 Loc3R resulted in 877bp and 990 bp respectively. +, positive control; -, water control; kb; kilo bases; M, marker.

SpCas9 expression was quantified by RT-qPCR. For this analysis unlike traditional RT-qPCR the primers were not separated by an intron, therefore RNA was also tested for DNA contamination. The RNA showed no PCR product. The results showed that *SpCas9* was expressed in the clones.

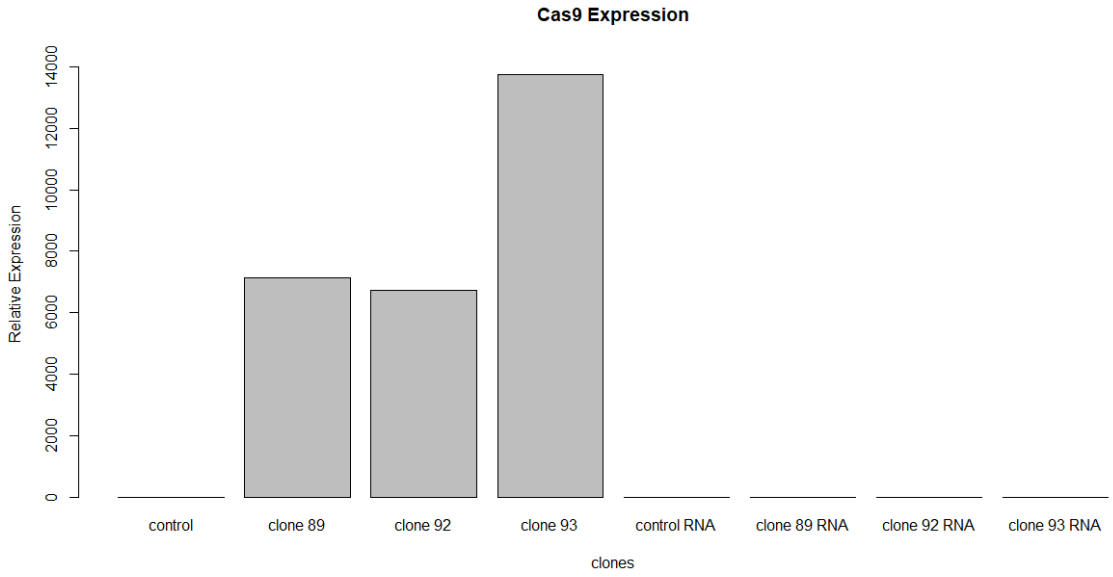


Figure 28 *Cas9* expression analysis of the clones 89, 92 and 93 using RT-qPCR results.

3.3.2.3 Protein Analysis

3.3.2.3.1 Western blot analysis

The presence of mRNA does not necessarily indicate the presence of functional protein. Therefore, protein was isolated from the clones 89, 92 and 93 to perform Western blot analysis. The SpCas9 protein of 160 kDa could be detected for all three clones. The negative control, protein from untransfected pKFs from *APC*^{1311} pig 73 showed only the loading control, GAPDH at 37 kDa.

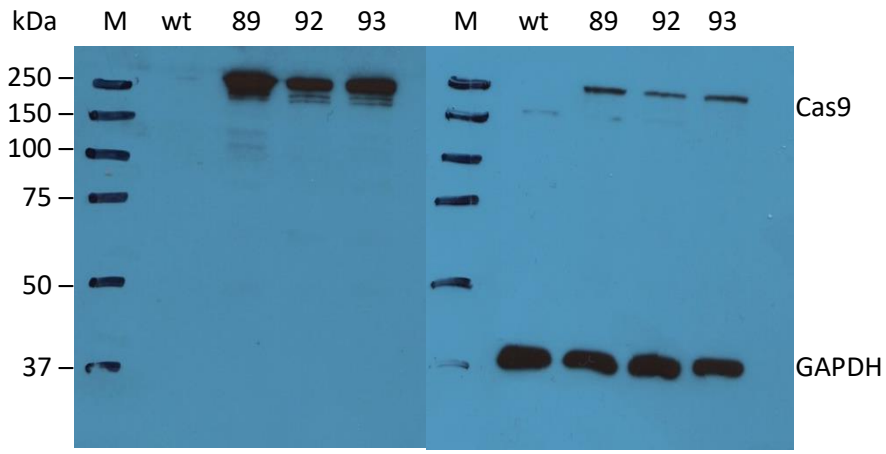


Figure 29 Western blot analysis of the clones 89, 92 and 93 visualising Cas9 and GAPDH.
kDa, kilo Dalton; M, marker; wt, wildtype control.

3.3.2.3.2 Functional Assay

SpCas9 is a nuclease, therefore it is essential to validate its activity. A reporter assay that can verify Cas9 functionality was established and indels were detected using open source online tool TIDE (https://tide.deskgen.com/).

Establishing gRNAs

First, gRNAs were generated, that target tumour suppressor genes that play a crucial role in human CRC development: *TP53*, as the "guardian of the genome", *APC* to induce loss of heterozygosity, *PTEN* as the controller of the PI3K pathway and *DCC*. These were inserted into Cas9-gRNA vector and transfected into pKFs. The cells were not selected. Therefore, gene editing efficency evaluated by TIDE analysis (https://tide.deskgen.com/) was dependent on transfection efficiency (Table 46). However, all gRNAs were able to induce cleavage and gene editing in the endogenous loci (Table 46).

Table 46 Tide analysis results of wildtype cells transfected with Cas9 and gRNA targeting TP53, PTEN, APC and DCC.

Cells	vector	Gene editing efficiency
pKF 73 *APC*^{1311}	pX330-Cas9-TP53	9.3
pKF 73 *APC*^{1311}	pX330-Cas9-Puro-PTEN	5.4
pKF 73 *APC*^{1311}	pX330-Cas9-APC	5.3
pKF 73 *APC*^{1311}	pX330-Cas9-DCC_Ex1-1	5.8

Reporter Assay

The gRNA sequence from the pX330 vector was added to the P119_pFUS_C_Check reporter plasmid carrying a dsRed sequence (rendering all transfected cells red) and a CRISPR target site including the

protospacer adjacent motif (PAM) between two homologous incomplete eGFP sequences (Figure 30). Upon transfection of cells expressing functional SpCas9 nuclease, the gRNA complexes with the Cas9 and travels to the endogenous target site but also to the one between the two homologous eGFP sequences and induces a double strand break. In case of homology directed repair (HDR), the homologous eGFP sequences will serve as templates, resulting in an intact eGFP. Succesful gene editing on the plasmid by cleavage-induced HDR is thus visible by fluorescence microscopy. Gene editing of the endogenous locus by the more frequent repair mechanism non-homologous end joining is detectable using the TIDE tool (https://tide.deskgen.com/).

The clones 89, 92 and 93 were transfected with pX330-cCheck-TP53 and both the fluorescence microscopy (Figure 31) and Tide analysis (Table 47) confirmed that all three clones generated a fully functional SpCas9 protein that was able to cleave both exogenous DNA (plasmid) and endogenous DNA and induce cellular DNA repair mechanisms.

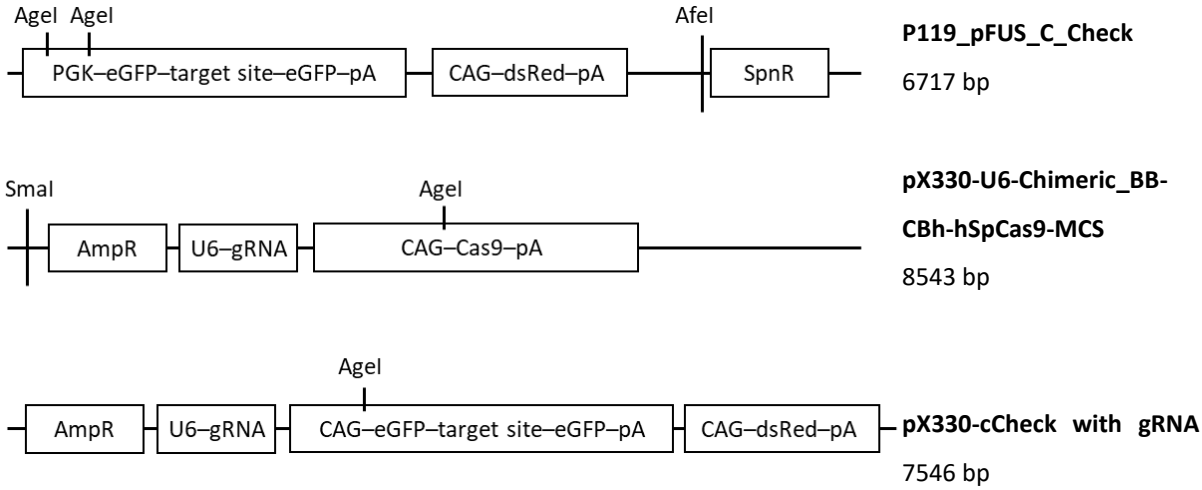


Figure 30 Cloning strategy of generating a reporter system for cleavage efficiency of *Cas9*-expressing cells. AmpR, ampicillin resistance gene; CAG, Chicken beta-actin promoter and cytomegalovirus enhancer element; dsRed, red fluorescent protein gene isolated from *Discosoma*; eGFP, enhanced green fluorescent protein gene; pA, polyadenylation signal; PGK, phosphoglycerate kinase promoter; SpnR, spectinomycine resistance gene; U6, U6 promoter.

Table 47 Tide analysis of the clones 89, 92 and 93 transfected with pX330-cCheck-TP53.

Cells	Vector	Gene editing efficiency
pKF 73 *APC*^{1311}	pX330-cCheck-TP53 ETOH	0
pKF 73 *APC*^{1311}/*Cas9* clone 89 P13	pX330-cCheck-TP53 ETOH	31.7
pKF 73 *APC*^{1311}/*Cas9* clone 92 P13	pX330-cCheck-TP53 ETOH	42.6
pKF 73 *APC*^{1311}/*Cas9* clone 93 P12	pX330-cCheck-TP53 ETOH	40.4

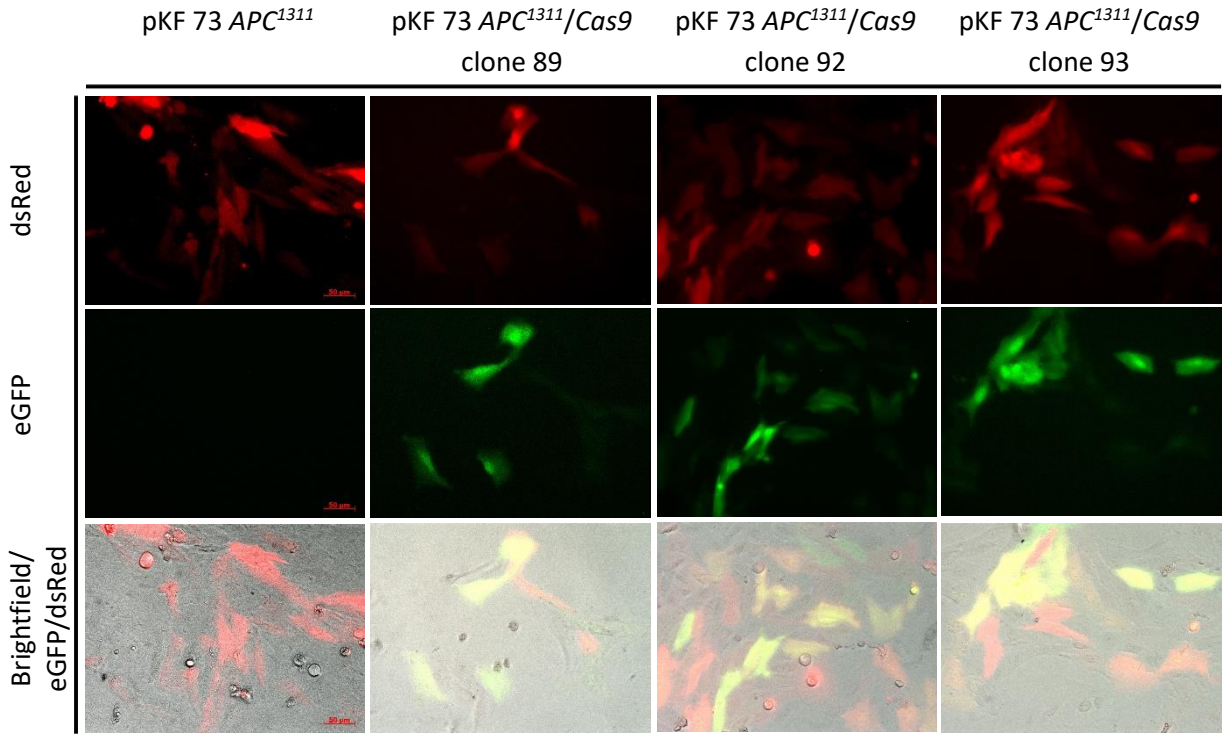


Figure 31 Fluorescence microscopy of the clones 89, 92 and 93 transfected with pX330-cCheck-TP53.

A vector carrying only the gRNA was generated, eradicating the competing exogenous target site of the cCheck plasmid (Figure 32) to efficiently compare the endogenous target cleavage efficacy in SpCas9 expressing cells to cells where both SpCas9 and gRNA are delivered.

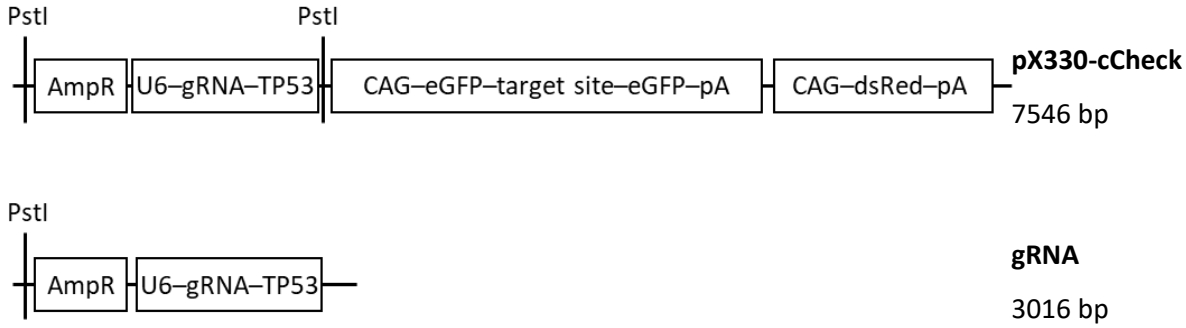


Figure 32 Cloning strategy of the generation of a vector carrying only the TP53 gRNA sequence for transfection of Cas9 expressing cells.

The clone 92 and pKF 73 *APC*^{1311}cells were transfected with pX330-cCheck-TP53 and gRNA-TP53 and the wildtype cells additionally with pX330-Cas-TP53. Tide analysis (https://tide.deskgen.com/) showed that the endogenous *TP53* was cut more efficiently when the cells expressed Cas9, than when pKF 73 *APC*^{1311}had to be transfected with both gRNA and *Cas9* (Table 48). However, as the determined gene editing efficiencies are dependent on transfection efficiency, gene editing in pKF 73 *APC*^{1311} transfected

with pX330-Cas9-TP53 was here three fold better than in the experiment inTable 46. Also, as expected the cleavage of the endogenous locus was improved, when the cells from clone 92 were transfected with the gRNA-TP53, without the competing target site. Furthermore, the cutting efficiency was influenced by the purification method of the DNA used. Plasmid DNA isolated using the plasmid DNA purification NucleoBond® Xtra Midi (MACHEREY-NAGEL GmbH & Co. KG) without any additional purification showed better cutting efficiency than those additionally purified using ethanol precipitation.

Table 48 Tide analysis of wildtype cells and cells from clone 92 transfected with vectors carrying a gRNA targeting TP53.

Cells	Vector	Gene editing efficiency [%]	Gene editing efficiency [%] EtOH
pKF 73 *APC*^{1311}	pX330-Cas9-TP53	30.1	19.5
pKF 73 *APC*^{1311}	pX330-cCheck-TP53	0	-
pKF 73 *APC*^{1311}	gRNA-TP53	0	-
pKF 73 *APC*^{1311}/*Cas9* clone 92	pX330-cCheck-TP53	62.7	45.2
pKF 73 *APC*^{1311}/*Cas9* clone 92	gRNA-TP53	73	-

3.3.2.4 Nuclear transfer

All positive clones were used in nuclear transfer. The clones 4, 5, 22, 89, 92 and 93 were frozen and pooled for the use in 20 nuclear transfers and subsequent embryo transfers of which none resulted in the birth of *APC*^{1311}/*Cas9* piglets. However, the clone 13 generated from wildtype pKFs resulted in birth of viable Cas9 pigs.

4. Discussion

Although mice are at the fore of mammalian research including modelling diseases such as cancer, they do not always replicate the human pathology, reducing their predictive value (Mak et al, 2014). The failure of many new drugs in clinical trials can be attributed to preclinical studies that fail to predict safety and effectiveness in human patients (Justice & Dhillon, 2016; Ledford, 2011). Non-rodent species such as pigs can provide additional information and improve the predictive value of preclinical studies (Bahr & Wolf, 2012; Perleberg et al, 2018). Work with laboratory animals requires to "replace, reduce and refine" their use whenever possible (Article 47 of Directive 2010/63/EU; National Center for the Replacement Refinement & Reduction of Animals in Research, https://www.nc3rs.org.uk/). It is thus important to ensure that all data gained is valuable and relevant to the disease studied. This is best achieved with well-defined animal models that replicate relevant aspects of human pathology as closely as possible. The *APC*^{1311} pigs, analysed in this work replicate hallmarks of human familial adenomatous polyposis (FAP) and colorectal cancer (CRC) including adenomatous polyps in the colorectum with loss of *APC* heterozygosity, β-catenin accumulation, upregulation of c-MYC, MAPK pathway activation and progression to carcinoma in situ and phenotypic variation in polyposis severity. More thorough characterisation of the model is required for use in translational biomedical research. In this work transcriptional analysis of normal mucosa and polyps from the *APC*^{1311} pigs, was aimed to identify elements in the genetic background such as single-nucleotide polymorphisms (SNPs), dysregulated genes, gene sets and microRNAs (miRNAs) that may contribute to susceptibility towards severe polyposis and tumour progression, respectively. This process should reveal both similarities to human molecular pathology and novel markers for early detection and drivers. A holistic approach was performed where whole mRNA and whole miRNA were sequenced using next generation sequencing technology and computational analysis pipelines to compare expression and distribution of genes, miRNAs and SNPs between high polyp and low polyp normal mucosa samples and between high grade and low grade intraepithelial neoplasia.

Therefore, the identity of genes modifying this susceptibility towards severe polyposis was investigated by comparative NGS-based expression analysis of mRNA and miRNA of normal mucosa samples between LP and HP animals. The generated data were used not only for comparative gene and miRNA expression analysis but also for gene set enrichment analysis, genome wide association studies of SNPs and allele-specific expression analysis of SNPs to detect molecular changes associated with severe polyposis of HP animals. The identification of modifier genes and the potential translation to human, may allow early screening of these genes via novel NGS methods in the future. A possible

screening method such as blood testing to identify individuals with susceptibility towards CRC could make the patients more aware of the risk and allow better preventive health care services.

To further validate the porcine *APC*^{1311} colorectal cancer model, comparative transcriptome analysis was performed between low and high-grade adenomas of the animals. Differentially expressed genes and miRNAs, SNPs and gene sets associated with high grade adenomas, were compared to human data. Further analysis was conducted analysing gene expression from laser microdissected low and high grade intraepithelial neoplasia, a study not yet performed in human, to identify drivers of CRC development that are masked by stromal expression patterns.

Diseases such as cancer, including CRC acquire multiple mutations and cancer promoting alterations over time. CRISPR/Cas9 technology was extended to the *APC*^{1311} pigs, allowing spatio-temporal knockout and mutation activation to model human CRC most accurately. *SpCas9* (isolated from *Streptococcus pyogenes*) was introduced into the *ROSA26* locus of porcine kidney-derived fibroblasts (pKFs) from *APC*^{1311} pigs to generate *APC*^{1311} pigs that express *SpCas9* ubiquitously to enable delivery of guide RNAs to the polyps to induce knockouts of tumour suppressor genes and oncogenic mutations by *in vivo* electroporation or adeno-associated viral vectors.

4.1 Characterisation of the porcine model for colorectal cancer by transcriptional analyses using next generation sequencing technology

The identification of both known and unknown genes and SNPs that may contribute to polyposis severity and CRC progression in normal mucosa and polyps of the *APC*^{1311} pigs respectively was performed in a holistic blind search approach to allow both, the search for specific targets and unknowns, at the same time. MRNA sequencing was preferred over genome sequencing because mRNA sequencing can give information not only about mutations and SNPs located in coding regions but also about changes in gene expression. Targeted analysis of DNA can follow later on. Computational analysis of sequencing data can be performed using different algorithms available, but transcriptional analysis of differential gene expression depends on alignment of sequenced reads to the porcine genome and assignment to features in the shape of a gene or transcript variant. If reads cannot be assigned to a feature in the shape of a gene or transcript variant they will not be compared between the analysed groups, and thus not detected even if they are differentially expressed. Furthermore, the lack and presence of different isoform annotations of a gene, influence

computational analysis strongly and determine if a gene is significantly differentially expressed or not. This is a disadvantage of mRNA sequencing compared to genome sequencing, specifically as the porcine genome is less well annotated than the human genome. However, the data sequenced remains extremely valuable, as the data can always be used at later time points for alignment to newer better annotated genome assemblies.

4.2 Attempt to identify modifier genes on mRNA level

Different degrees of polyposis severity among familial adenomatous polyposis (FAP) patients have been found to correlate in part with the site of APC mutation (Crabtree et al, 2002). However, variation in polyposis severity in patients with the same APC mutation have also been observed in humans evidencing the existence of modifier genes(Crabtree et al, 2002; Houlston et al, 2001). This severity in polyposis is attributed to increased initiation events rather than accelerated progression (Crabtree et al, 2001). Studies to investigate susceptibility towards colorectal cancer (CRC) have been performed in humans by searching for SNPs that associate with diseased patients versus controls (Broderick et al, 2007; Tomlinson et al, 2007; Tomlinson et al, 2008; Whiffin et al, 2014). Two SNPs associated with higher CRC risk (rs16892766 at 8q23.3 and rs3802842 at 11q23.1) have also been found to associate with more severe polyposis in FAP (Ghorbanoghli et al, 2016). Studies to identify modifier genes of severe FAP have been analysed in mice by cross-breeding of different inbred strains carrying Apc^{Min} alleles rather than genome wide association studies (Karim & Huso, 2013). Modifiers such as modifier of Min 1 (*Mom1*), *Mom5, Mom7, Mom12, Mom13* and *Pla2g2a* have been identified but showed no association with human FAP severity (Karim & Huso, 2013; Talseth-Palmer, 2017). This highlighted the dependence of polyposis severity to the genetic background, thus the status of genes that do not exhibit high penetrance and the need of other model organisms such as the APC^{1311} pigs that may help identify modifiers that associate with human polyposis severity.

Human studies to identify modifier genes have mainly been performed using blood samples and thus only genetic analyses were possible searching for SNPs. The affected tissue, the colonic mucosa has not been sampled, likely due to safety reasons as the sampling of supposedly normal mucosa might pose a health risk for the patients. The use of the APC^{1311} pig model allows comparative mRNA or miRNA sequencing analysis of colonic normal mucosa between animals with more and less severe polyposis in FAP, that has not been conducted in humans.

4.2.1 Gene expression analysis

MRNA Sequencing and two independent subsequent gene expression analyses of 16 high polyp (HP) versus 19 low polyp (LP) normal mucosa samples from animals between 3-9 months identified the genes *CYP7A1* and *SATB1* highly expressed and *SFRP5* lower expressed in HP animals. The gene *CYP7A1* was however the only gene significantly differentially expressed after multiple component testing (adjusted p-value). Significantly higher *CYP7A1* expression was confirmed using RT-qPCR. The reason for increased expression of *CYP7A1* was investigated by analysing cis regulatory factors such as SNP analysis of 5742 bp 5' of the ATG and CpG island (CGI) methylation status of two islands 3459 bp 5' to the ATG. 78 SNPs detected showed no significant association with either LP or HP animals. CGI methylation analysis however revealed significantly reduced methylation of CpG1 site in CGI2, which may influence the binding capacities of the transcription factor (TF) families signal transducer and activator of transcription 3 (STAT3), GLIS family zinc finger 1 (GLIS1), Ccaat/enhancer binding protein beta (CEBPB) and GA binding protein TF alpha (GABPA). This may lead to increased transcription activation in HP animals. Further by elucidating the gene structure, a skipping of exon 5 was identified. The differential expression was observed in the transcript variant with exon 5. The frequency of skipping might contribute to the differential expression. Therefore, Sanger sequencing of intron 4 was performed but analysis was not possible. However, this region is clearly relevant for exon skipping and needs to be thoroughly analysed with a new primer setup for Sanger sequencing. In Sanger sequencing, the quality of the first 15-40 bp and after 700-900 bp is problematic due to primer binding and lack of separation power of large fragments that differ in only 1 base. Heterozygous deletions or insertions are very difficult to analyse, and quality is reduced even earlier. The previous set up of primer walking where 700 bp fragments were sequenced from both sides with overlaps between fragments of 100 bp was not sufficient. Shorter fragments of 400-500 bp are required to elucidate intron 4 sequence.

By laser microdissection, differential expression of *CYP7A1* could be located to stromal rather than crypt composing cells. However, the set of 12 samples (5 HP and 7 LP) was small and the expression of the gene very low.

Cytochrome P450 family 7 subfamily A member 1 (CYP7A1) is a monooxygenase that catalyses bile acid synthesis from cholesterol in the liver. SNPs and high hepatic *CYP7A1* expression promote CRC via increased bile acid synthesis that promote CRC development (Gadaleta et al, 2017; Hagiwara et al, 2005; Wertheim et al, 2012). Increased *CYP7A1* expression in colon however has not been mentioned. To determine whether *CYP7A1* functions indeed as a modifier of FAP severity, it first necessary to identify the origion of the increased *CYP7A1* expression in HP animals. The identified increased *CYP7A1* expression could be attributed to three different situations in the APC^{1311} pigs. 1) The high *CYP7A1* expression in HP could reflect ubiquitous high *CYP7A1* levels (suggested by higher *CYP7A1* levels in HP

crypts (Figure 12)), causing also higher hepatic *CYP7A1* expression leading to increased cancer-promoting bile acids in the colon. To test this hypothesis RNA from other tissues including liver would need to be analysed for *CYP7A1* expression levels between HP and LP animals. 2) The *CYP7A1* expression in the colon may exhibit an unknown carcinogenic effect inducing increased polyposis in HP animals. 3) The increased *CYP7A1* expression detected in the HP normal mucosa samples may be caused by an increased immune cell infiltration, including macrophages that were associated with *CYP7A1* expression for cholesterol efflux (Bao et al, 2015), in response to severe polyposis. This hypothesis was inspired by the observation of increased immune infiltration in analysed polyps (Figure 16) and literature where specifically alveolar macrophages showed high *CYP7A1* expression in pigs (Freeman et al, 2012). The third hypothesis, would however mean that high *CYP7A1* is not a cause not a cause but a consequence of high polyposis that also affects areas of the mucosa not closely associated with polyps. To identify whether hypothesis two or three are more likely more investigations including immunohistochemistry of normal mucosa samples are essential to identify which cells of the stroma express increased *CYP7A1* levels.

Further the analysis of other genes contributing to the classic synthesis pathway of carcinogenic bile acids (deoxycholic acid and lithocholic acid) such as *CYP8B1*, *CYP27A1* and *HSD3B7* (Li & Chiang, 2014) in the colon of HP and LP animals could elucidate, whether increased CYP7A1 expression contributes to bile acids formation in the gut or whether the increased CYP7A1 expression found is independent of bile acid synthesis.

4.2.2 Gene set enrichment analysis

Gene set enrichment analysis has revealed the expression of few genes enriched in HP versus many gene sets enriched in LP samples. This is likely due to the fact that the LP group is more homogeneous than the HP group. The gene set of oestrogen response was enriched in the HP group analysed using pipeline 1 and even significantly with pipeline 2 (FDR< 0.25). Therefore, the distribution of male and female animals in the HP and LP group was analysed. Contradictory to the initial hypothesis HP animals were composed to 62.5 % of males (25% castrated; 37.5% non-castrated males) and the LP group to 26 % males (10% castrated; 16% non-castrated males). Meaning the HP group contained more non-castrated males in number and in proportion than the LP group, however, the HP group also contained more castrated males. Oestrogen signalling can exhibit a tumour-promoting, via oestrogen receptor α (ERα), or a tumour-suppressing capacity via ERβ (Caiazza et al, 2015). ERα levels in colon were generally reported to be low and ERβ the predominant oestrogen receptor, which is reduced during carcinogenesis compared to normal mucosa (Caiazza et al, 2015). However, components of the

oestrogen response gene set enriched in HP were *JAK2*, *PDZK1* and *SYBU*. Activation of JAK2/STAT3 signalling exhibits oncogenic potential (Alvarez et al, 2006; Yu et al, 2014) and that it's inhibition induced apoptosis in CRC cells (Du et al, 2012). Increased *JAK2* expression together with the reduced methylation of a potential STAT3 binding site in *CYP7A1* in the HP animals suggests oncogenic potential through increased activation in the HP group.

4.2.3 Single-nucleotide polymorphisms

Genome-wide association studies identifying SNPs in humans have been carried out in CRC-diseased patients compared to healthy individuals (Tomlinson et al, 2007; Tomlinson et al, 2008; Whiffin et al, 2014). The comparison of SNPs between pigs and humans specifically those outside of exons and even more so outside of genes is difficult as non-coding sequences are not as conserved as those coding for proteins. Nevertheless, SNP analysis was performed on the basis of the mRNA sequencing, comparing SNPs found in transcribed regions of the porcine genome between HP and LP animals. For this analysis the sequencing data was aligned to the Sscrofa11.1 genome assembly. This was performed under the premise that the novel assembly may have fewer gaps in the genomic sequence and possible incorrect gene annotation is irrelevant for this analysis.

However, the analysis of SNPs on RNA level has to be performed with caution due to a rare phenomenon in vertebrates called RNA-editing. These nucleotide substitutions occur mainly in repetitive elements but can also occur in coding sequences (Ramaswami et al, 2012). Therefore, the analysis of SNPs on RNA level should always be confirmed on DNA level as well, to avoid false positives (Kamps et al, 2017).

4.2.4 Computational analysis

With the increasing ease and reduced costs of genome and transcriptome sequencing, it has become a frequently used method for genome wide association studies but also gene expression experiments. RNA sequencing in specific gives tremendous amount of information of the gene expression without knowledge of the sequence. The management and analysis of these data has been tackled by the invention of numerous different pipelines using different algorythms and software. However, no one gold standard or consensus of analysis has been identified. The analysis of the data obtained by mRNA sequencing in this work was performed using two different pipelines. The pipelines worked with the same porcine genome annotation Sscrofa10.2 but different analysis algorythms and software were used to process and analyse the data. Both pipelines omitted adapter clipping as it can alter the gene

expression estimates (Williams et al, 2016). The main difference between the two pipelines applied was the alignment method. Pipeline 1 utilising the STAR aligner (Dobin et al, 2013), performs alignment of the reads directly to the genome, whereas pipeline 2 uses the pseudo alignment tool kallisto (Bray et al, 2016). Here the reads are broken up into k-mers and assigned to transcripts. However, it has been found that the alignment method does not necessarily have a major impact on the differential gene expression analysis. The specificity, accuracy and true positive rate seem to be really determined by the software utilised for the identification of the differential gene expression (Costa-Silva et al, 2017). Pipeline 1 used after duplicate marking (http://broadinstitute.github.io/picard) and gene assignment (Liao et al, 2014) DESeq.2 (Love et al, 2014). Pipeline two applied after transcript quantification sleuth software (Pimentel et al, 2017). Both software showed high specificity, accuracy and true positive rates in comparative studies, although the DESeq.2 performed slightly better (Costa-Silva et al, 2017). The application of more than one pipeline for differential gene expression analysis, as performed here, was found beneficial for the increase of specificity and true positive rate of the resulting differentially expressed genes (Costa-Silva et al, 2017).

4.3 Attempt to identify modifier genes on miRNA level

MiRNA sequencing was performed to identify other mechanisms that could cause differential protein levels between HP and LP animals that cannot be visualised using mRNA sequencing. MiRNAs target the 3' untranslated region of mRNAs and induce mRNA degradation which reduces the number of mRNA and protein or translational repression of their target mRNA (Shirafkan et al, 2018). Here the number of proteins is reduced but the mRNA amount is unaltered. It is believed that translational repression is less common than mRNA degradation (Guo et al, 2010; Hendrickson et al, 2009). However, miRNA-mediated translational repression cannot be detected by means of mRNA sequencing and can only be discovered by miRNA sequencing.

MiRNA sequencing and differential expression analysis of 9 HP and 10 LP normal mucosa samples from animals between 3-9 months revealed higher expression of miR-215 and 194b and lower expression of miR-27a-3p and 146a-5p in HP animals. Only the increased expression of miR-215 was significant (adjusted p-value < 0.05). The high expression of miR-215 and 194b was confirmed using RT-qPCR. Even though miR-215 and 194b have been associated with tumour suppressor functions, miR-215 has also been discussed to promote gastric cancer (Zang et al, 2017). Especially since no data of these early events during carcinogenesis initiation are available, it is likely that miRNAs might exhibit oncogenic and tumour suppressor functions at different time points during carcinogenesis. Additionally, *in silico*

target analysis using all differentially expressed miRNAs with a p-value <0.05 (13 miRNAs) revealed many cancer-associated pathways targeted. Also oestrogen signalling pathway was targeted by the differentially expressed miRNAs according to *in silico* analysis. This is particularly interesting as gene set enrichment analysis of the mRNA sequencing data, revealed oestrogen early response gene set enriched in HP animals.

In this sequencing project, data analysis presented two differential expression tables. One table contained all different isoforms. These isoforms are variants of mature canonical miRNAs and differ in length and sequence by 1 or 2 bp at the ends or in the middle (Guo & Chen, 2014), which is very hard to distinguish. Specifically, in the primer-based RT-qPCR distinguishability of the detected isoforms is very difficult. Primers could bind sequences with a SNP easily, thus detecting not only one isoform but several or all. As expression of canonical miRNAs and isoforms are strongly correlated as cooperative partners (Cloonan et al, 2011), the second table listing all canonical miRNAs according to differential expression of canonical and isoforms collectively was used for all analyses. This was also important for the *in silico* target analysis as only stem sequences were selectable.

Further, just like mRNA sequencing, miRNA sequencing is limited by the quality of the annotated miRNA database (MiRbase). The MiRbase release 22 documented 48885 mature miRNAs in 271 species. That includes 2654 mature miRNAs in homo sapiens and only 457 mature miRNAs in pigs (http://www.mirbase.org/cgi-bin/browse.pl). This highlights that the annotation of miRNAs in pigs is only about 1/5 of the human annotation, resulting in loss of differentially expressed miRNAs due to no feature assignment. But again the data acquired by sequencing can always be reanalysed using newer better annotated databases at later time points.

4.4 Analysis of tumour progression on mRNA level

Carcinogenesis, the development from normal tissue to an adenoma and even carcinoma is a multistep process that is characterised by the acquisition of certain key properties or hallmarks. Genome instability and increased proliferation are the first essential steps followed by the evasion of growth repressors and apoptosis. The pathways essential for the acquisition of these properties such as WNT, MAPK, TGFβ and TP53 are well known. However, the factors that can modulate these pathways are numerous. Therefore, targeted analysis for the identification for factors, was not possible especially because there are factors that might contribute to carcinogenesis that have not yet been identified. To perform a blind search for factors driving CRC carcinogenesis high grade (HG-IEN) and low grade intraepithelial neoplasia (LG-IEN) were compared. The reduction of *AHNAK* in HG-IEN bulk samples and

thus a reduced TGFβ signalling indicates that the evasion of growth repression has already occurred in HG-IEN compared to LG-IEN. Bulk sample analysis further revealed many genes lower expressed in HG that were associated with the immune system. This can certainly be attributed to the infiltrating immune cells that were detected in HG (Figure 16). However, as the immune related genes were lower expressed, the immune infiltration in LG must be stronger or the immune cell composition is different. One of the hallmarks of cancer is the evasion or repression of immune-mediated destruction. *IL7* which was significantly reduced in HG-IEN is an activation, growth and survival factor of T-cells (Shalapour et al, 2012). Therefore, a reduction of IL-7 would suggest also reduced mature T-cells and thus a less tumour suppressing immune milieu in the HG-IEN. However, the immune system is very complex and many more immune associated genes such as *IL21*, *CD40* and *IL20RA* were found to be differentially expressed. Therefore, it is certainly evident that the immune infiltration in HG-IEN differs from that in LG-IEN. Whether the change can be associated with increased immune response or the acquisition of immune repression requires more thorough analyses such as immunohistochemistry and flow cytometry to characterise the infiltrating immune cells.

The differential expression analysis of laser microdissected LG-IEN and HG-IEN then presented a different set of differentially expressed genes specifically excluding immune related genes that were seen in the bulk samples. This analysis deducted all stromal gene expression and only the gene expression coming from adenomatous epithelium were compared between LG and HG-IEN. Among the genes higher expressed in HG-IEN were genes associated with anti-inflammatory activity (*SLC30A1*), WNT regulation (*SMARCD3*) and p53 regulation and cell proliferation (*VASH1*). Furthermore, the gene *PLXND1* found highly expressed in HG-IEN, has been reported as highly expressed in human tumours (Roodink et al, 2009) associated with epithelial mesenchymal transition promoting invasiveness and metastasis (Casazza et al, 2010; Tseng et al, 2011). The results taken together validate on molecular level the advanced stage of the HG-IEN compared to the LG-IEN and present novel targets in treating CRC.

The genes differentially expressed in the HG-IEN compared to LG-IEN were numerous both in bulk and in laser microdissected samples, specifically compared to the susceptibility study where normal mucosa samples of HP and LP animals were compared. The cellular physiology is much more altered at the stage of neoplasia. Therefore, differential expression can reflect on the altered growth and differentiation properties (Fearon, 2011). Furthermore, many of the genes differentially expressed may be downstream targets of oncogenes or tumour suppressor genes essential for the cellular development of neoplasia, thus increase the amount of genes differentially expressed.

4.5 Optimisation of the CRC model

Diseases such as cancer, including colorectal cancer (CRC) acquire multiple mutations and cancer promoting alterations over time. To model CRC most accurately, animal models need to be generated that allow spatio-temporal knockout and mutation activation. So far recombinase systems such as the flipase-FRT (Flp-FRT) or Cre-loxP system have been utilised for sequential activation of mutations to model cancers (Schonhuber et al, 2014). The design of different lox sites that can be recognised by Cre recombinase enables multiple independent recombinations simultaneously (Sauer, 1996). Cre-expressing mice and pigs have been generated to perform tissue-specific activation of knockouts or mutations (Chen et al, 2010; Schonhuber et al, 2014). A drawback ofthis system specifically in large animals such as pigs, is however, that the generation of both the conditionally mutated animal and the particular Cre line is required, which is extremely time consuming in pigs. The CRISPR/Cas9 system is much more flexible, and requires generation of only the Cas9 pig, with which all kinds of mutations can be introduced. Further, it offers the opportunity to multiplex by the combination of the endonuclease Cas9 with multiple different gRNAs (Cong et al, 2013). By administration of the gRNAs at different time points, the CRISPR/Cas9 system also allows time specific activation of mutations and knockouts.

A *Cas9* expressing mouse was generated in combination with and without the Cre-loxP system (Platt et al, 2014). Thus, Cas9 expression is Cre-dependent. The potential of the model was shown by lung tumour induction by administering gRNAs targeting the three most frequently mutated genes in lung cancer by adeno associated viral vectors (AAVs). To expand this technology also to pigs, to enable acceleration of the CRC carcinogenesis and to better model CRC dynamics, *Cas9* was introduced into the genome of somatic cells from an APC^{1311} pig with the aim to generate APC^{1311}*/Cas9* pigs by nuclear transfer.

When introducing any transgene into the genome, it has to be made sure that the integration does not disrupt any essential genes. This is not only important for avoiding lethality caused by disruption of essential genes but also omitting phenotypes mediated by the mere introduction of a transgene. Therefore, Cas9 was introduced by gene targeting into the porcine *ROSA26* locus. The *Rosa26* locus has first been identified in mice (Friedrich & Soriano, 1991) and has quickly become the locus of choice when introducing transgenes in mice. The *Rosa26* gene does not translate into protein and targeting intron 1 of this locus resulted in viable fertile animals with constitutive ubiquitous expression of the integrated transgene (Zambrowicz et al, 1997). Silencing of transgenes integrated into the *Rosa26* locus has not been reported so far. Homologues, sharing the same properties, have been identified in

both human and pig (Irion et al, 2007; Kong et al, 2014; Li et al, 2014). Thus, viable, gene targeted pigs were generated by transgene introduction into the *ROSA26* locus (Kong et al, 2014; Li et al, 2014). In this work Cas9 was introduced into the porcine *ROSA26* locus via gene targeting using a promoter trap vector.

4.5.1 Generation and analysis of Cas9 targeted clones

Gene targeting via homologous recombination (HR) in murine embryonic stem (ES) cells has been very efficient (Capecchi, 1989). In pigs, however, no equivalent cells exist (Nowak-Imialek & Niemann, 2012). Gene targeting in pigs has to be performed in somatic cells, where HR is less efficient. Gene targeting using homologous arms that mediate homologous recombination of the vector and the genomic *ROSA26* locus was combined with the promoter trap strategy (Friedel et al, 2005).

A neomycin resistance gene (*neo*) with a splice acceptor (SA) was placed between the homologous arms of the promoter trap vector. As mice expressing CAG-controlled *SpCas9* (isolated from *Streptococcus pyogenes*) showed no Cas9-associated toxicity (Platt et al, 2014), CAG- driven *SpCas9* was also placed between the homologous arms and behind the SA-neo-polyA construct resulting in the GEMT-Rosa26-Neo-Cas9 vector.

Porcine kidney fibroblasts from *APC*^{1311} pigs were isolated and targeted using the linearized GEMT-Rosa26-Neo-Cas9 vector to further increase HR (Kucherlapati et al, 1984). PCR and subsequent sequencing of the PCR product of the 5' junction and 3' junction between endogenous locus and the cassette introduced, showed correct targeting. Southern blot and digital droplet PCR gave discordant results. Cell clones were characterised on mRNA and protein level, showing correct splicing, sufficient expression and protein translation of fully functional SpCas9 nuclease, capable to traffic into the nucleus, evidenced by TIDE analysis (https://tide.deskgen.com/). All four cell clones were used for nuclear transfer to generate *SpCas9* expressing pigs. The cell clones carrying both an *APC*^{1311} mutation and *SpCas9* expression have failed to generate a viable pig, while the clone 13, that only expresses *SpCas9* resulted in a viable Cas9 pig.

While the Cas9 expressing pig was being generated, a Cre-dependent *SpCas9* expressing pig was generated by a Chinese group (Wang et al, 2017) and proved that the Cas9 pig is a very powerful tool to generate porcine tumour models. They induced lung tumours by generating gRNAs targeting well known tumour suppressor genes including genes that were also targeted in this work *TP53*, *APC* and *PTEN*.

4.5.2 Application of *APC*^{1311}/*Cas9* pigs

The use of Cas9 pigs alone enable more accurate modelling of human diseases specifically for diseases such as cancer.

An *APC*^{1311}/*Cas9* pig, that can be generated by breeding the *SpCas9* pig to the *APC*^{1311} pig, will therefore enable a more accurate modelling of CRC and accelerate the carcinogenesis by sequential introduction of mutations into polyps *in vivo*. Such site specific introduction of mutations or knockouts in polyps can be achieved by *in vivo* electroporation and injection of the gRNA DNA sequences and single-stranded oligodeoxynucleotides (ssODNs) into the polyp. Once this will have led to full recapitulation of the human CRC carcinogenesis to an invasive carcinoma, the *APC*^{1311}/*Cas9* pig can be used for testing gene therapy to stop or reverse the process of carcinogenesis. This could be done by tissue specific or global application for gRNAs with ssODNs to correct CRC-causative mutations. Tissue specific and global delivery of gRNAs is more challenging than site-specific administration, but can be accomplished by the use of adeno associated viral vectors (AAVs).

Successful gene therapy in disease models is aimed at the application in humans that could allow correction of oncogenic mutations such as spontaneous but also hereditary APC mutations (Cooney et al, 2016; Steines et al, 2016). The CRISPR/Cas9 is a potential tool for gene therapy. However, the problem of off-target cleavage of the widely used *Streptococcus pyogenes*-isolated Cas9 (SpCas9) remains a safety risk (Cradick et al, 2013; Frock et al, 2015; Fu et al, 2013; Hsu et al, 2013; Kim et al, 2015; Pattanayak et al, 2013; Tsai et al, 2015). To reduce these off-targets the SpCas9 enzyme has been modified in their amino acid sequence rationally or randomly, resulting in novel versions of Cas9, SpCas9-HF1, eSpCas9(1.1), HypaCas9 and evoCas9 (Casini et al, 2018; Chen et al, 2017; Kleinstiver et al, 2016; Slaymaker et al, 2016). Comparison to the unmodified SpCas9 showed that off-target activity was markedly reduced. Therefore, for gene therapy one of the optimised Cas9 enzymes are more likely to be used.

However, for gene therapy use in humans using the CRISPR/Cas9 system, not only the gRNA and ssODNs would need to be delivered but also the enzyme itself. AAVs can be utilised for this purpose, however, they have a limited capacity of 4.7 kb in wildtype AAVs and up to 6 kb for AAVs with only 2 capsid subunits (Grieger & Samulski, 2005). SpCas9 and the modified versions of about 4000 bp are quite large, together with regulatory elements, gRNA and ssODN sequences breaching AAV capacity. Recent development of intein-mediated split-Cas9 and discovery of a smaller Cas9 orthologue isolated from *Campylobacter jejuni* (Kim et al, 2017; Truong et al, 2015), allow better AAV-mediated delivery. Therefore, the optimisation strategy of reducing off-target activity needs to be applied to these smaller versions and tested in animal models for diseases in order to make gene therapy a possibility.

5. Final remarks and outlook

This work has generated a vast amount of data and selected results require further validation and investigation.

Using Next generation RNA sequencing genes and miRNAs differentially expressed between *APC*^{1311} pigs with severe polyposis or very few polyps were identified. These potential modifiers of polyposis severity *CYP7A1* and miR-215 and 194b that were validated using PCR-based methods require further analysis to determine their functions in polyposis severity. *In silico* analyses have revealed connection to the oestrogen signalling in both the results of mRNA and miRNA sequencing. However, the CYP7A1 expressing cell compartment has to be identified via immunohistochemistry and its tumour promoting potential by increased bile acid production in the colon needs to be evidenced via bile acid analysis in the colon. The targets of miRNAs miR-215 and 194b need to be analysed *in vitro* and their contribution to polyposis promoting pathways need to be validated. The SNPs differentially distributed and expressed, also require further analysis elucidating their function on gene expression or miRNA expression and on cellular processes. Once, mechanisms are confirmed and better understood, these findings might be translatable to human CRC research, where the knowledge obtained may help screen patients for susceptibility factors and to provide susceptible patients with better preventive care.

The identification of genes and miRNAs differentially expressed between HG-IEN and LG-IEN by RNA sequencing in this work helped validate the potential of the porcine model to replicate molecular changes of human CRC and to identify novel drivers of CRC that mediate the progression from LG to HG-IEN. The data revealed that the change of immune cells in the stroma of IENs may mediate the progression. However closer immunological analyses including determination of the identity of the immune cells in LG and HG-IEN stroma by immunohistochemistry is vital for the understanding of this process. MiRNAs with higher expression in HG-IENs need to be analysed *in vitro* to identify their targets and their role in CRC progression. By RNA sequencing analysis of laser microdissected LG and HG-IEN, highlighted the influence of the tumour stroma and identified potential novel drivers of CRC such as *PLXND1* that was masked when hole polyps were analysed. Analysis of its migration and invasion-promoting properties require validation by *in vitro* migration assays.

The improvement of the *APC*^{1311} pig by crossbreeding with the SpCas9 pig generated during the time span of this work, will allow better modelling of human CRC and FAP by sequential introduction of oncogenic mutations via administration of gRNA locally or with adeno associated viral vectors. But also it will help to better understand the disease progression and to identify which potential driver mutation are really essential for the progression to CRC.

6. List of abbreviations

%	percent
°C	degrees celsius
µg	micro gram
µl	micro litre
µm	micro meter
µM	micro molar
AHNAK	AHNAK nucleoprotein
AKT	AKT serine/threonine kinase
APC	adenomatous polyposis coli
APCDD1	APC down-regulated 1
APS	ammonium persulphate
ASCL2	achaete-scute family bHLH transcription factor 2
bcDNA	bisulphite converted DNA
BCL2	BCL2, Apoptosis Regulator
bp	base pair
CAG	chicken beta-actin promoter and cytomegalovirus enhancer element
CDK4	cyclin dependent kinase 4
CDK6	cyclin dependent kinase 6
cDNA	complementary DNA
CEBPB	ccaat/enhancer binding protein beta
CGI	CpG island
CIMP	CGI methylator phenotype
CIN	chromosomal instability
c-MYC	myelocytomatosis proto-oncogene
CRC	colorectal cancer
Cre	Cre recombinase
CRE	cis-regulatry element
CRISPR	clustered regularly interspaced short palindromic repeat
crRNA	CRISPR RNA
CTNNB1	β-catenin
CYP7A1	cytochrome P450 family 7 subfamily A member 1
DCC	deleted in colorectal carcinoma
ddNTP	dideoxyribonucleotide triphosphate
ddPCR	digital droplet PCR
DKK1	dickkopf WNT signaling pathway inhibitor
DMEM	dulbecco's modified eagle's medium
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
D-PBS	dulbecco's phosphate-buffered saline
dUTP	deoxyuridine triphosphate
DVL	dishevelled

ECM	extracellular matrix
EDTA	ethylene diamine tetracetic acid
EGF	epithelial growth factor
EGFR	epithelial growth factor receptor
EMT	extra cellular matrix
EtOH	ethanol
EZH2	enhancer of zeste 2 polycomb repressive complex 2 subunit
FAP	familial adenomatous polyposis
FCS	fetal calf serum
FIT	faecal immunochemical test
FLP	flipase
FOBT	faecal occult blood test
FZD	frizzled
g	gram
g	gravitational force
G418	geneticin
GABPA	GA binding protein transcription factor alpha
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
gDNA	genomic DNA
GFP	green fluorescent protein
GLIS1	GLIS family zinc finger 1
gRNA	guide RNA
GSK3β	glycogen synthase kinase 3 beta
h	hour
HR	homologous recombination
HRP	horse radish peroxidase
IL-7	interleukin 7
kb	kilo base
kDa	kilo Dalton
KRAS	Kirsten rat sarcoma viral oncogene homolog proto-oncogene
l	litre
LEF	lymphoid enhancer binding factor
LGR5	Leucine Rich Repeat Containing G Protein-Coupled Receptor 5
LSL	loxP-stop-loxP
MAPK	mitogen-activated protein kinase
MDM2	MDM2 proto-oncogene
MEK	mitogen-activated protein kinase kinase
min	minute
Min	multiple intestinal neoplasia
miRNA	microRNA
ml	millilitre
mM	millimolar
MMP	matrix metallo proteinase
MMR	mismatch repair
mRNA	messenger RNA

MSI	microsatellite instable
MSS	microsatellite stable
mTOR	mechanistic target of rapamycin kinase
neo	neomycin
ng	nanogram
NGS	next generation sequencing
NRAS	neuroblastoma ras oncogene
NT	nuclear transfer
P14	cyclin dependent kinase inhibitor 2A
P21	cyclin dependent kinase inhibitor 1A
pADMSCs	porcine adipose derived mesenchymal stem cells
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PDCD4	programmed cell death 4
PGK	phosphoglycerate kinase
PI3K	phosphoinositide 3-kinase
pKFs	porcine kidney fibroblast
PLXND1	plexin D1
pM	picomolar
PTEN	phosphatase and tensin homolog
qPCR	quantitative PCR
RECK	reversion inducing cysteine rich protein with kazal motifs
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RT-qPCR	quantitative reverse transcription PCR
RUNX3	runt related transcription factor 3
SATB1	SATB homeobox 1
SCNT	somatic cell nuclear transfer
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	second
SFRP	secreted frizzled related protein
SIRT1	sirtuin 1
SNP	single-nucleotide polymorphism
ssODN	single-stranded oligodeoxynucleotide
STAT3	signal transducer and activator of transcription 3
TALEN	transcription activator-like effector nuclease
TCF	T-cell factor
TF	transcription factor
TGFβ	transforming growth factor beta
TGFβRI	transforming growth factor beta receptor 1
TGFβRII	transforming growth factor beta receptor 2
TIMP3	TIMP metallopeptidase inhibitor 3
TP53	tumour protein p53
tracrRNA	transactivating crRNA

TSP-1	thrombospondin 1
uPAR	plasminogen activator, urokinase receptor
ZEB1/2	zinc finger E-box binding homeobox 1/2
ZFN	zinc finger nuclease

7. List of figures

8. List of tables

9. Bibliography

Akao, Y., Nakagawa, Y. & Naoe, T. (2006) let-7 microRNA functions as a potential growth suppressor in human colon cancer cells. *Biol Pharm Bull*, 29(5), 903-6.

Akao, Y., Noguchi, S., Iio, A., Kojima, K., Takagi, T. & Naoe, T. (2011) Dysregulation of microRNA-34a expression causes drug-resistance to 5-FU in human colon cancer DLD-1 cells. *Cancer Lett*, 300(2), 197-204.

Akino, K., Toyota, M., Suzuki, H., Mita, H., Sasaki, Y., Ohe-Toyota, M., Issa, J. P., Hinoda, Y., Imai, K. & Tokino, T. (2005) The Ras effector RASSF2 is a novel tumor-suppressor gene in human colorectal cancer. *Gastroenterology*, 129(1), 156-69.

Al-Sohaily, S., Henderson, C., Selinger, C., Pangon, L., Segelov, E., Kohonen-Corish, M. R. & Warusavitarne, J. (2014) Loss of special AT-rich sequence-binding protein 1 (SATB1) predicts poor survival in patients with colorectal cancer. *Histopathology*, 65(2), 155-63.

Alberici, P. & Fodde, R. (2006) The role of the APC tumor suppressor in chromosomal instability. *Genome Dyn*, 1, 149-70.

Albuquerque, C., Breukel, C., van der Luijt, R., Fidalgo, P., Lage, P., Slors, F. J., Leitao, C. N., Fodde, R. & Smits, R. (2002) The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. *Hum Mol Genet*, 11(13), 1549-60.

Alvarez, J. V., Greulich, H., Sellers, W. R., Meyerson, M. & Frank, D. A. (2006) Signal transducer and activator of transcription 3 is required for the oncogenic effects of non-small-cell lung cancer-associated mutations of the epidermal growth factor receptor. *Cancer Res*, 66(6), 3162-8.

Andersen, C. L., Jensen, J. L. & Orntoft, T. F. (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res*, 64(15), 5245-50.

Anderson, S. (1981) Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res*, 9(13), 3015-27.

Anwar, S. L., Wulaningsih, W. & Lehmann, U. (2017) Transposable Elements in Human Cancer: Causes and Consequences of Deregulation. *Int J Mol Sci*, 18(5).

Article 47 of Directive 2010/63/EU (http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32010L0063&from=EN), retrieved 02.09.2018

Baba, Y., Nosho, K., Shima, K., Freed, E., Irahara, N., Philips, J., Meyerhardt, J. A., Hornick, J. L., Shivdasani, R. A., Fuchs, C. S. & Ogino, S. (2009) Relationship of CDX2 loss with molecular features and prognosis in colorectal cancer. *Clin Cancer Res*, 15(14), 4665-73.

Bahr, A. & Wolf, E. (2012) Domestic animal models for biomedical research. *Reprod Domest Anim*, 47 Suppl 4, 59-71.

Bandres, E., Agirre, X., Bitarte, N., Ramirez, N., Zarate, R., Roman-Gomez, J., Prosper, F. & Garcia-Foncillas, J. (2009) Epigenetic regulation of microRNA expression in colorectal cancer. *Int J Cancer*, 125(11), 2737-43.

Bandres, E., Cubedo, E., Agirre, X., Malumbres, R., Zarate, R., Ramirez, N., Abajo, A., Navarro, A., Moreno, I., Monzo, M. & Garcia-Foncillas, J. (2006) Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. *Mol Cancer*, 5, 29.

Bannister, A. J. & Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res*, 21(3), 381-95.

Bao, L. D., Li, C. Q., Peng, R., Ren, X. H., Ma, R. L., Wang, Y. & Lv, H. J. (2015) Correlation between the decrease of cholesterol efflux from macrophages in patients with type II diabetes mellitus and down-regulated CYP7A1 expression. *Genet Mol Res*, 14(3), 8716-24.

Barber, T. D., McManus, K., Yuen, K. W., Reis, M., Parmigiani, G., Shen, D., Barrett, I., Nouhi, Y., Spencer, F., Markowitz, S., Velculescu, V. E., Kinzler, K. W., Vogelstein, B., Lengauer, C. & Hieter, P. (2008) Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci U S A*, 105(9), 3443-8.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. & Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315(5819), 1709-12.

Behrens, J., von Kries, J. P., Kuhl, M., Bruhn, L., Wedlich, D., Grosschedl, R. & Birchmeier, W. (1996) Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature*, 382(6592), 638-42.

Beroud, C. & Soussi, T. (1996) APC gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Res*, 24(1), 121-4.

Bettington, M., Walker, N., Clouston, A., Brown, I., Leggett, B. & Whitehall, V. (2013) The serrated pathway to colorectal carcinoma: current concepts and challenges. *Histopathology*, 62(3), 367-86.

Black, J. C., Van Rechem, C. & Whetstine, J. R. (2012) Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell*, 48(4), 491-507.

Blanc, R. S. & Richard, S. (2017) Arginine Methylation: The Coming of Age. *Mol Cell*, 65(1), 8-24.

Bogdanove, A. J. & Voytas, D. F. (2011) TAL effectors: customizable proteins for DNA targeting. *Science*, 333(6051), 1843-6.

Boon, E. M., van der Neut, R., van de Wetering, M., Clevers, H. & Pals, S. T. (2002) Wnt signaling regulates expression of the receptor tyrosine kinase met in colorectal cancer. *Cancer Res*, 62(18), 5126-8.

Braun, C. J., Zhang, X., Savelyeva, I., Wolff, S., Moll, U. M., Schepeler, T., Orntoft, T. F., Andersen, C. L. & Dobbelstein, M. (2008) p53-Responsive micrornas 192 and 215 are capable of inducing cell cycle arrest. *Cancer Res*, 68(24), 10094-104.

Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*, 34(5), 525-7.

Brenner, H. & Tao, S. (2013) Superior diagnostic performance of faecal immunochemical tests for haemoglobin in a head-to-head comparison with guaiac based faecal occult blood test among 2235 participants of screening colonoscopy. *Eur J Cancer*, 49(14), 3049-54.

Brinster, R. L., Braun, R. E., Lo, D., Avarbock, M. R., Oram, F. & Palmiter, R. D. (1989) Targeted correction of a major histocompatibility class II E alpha gene by DNA microinjected into mouse eggs. *Proc Natl Acad Sci U S A*, 86(18), 7087-91.

Brocato, J. & Costa, M. (2015) SATB1 and 2 in colorectal cancer. *Carcinogenesis*, 36(2), 186-91.

Broderick, P., Carvajal-Carmona, L., Pittman, A. M., Webb, E., Howarth, K., Rowan, A., Lubbe, S., Spain, S., Sullivan, K., Fielding, S., Jaeger, E., Vijayakrishnan, J., Kemp, Z., Gorman, M., Chandler, I., Papaemmanuil, E., Penegar, S., Wood, W., Sellick, G., Qureshi, M., Teixeira, A., Domingo, E., Barclay, E., Martin, L., Sieber, O., Consortium, C., Kerr, D., Gray, R., Peto, J., Cazier, J. B., Tomlinson, I. & Houlston, R. S. (2007) A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat Genet*, 39(11), 1315-7.

Caiazza, F., Ryan, E. J., Doherty, G., Winter, D. C. & Sheahan, K. (2015) Estrogen receptors and their implications in colorectal carcinogenesis. *Front Oncol*, 5, 19.

Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M. & Croce, C. M. (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A*, 101(9), 2999-3004.

Callesen, H., Liu, Y., Pedersen, H. S., Li, R. & Schmidt, M. (2014) Increasing efficiency in production of cloned piglets. *Cell Reprogram*, 16(6), 407-10.

Canard, B. & Sarfati, R. S. (1994) DNA polymerase fluorescent substrates with reversible 3'-tags. *Gene*, 148(1), 1-6.

Cancer Genome Atlas, N. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, 487(7407), 330-7.

Capecchi, M. R. (1989) Altering the genome by homologous recombination. *Science*, 244(4910), 1288-92.

Carlson, D. F., Tan, W., Lillico, S. G., Stverakova, D., Proudfoot, C., Christian, M., Voytas, D. F., Long, C. R., Whitelaw, C. B. & Fahrenkrug, S. C. (2012) Efficient TALEN-mediated gene knockout in livestock. *Proc Natl Acad Sci U S A*, 109(43), 17382-7.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. & Werner, T. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, 21(13), 2933-42.

Casazza, A., Finisguerra, V., Capparuccia, L., Camperi, A., Swiercz, J. M., Rizzolio, S., Rolny, C., Christensen, C., Bertotti, A., Sarotto, I., Risio, M., Trusolino, L., Weitz, J., Schneider, M., Mazzone, M., Comoglio, P. M. & Tamagnone, L. (2010) Sema3E-Plexin D1 signaling drives human cancer cell invasiveness and metastatic spreading in mice. *J Clin Invest*, 120(8), 2684-98.

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demichelis, F., Inga, A. & Cereseto, A. (2018) A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat Biotechnol*, 36(3), 265-271.

Cassinotti, E., Boni, L., Segato, S., Rausei, S., Marzorati, A., Rovera, F., Dionigi, G., David, G., Mangano, A., Sambucci, D. & Dionigi, R. (2013) Free circulating DNA as a biomarker of colorectal cancer. *Int J Surg*, 11 Suppl 1, S54-7.

Cekaite, L., Eide, P. W., Lind, G. E., Skotheim, R. I. & Lothe, R. A. (2016) MicroRNAs as growth regulators, their function and biomarker status in colorectal cancer. *Oncotarget*, 7(6), 6476-505.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M. & Lee, J. J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4, 7.

Chang, K. H., Mestdagh, P., Vandesompele, J., Kerin, M. J. & Miller, N. (2010) MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. *BMC Cancer*, 10, 173.

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., Sternberg, S. H., Joung, J. K., Yildiz, A. & Doudna, J. A. (2017) Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*, 550(7676), 407-410.

Chen, L., Li, L., Pang, D., Li, Z., Wang, T., Zhang, M., Song, N., Yan, S., Lai, L. X. & Ouyang, H. (2010) Construction of transgenic swine with induced expression of Cre recombinase. *Animal*, 4(5), 767-71.

Chen, W. Y., Zhao, X. J., Yu, Z. F., Hu, F. L., Liu, Y. P., Cui, B. B., Dong, X. S. & Zhao, Y. S. (2015) The potential of plasma miRNAs for diagnosis and risk estimation of colorectal cancer. *Int J Clin Exp Pathol*, 8(6), 7092-101.

Chen, X., Guo, X., Zhang, H., Xiang, Y., Chen, J., Yin, Y., Cai, X., Wang, K., Wang, G., Ba, Y., Zhu, L., Wang, J., Yang, R., Zhang, Y., Ren, Z., Zen, K., Zhang, J. & Zhang, C. Y. (2009) Role of miR-143 targeting KRAS in colorectal tumorigenesis. *Oncogene*, 28(10), 1385-92.

Chiang, Y., Song, Y., Wang, Z., Liu, Z., Gao, P., Liang, J., Zhu, J., Xing, C. & Xu, H. (2012) microRNA-192, -194 and -215 are frequently downregulated in colorectal cancer. *Exp Ther Med*, 3(3), 560-566.

Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S. & Kim, J. S. (2013) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*.

Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S. & Kim, J. S. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*, 24(1), 132-41.

Chu, V. T., Weber, T., Graf, R., Sommermann, T., Petsch, K., Sack, U., Volchkov, P., Rajewsky, K. & Kuhn, R. (2016) Efficient generation of Rosa26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes. *BMC Biotechnol*, 16, 4.

Clark, K. J., Carlson, D. F., Foster, L. K., Kong, B. W., Foster, D. N. & Fahrenkrug, S. C. (2007) Enzymatic engineering of the porcine genome with transposons and recombinases. *BMC Biotechnol*, 7, 42.

Cloonan, N., Wani, S., Xu, Q., Gu, J., Lea, K., Heater, S., Barbacioru, C., Steptoe, A. L., Martin, H. C., Nourbakhsh, E., Krishnan, K., Gardiner, B., Wang, X., Nones, K., Steen, J. A., Matigian, N. A., Wood, D. L., Kassahn, K. S., Waddell, N., Shepherd, J., Lee, C., Ichikawa, J., McKernan, K., Bramlett, K., Kuersten, S. & Grimmond, S. M. (2011) MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biol*, 12(12), R126.

Cock, P. J., Chilton, J. M., Gruning, B., Johnson, J. E. & Soranzo, N. (2015) NCBI BLAST+ integrated into Galaxy. *Gigascience*, 4, 39.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121), 819-23.

Cooney, A. L., Abou Alaiwa, M. H., Shah, V. S., Bouzek, D. C., Stroik, M. R., Powers, L. S., Gansemer, N. D., Meyerholz, D. K., Welsh, M. J., Stoltz, D. A., Sinn, P. L. & McCray, P. B., Jr. (2016) Lentiviral-mediated phenotypic correction of cystic fibrosis pigs. *JCI Insight*, 1(14).

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., Abramowitz, L. K., Bartolomei, M. S., Rambow, F., Bassi, M. R., Bruno, T., Fanciulli, M., Renner, C., Klein-Szanto, A. J., Matsumoto, Y., Kobi, D., Davidson, I., Alberti, C., Larue, L. & Bellacosa, A. (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*, 146(1), 67-79.

Costa-Silva, J., Domingues, D. & Lopes, F. M. (2017) RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One*, 12(12), e0190152.

Crabtree, M. D., Tomlinson, I. P., Hodgson, S. V., Neale, K., Phillips, R. K. & Houlston, R. S. (2002) Explaining variation in familial adenomatous polyposis: relationship between genotype and phenotype and evidence for modifier genes. *Gut*, 51(3), 420-3.

Crabtree, M. D., Tomlinson, I. P., Talbot, I. C. & Phillips, R. K. (2001) Variability in the severity of colonic disease in familial adenomatous polyposis results from differences in tumour initiation rather than progression and depends relatively little on patient age. *Gut*, 49(4), 540-3.

Cradick, T. J., Fine, E. J., Antico, C. J. & Bao, G. (2013) CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*, 41(20), 9584-92.

Criscione, S. W., Zhang, Y., Thompson, W., Sedivy, J. M. & Neretti, N. (2014) Transcriptional landscape of repetitive elements in normal and cancer human cells. *BMC Genomics*, 15, 583.

Croner, R. S., Brueckl, W. M., Reingruber, B., Hohenberger, W. & Guenther, K. (2005) Age and manifestation related symptoms in familial adenomatous polyposis. *BMC Cancer*, 5, 24.

Dai, Y., Vaught, T. D., Boone, J., Chen, S. H., Phelps, C. J., Ball, S., Monahan, J. A., Jobst, P. M., McCreath, K. J., Lamborn, A. E., Cowell-Lucero, J. L., Wells, K. D., Colman, A., Polejaeva, I. A. & Ayares, D. L. (2002) Targeted disruption of the alpha1,3-galactosyltransferase gene in cloned pigs. *Nat Biotechnol*, 20(3), 251-5.

Daskalos, A., Nikolaidis, G., Xinarianos, G., Savvari, P., Cassidy, A., Zakopoulou, R., Kotsinas, A., Gorgoulis, V., Field, J. K. & Liloglou, T. (2009) Hypomethylation of retrotransposable elements correlates with genomic instability in non-small cell lung cancer. *Int J Cancer*, 124(1), 81-7.

Dawson, M. A. & Kouzarides, T. (2012) Cancer epigenetics: from mechanism to therapy. *Cell*, 150(1), 12-27.

de la Chapelle, A. & Hampel, H. (2010) Clinical relevance of microsatellite instability in colorectal cancer. *J Clin Oncol*, 28(20), 3380-7.

de Sousa, E. M. F., Colak, S., Buikhuisen, J., Koster, J., Cameron, K., de Jong, J. H., Tuynman, J. B., Prasetyanti, P. R., Fessler, E., van den Bergh, S. P., Rodermond, H., Dekker, E., van der Loos, C. M., Pals, S. T., van de Vijver, M. J., Versteeg, R., Richel, D. J., Vermeulen, L. & Medema, J. P. (2011) Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients. *Cell Stem Cell*, 9(5), 476-85.

Deaton, A. M. & Bird, A. (2011) CpG islands and the regulation of transcription. *Genes Dev*, 25(10), 1010-22.

Debinski, H. S., Love, S., Spigelman, A. D. & Phillips, R. K. (1996) Colorectal polyp counts and cancer risk in familial adenomatous polyposis. *Gastroenterology*, 110(4), 1028-30.

Deng, G., Chen, A., Hong, J., Chae, H. S. & Kim, Y. S. (1999) Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. *Cancer Res*, 59(9), 2029-33.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D. & Daly, M. J. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*, 43(5), 491-8.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15-21.

Du, W., Hong, J., Wang, Y. C., Zhang, Y. J., Wang, P., Su, W. Y., Lin, Y. W., Lu, R., Zou, W. P., Xiong, H. & Fang, J. Y. (2012) Inhibition of JAK2/STAT3 signalling induces colorectal cancer cell apoptosis via mitochondrial pathway. *J Cell Mol Med*, 16(8), 1878-88.

Dunlop, M. G., Dobbins, S. E., Farrington, S. M., Jones, A. M., Palles, C., Whiffin, N., Tenesa, A., Spain, S., Broderick, P., Ooi, L. Y., Domingo, E., Smillie, C., Henrion, M., Frampton, M., Martin, L., Grimes, G., Gorman, M., Semple, C., Ma, Y. P., Barclay, E., Prendergast, J., Cazier, J. B., Olver, B., Penegar, S., Lubbe, S., Chander, I., Carvajal-Carmona, L. G., Ballereau, S., Lloyd, A., Vijayakrishnan, J., Zgaga, L., Rudan, I., Theodoratou, E., Colorectal Tumour Gene Identification, C., Starr, J. M., Deary, I., Kirac, I., Kovacevic, D., Aaltonen, L. A., Renkonen-Sinisalo, L., Mecklin, J. P., Matsuda, K., Nakamura, Y., Okada, Y., Gallinger, S., Duggan, D. J., Conti, D., Newcomb, P., Hopper, J., Jenkins, M. A., Schumacher, F., Casey, G., Easton, D., Shah, M., Pharoah, P., Lindblom, A., Liu, T., Swedish Low-Risk Colorectal Cancer Study, G., Smith, C. G., West, H., Cheadle, J. P., Group, C. C., Midgley, R., Kerr, D. J., Campbell, H., Tomlinson, I. P. & Houlston, R. S. (2012) Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nat Genet*, 44(7), 770-6.

Duval, A. & Hamelin, R. (2002) Mutations at Coding Repeat Sequences in Mismatch Repair-deficient Human Cancers. *Toward a New Concept of Target Genes for Instability*, 62(9), 2447-2454.

Dziki, L., Pula, A., Stawiski, K., Mudza, B., Wlodarczyk, M. & Dziki, A. (2015) Patients' Awareness Of The Prevention And Treatment Of Colorectal Cancer. *Pol Przegl Chir*, 87(9), 459-63.

Earle, J. S., Luthra, R., Romans, A., Abraham, R., Ensor, J., Yao, H. & Hamilton, S. R. (2010) Association of microRNA expression with microsatellite instability status in colorectal adenocarcinoma. *J Mol Diagn*, 12(4), 433-40.

Easow, G., Teleman, A. A. & Cohen, S. M. (2007) Isolation of microRNA targets by miRNP immunopurification. *RNA*, 13(8), 1198-204.

Edwards, A., Voss, H., Rice, P., Civitello, A., Stegemann, J., Schwager, C., Zimmermann, J., Erfle, H., Caskey, C. T. & Ansorge, W. (1990) Automated DNA sequencing of the human HPRT locus. *Genomics*, 6(4), 593-608.

Estecio, M. R., Gharibyan, V., Shen, L., Ibrahim, A. E., Doshi, K., He, R., Jelinek, J., Yang, A. S., Yan, P. S., Huang, T. H., Tajara, E. H. & Issa, J. P. (2007) LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability. *PLoS One*, 2(5), e399.

Esteller, M., Hamilton, S. R., Burger, P. C., Baylin, S. B. & Herman, J. G. (1999) Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. *Cancer Res*, 59(4), 793-7.

Esteller, M., Sparks, A., Toyota, M., Sanchez-Cespedes, M., Capella, G., Peinado, M. A., Gonzalez, S., Tarafa, G., Sidransky, D., Meltzer, S. J., Baylin, S. B. & Herman, J. G. (2000) Analysis of adenomatous polyposis coli promoter hypermethylation in human cancer. *Cancer Res*, 60(16), 4366-71.

Evans, M. J. & Kaufman, M. H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292(5819), 154-6.

Ewing, B. & Green, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*, 8(3), 186-94.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*, 8(3), 175-85.

Fabian, M. R., Sonenberg, N. & Filipowicz, W. (2010) Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem*, 79, 351-79.

Fagotto, F., Jho, E., Zeng, L., Kurth, T., Joos, T., Kaufmann, C. & Costantini, F. (1999) Domains of axin involved in protein-protein interactions, Wnt pathway inhibition, and intracellular localization. *J Cell Biol*, 145(4), 741-56.

Fan, D., Lin, X., Zhang, F., Zhong, W., Hu, J., Chen, Y., Cai, Z., Zou, Y., He, X., Chen, X., Lan, P. & Wu, X. (2018) MicroRNA 26b promotes colorectal cancer metastasis by downregulating phosphatase and tensin homolog and wingless-type MMTV integration site family member 5A. *Cancer Sci*, 109(2), 354-362.

Fearnhead, N. S., Britton, M. P. & Bodmer, W. F. (2001) The ABC of APC. *Hum Mol Genet*, 10(7), 721-33.

Fearon, E. R. (2011) Molecular genetics of colorectal cancer. *Annu Rev Pathol*, 6, 479-507.

Fearon, E. R. & Vogelstein, B. (1990) A genetic model for colorectal tumorigenesis. *Cell*, 61(5), 759-67.

Feinberg, A. P. & Vogelstein, B. (1983a) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, 301(5895), 89-92.

Feinberg, A. P. & Vogelstein, B. (1983b) Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun*, 111(1), 47-54.

Feng, B., Dong, T. T., Wang, L. L., Zhou, H. M., Zhao, H. C., Dong, F. & Zheng, M. H. (2012) Colorectal cancer migration and invasion initiated by microRNA-106a. *PLoS One*, 7(8), e43452.

Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C. & Parkin, D. M. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*, 127(12), 2893-917.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D. & Bray, F. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*, 136(5), E359-86.

Fernandes, M. S., Carneiro, F., Oliveira, C. & Seruca, R. (2013) Colorectal cancer and RASSF family--a special emphasis on RASSF1A. *Int J Cancer*, 132(2), 251-8.

Fernandez-Rozadilla, C., Cazier, J. B., Tomlinson, I., Brea-Fernandez, A., Lamas, M. J., Baiget, M., Lopez-Fernandez, L. A., Clofent, J., Bujanda, L., Gonzalez, D., de Castro, L., Consortium, E., Hemminki, K., Bessa, X., Andreu, M., Jover, R., Xicola, R., Llor, X., Moreno, V., Castells, A., Castellvi-Bel, S., Carracedo, A. & Ruiz-Ponte, C. (2014) A genome-wide association study on copy-number variation identifies a 11q11 loss as a candidate susceptibility variant for colorectal cancer. *Hum Genet*, 133(5), 525-34.

Flisikowska, T., Merkl, C., Landmann, M., Eser, S., Rezaei, N., Cui, X., Kurome, M., Zakhartchenko, V., Kessler, B., Wieland, H., Rottmann, O., Schmid, R. M., Schneider, G., Kind, A., Wolf, E., Saur, D. &

Schnieke, A. (2012) A porcine model of familial adenomatous polyposis. *Gastroenterology*, 143(5), 1173-5 e1-7.

Flisikowska, T., Stachowiak, M., Xu, H., Wagner, A., Hernandez-Caceres, A., Wurmser, C., Perleberg, C., Pausch, H., Perkowska, A., Fischer, K., Frishman, D., Fries, R., Switonski, M., Kind, A., Saur, D., Schnieke, A. & Flisikowski, K. (2017) Porcine familial adenomatous polyposis model enables systematic analysis of early events in adenoma progression. *Sci Rep*, 7(1), 6613.

Fodde, R. & Smits, R. (2001) Disease model: familial adenomatous polyposis. *Trends Mol Med*, 7(8), 369-73.

Fodde, R., Smits, R. & Clevers, H. (2001) APC, signal transduction and genetic instability in colorectal cancer. *Nat Rev Cancer*, 1(1), 55-67.

Font, J. & Mackay, J. P. (2010) Beyond DNA: zinc finger domains as RNA-binding modules. *Methods Mol Biol*, 649, 479-91.

Freeman, T. C., Ivens, A., Baillie, J. K., Beraldi, D., Barnett, M. W., Dorward, D., Downing, A., Fairbairn, L., Kapetanovic, R., Raza, S., Tomoiu, A., Alberio, R., Wu, C., Su, A. I., Summers, K. M., Tuggle, C. K., Archibald, A. L. & Hume, D. A. (2012) A gene expression atlas of the domestic pig. *BMC Biol*, 10, 90.

Friedel, R. H., Plump, A., Lu, X., Spilker, K., Jolicoeur, C., Wong, K., Venkatesh, T. R., Yaron, A., Hynes, M., Chen, B., Okada, A., McConnell, S. K., Rayburn, H. & Tessier-Lavigne, M. (2005) Gene targeting using a promoterless gene trap vector ("targeted trapping") is an efficient method to mutate a large fraction of genes. *Proc Natl Acad Sci U S A*, 102(37), 13188-93.

Friedrich, G. & Soriano, P. (1991) Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev*, 5(9), 1513-23.

Frock, R. L., Hu, J., Meyers, R. M., Ho, Y. J., Kii, E. & Alt, F. W. (2015) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol*, 33(2), 179-86.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K. & Sander, J. D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*.

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*, 32(3), 279-84.

Gabriel, R., Lombardo, A., Arens, A., Miller, J. C., Genovese, P., Kaeppel, C., Nowrouzi, A., Bartholomae, C. C., Wang, J., Friedman, G., Holmes, M. C., Gregory, P. D., Glimm, H., Schmidt, M., Naldini, L. & von Kalle, C. (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol*, 29(9), 816-23.

Gadaleta, R. M., Garcia-Irigoyen, O. & Moschetta, A. (2017) Bile acids and colon cancer: Is FXR the solution of the conundrum? *Mol Aspects Med*, 56, 66-74.

Gama-Sosa, M. A., Slagel, V. A., Trewyn, R. W., Oxenhandler, R., Kuo, K. C., Gehrke, C. W. & Ehrlich, M. (1983) The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res*, 11(19), 6883-94.

Gardner, R. C., Howarth, A. J., Hahn, P., Brown-Luedi, M., Shepherd, R. J. & Messing, J. (1981) The complete nucleotide sequence of an infectious clone of cauliflower mosaic virus by M13mp7 shotgun sequencing. *Nucleic Acids Res*, 9(12), 2871-88.

Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A. H. & Moineau, S. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 468(7320), 67-71.

Garnett, M. J., Edelman, E. J., Heidorn, S. J., Greenman, C. D., Dastur, A., Lau, K. W., Greninger, P., Thompson, I. R., Luo, X., Soares, J., Liu, Q., Iorio, F., Surdez, D., Chen, L., Milano, R. J., Bignell, G. R., Tam, A. T., Davies, H., Stevenson, J. A., Barthorpe, S., Lutz, S. R., Kogera, F., Lawrence, K., McLaren-Douglas, A., Mitropoulos, X., Mironenko, T., Thi, H., Richardson, L., Zhou, W., Jewitt, F., Zhang, T., O'Brien, P., Boisvert, J. L., Price, S., Hur, W., Yang, W., Deng, X., Butler, A., Choi, H. G., Chang, J. W., Baselga, J., Stamenkovic, I., Engelman, J. A., Sharma, S. V., Delattre, O., Saez-Rodriguez, J., Gray, N. S.,

Settleman, J., Futreal, P. A., Haber, D. A., Stratton, M. R., Ramaswamy, S., McDermott, U. & Benes, C. H. (2012) Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*, 483(7391), 570-5.

Garofalo, M., Di Leva, G., Romano, G., Nuovo, G., Suh, S. S., Ngankeu, A., Taccioli, C., Pichiorri, F., Alder, H., Secchiero, P., Gasparini, P., Gonelli, A., Costinean, S., Acunzo, M., Condorelli, G. & Croce, C. M. (2009) miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation. *Cancer Cell*, 16(6), 498-509.

Garrels, W., Mates, L., Holler, S., Dalda, A., Taylor, U., Petersen, B., Niemann, H., Izsvak, Z., Ivics, Z. & Kues, W. A. (2011) Germline transgenic pigs by Sleeping Beauty transposition in porcine zygotes and targeted integration in the pig genome. *PLoS One*, 6(8), e23573.

Ghorbanoghli, Z., Nieuwenhuis, M. H., Houwing-Duistermaat, J. J., Jagmohan-Changur, S., Hes, F. J., Tops, C. M., Wagner, A., Aalfs, C. M., Verhoef, S., Gomez Garcia, E. B., Sijmons, R. H., Menko, F. H., Letteboer, T. G., Hoogerbrugge, N., van Wezel, T., Vasen, H. F. & Wijnen, J. T. (2016) Colorectal cancer risk variants at 8q23.3 and 11q23.1 are associated with disease phenotype in APC mutation carriers. *Fam Cancer*, 15(4), 563-70.

Goel, A., Arnold, C. N., Niedzwiecki, D., Carethers, J. M., Dowell, J. M., Wasserman, L., Compton, C., Mayer, R. J., Bertagnolli, M. M. & Boland, C. R. (2004) Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers. *Cancer Res*, 64(9), 3014-21.

Gordon, D., Abajian, C. & Green, P. (1998) Consed: a graphical tool for sequence finishing. *Genome Res*, 8(3), 195-202.

Grady, W. M. (2004) Genomic instability and colon cancer. *Cancer Metastasis Rev*, 23(1-2), 11-27.

Grady, W. M., Myeroff, L. L., Swinler, S. E., Rajput, A., Thiagalingam, S., Lutterbaugh, J. D., Neumann, A., Brattain, M. G., Chang, J., Kim, S. J., Kinzler, K. W., Vogelstein, B., Willson, J. K. & Markowitz, S. (1999) Mutational inactivation of transforming growth factor beta receptor type II in microsatellite stable colon cancers. *Cancer Res*, 59(2), 320-4.

Gregory, P. A., Bert, A. G., Paterson, E. L., Barry, S. C., Tsykin, A., Farshid, G., Vadas, M. A., Khew-Goodall, Y. & Goodall, G. J. (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*, 10(5), 593-601.

Grieger, J. C. & Samulski, R. J. (2005) Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J Virol*, 79(15), 9933-44.

Guanti, G., Resta, N., Simone, C., Cariola, F., Demma, I., Fiorente, P. & Gentile, M. (2000) Involvement of PTEN mutations in the genetic pathways of colorectal cancerogenesis. *Hum Mol Genet*, 9(2), 283-7.

Guinney, J., Dienstmann, R., Wang, X., de Reynies, A., Schlicker, A., Soneson, C., Marisa, L., Roepman, P., Nyamundanda, G., Angelino, P., Bot, B. M., Morris, J. S., Simon, I. M., Gerster, S., Fessler, E., De Sousa, E. M. F., Missiaglia, E., Ramay, H., Barras, D., Homicsko, K., Maru, D., Manyam, G. C., Broom, B., Boige, V., Perez-Villamil, B., Laderas, T., Salazar, R., Gray, J. W., Hanahan, D., Tabernero, J., Bernards, R., Friend, S. H., Laurent-Puig, P., Medema, J. P., Sadanandam, A., Wessels, L., Delorenzi, M., Kopetz, S., Vermeulen, L. & Tejpar, S. (2015) The consensus molecular subtypes of colorectal cancer. *Nat Med*, 21(11), 1350-6.

Guo, C., Sah, J. F., Beard, L., Willson, J. K., Markowitz, S. D. & Guda, K. (2008) The noncoding RNA, miR-126, suppresses the growth of neoplastic cells by targeting phosphatidylinositol 3-kinase signaling and is frequently lost in colon cancers. *Genes Chromosomes Cancer*, 47(11), 939-46.

Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, 466(7308), 835-40.

Guo, L. & Chen, F. (2014) A challenge for miRNA: multiple isomiRs in miRNAomics. *Gene*, 544(1), 1-7.

Ha, M. & Kim, V. N. (2014) Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol*, 15(8), 509-24.

Hagiwara, T., Kono, S., Yin, G., Toyomura, K., Nagano, J., Mizoue, T., Mibu, R., Tanaka, M., Kakeji, Y., Maehara, Y., Okamura, T., Ikejiri, K., Futami, K., Yasunami, Y., Maekawa, T., Takenaka, K., Ichimiya, H. & Imaizumi, N. (2005) Genetic polymorphism in cytochrome P450 7A1 and risk of colorectal cancer: the Fukuoka Colorectal Cancer Study. *Cancer Res*, 65(7), 2979-82.

Hai, T., Teng, F., Guo, R., Li, W. & Zhou, Q. (2014) One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. *Cell Res*, 24(3), 372-5.

Hammer, R. E., Pursel, V. G., Rexroad, C. E., Jr., Wall, R. J., Bolt, D. J., Ebert, K. M., Palmiter, R. D. & Brinster, R. L. (1985) Production of transgenic rabbits, sheep and pigs by microinjection. *Nature*, 315(6021), 680-3.

Harada, K., Hiraoka, S., Kato, J., Horii, J., Fujita, H., Sakaguchi, K. & Shiratori, Y. (2007) Genetic and epigenetic alterations of Ras signalling pathway in colorectal neoplasia: analysis based on tumour clinicopathological features. *Br J Cancer*, 97(10), 1425-31.

Hauschild, J., Petersen, B., Santiago, Y., Queisser, A. L., Carnwath, J. W., Lucas-Hahn, A., Zhang, L., Meng, X., Gregory, P. D., Schwinzer, R., Cost, G. J. & Niemann, H. (2011) Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proc Natl Acad Sci U S A*, 108(29), 12013-7.

Hayes, J., Peruzzi, P. P. & Lawler, S. (2014) MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol Med*, 20(8), 460-9.

He, L., He, X., Lim, L. P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., Jackson, A. L., Linsley, P. S., Chen, C., Lowe, S. W., Cleary, M. A. & Hannon, G. J. (2007) A microRNA component of the p53 tumour suppressor network. *Nature*, 447(7148), 1130-4.

Heinritz, S. N., Mosenthin, R. & Weiss, E. (2013) Use of pigs as a potential model for research into dietary modulation of the human gut microbiota. *Nutr Res Rev*, 26(2), 191-209.

Hellebrekers, D. M., Lentjes, M. H., van den Bosch, S. M., Melotte, V., Wouters, K. A., Daenen, K. L., Smits, K. M., Akiyama, Y., Yuasa, Y., Sanduleanu, S., Khalid-de Bakker, C. A., Jonkers, D., Weijenberg, M. P., Louwagie, J., van Criekinge, W., Carvalho, B., Meijer, G. A., Baylin, S. B., Herman, J. G., de Bruine, A. P. & van Engeland, M. (2009) GATA4 and GATA5 are potential tumor suppressors and biomarkers in colorectal cancer. *Clin Cancer Res*, 15(12), 3990-7.

Hendrickson, D. G., Hogan, D. J., McCullough, H. L., Myers, J. W., Herschlag, D., Ferrell, J. E. & Brown, P. O. (2009) Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. *PLoS Biol*, 7(11), e1000238.

Herman, J. G., Umar, A., Polyak, K., Graff, J. R., Ahuja, N., Issa, J. P., Markowitz, S., Willson, J. K., Hamilton, S. R., Kinzler, K. W., Kane, M. F., Kolodner, R. D., Vogelstein, B., Kunkel, T. A. & Baylin, S. B. (1998) Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A*, 95(12), 6870-5.

Heuberger, J. & Birchmeier, W. (2010) Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. *Cold Spring Harb Perspect Biol*, 2(2), a002915.

Hinoue, T., Weisenberger, D. J., Lange, C. P., Shen, H., Byun, H. M., Van Den Berg, D., Malik, S., Pan, F., Noushmehr, H., van Dijk, C. M., Tollenaar, R. A. & Laird, P. W. (2012) Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Res*, 22(2), 271-82.

Hofmann, A., Kessler, B., Ewerling, S., Weppert, M., Vogg, B., Ludwig, H., Stojkovic, M., Boelhauve, M., Brem, G., Wolf, E. & Pfeifer, A. (2003) Efficient transgenesis in farm animals by lentiviral vectors. *EMBO Rep*, 4(11), 1054-60.

Hol, L., Wilschut, J. A., van Ballegooijen, M., van Vuuren, A. J., van der Valk, H., Reijerink, J. C., van der Togt, A. C., Kuipers, E. J., Habbema, J. D. & van Leerdam, M. E. (2009) Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. *Br J Cancer*, 100(7), 1103-10.

Houlston, R., Crabtree, M., Phillips, R., Crabtree, M. & Tomlinson, I. (2001) Explaining differences in the severity of familial adenomatous polyposis and the search for modifier genes. *Gut*, 48(1), 1-5.

Hryniuk, A., Grainger, S., Savory, J. G. & Lohnes, D. (2014) Cdx1 and Cdx2 function as tumor suppressors. *J Biol Chem*, 289(48), 33343-54.

Hsu, H. C., Thiam, T. K., Lu, Y. J., Yeh, C. Y., Tsai, W. S., You, J. F., Hung, H. Y., Tsai, C. N., Hsu, A., Chen, H. C., Chen, S. J. & Yang, T. S. (2016) Mutations of KRAS/NRAS/BRAF predict cetuximab resistance in metastatic colorectal cancer patients. *Oncotarget*, 7(16), 22257-70.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G. & Zhang, F. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*, 31(9), 827-32.

Hung, K. E., Maricevich, M. A., Richard, L. G., Chen, W. Y., Richardson, M. P., Kunin, A., Bronson, R. T., Mahmood, U. & Kucherlapati, R. (2010) Development of a mouse model for sporadic and metastatic colon tumors and its use in assessing drug treatment. *Proc Natl Acad Sci U S A*, 107(4), 1565-70.

Hur, K., Cejas, P., Feliu, J., Moreno-Rubio, J., Burgos, E., Boland, C. R. & Goel, A. (2014) Hypomethylation of long interspersed nuclear element-1 (LINE-1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. *Gut*, 63(4), 635-46.

Huth, L., Jakel, J. & Dahl, E. (2014) Molecular Diagnostic Applications in Colorectal Cancer. *Microarrays (Basel)*, 3(3), 168-79.

Hwang, W. L., Jiang, J. K., Yang, S. H., Huang, T. S., Lan, H. Y., Teng, H. W., Yang, C. Y., Tsai, Y. P., Lin, C. H., Wang, H. W. & Yang, M. H. (2014) MicroRNA-146a directs the symmetric division of Snail-dominant colorectal cancer stem cells. *Nat Cell Biol*, 16(3), 268-80.

Imamura, Y., Hibi, K., Koike, M., Fujiwara, M., Kodera, Y., Ito, K. & Nakao, A. (2005) RUNX3 promoter region is specifically methylated in poorly-differentiated colorectal cancer. *Anticancer Res*, 25(4), 2627-30.

Imperiale, T. F., Ransohoff, D. F., Itzkowitz, S. H., Levin, T. R., Lavin, P., Lidgard, G. P., Ahlquist, D. A. & Berger, B. M. (2014) Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med*, 370(14), 1287-97.

Irahara, N., Baba, Y., Nosho, K., Shima, K., Yan, L., Dias-Santagata, D., Iafrate, A. J., Fuchs, C. S., Haigis, K. M. & Ogino, S. (2010) NRAS mutations are rare in colorectal cancer. *Diagn Mol Pathol*, 19(3), 157-63.

Irion, S., Luche, H., Gadue, P., Fehling, H. J., Kennedy, M. & Keller, G. (2007) Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat Biotechnol*, 25(12), 1477-82.

Ivics, Z., Garrels, W., Mates, L., Yau, T. Y., Bashir, S., Zidek, V., Landa, V., Geurts, A., Pravenec, M., Rulicke, T., Kues, W. A. & Izsvak, Z. (2014) Germline transgenesis in pigs by cytoplasmic microinjection of Sleeping Beauty transposons. *Nat Protoc*, 9(4), 810-27.

Jakobsen, J. E., Johansen, M. G., Schmidt, M., Dagnaes-Hansen, F., Dam, K., Gunnarsson, A., Liu, Y., Kragh, P. M., Li, R., Holm, I. E., Callesen, H., Mikkelsen, J. G., Nielsen, A. L. & Jorgensen, A. L. (2013) Generation of minipigs with targeted transgene insertion by recombinase-mediated cassette exchange (RMCE) and somatic cell nuclear transfer (SCNT). *Transgenic Res*, 22(4), 709-23.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E. & Forman, D. (2011) Global cancer statistics. *CA Cancer J Clin*, 61(2), 69-90.

Jesinghaus, M., Pfarr, N., Endris, V., Kloor, M., Volckmar, A. L., Brandt, R., Herpel, E., Muckenhuber, A., Lasitschka, F., Schirmacher, P., Penzel, R., Weichert, W. & Stenzinger, A. (2016) Genotyping of colorectal cancer for cancer precision medicine: Results from the IPH Center for Molecular Pathology. *Genes Chromosomes Cancer*, 55(6), 505-21.

Jetten, A. M. (2018) GLIS1-3 transcription factors: critical roles in the regulation of multiple physiological processes and diseases. *Cell Mol Life Sci*.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096), 816-21.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. & Doudna, J. (2013) RNA-programmed genome editing in human cells. *Elife*, 2, e00471.

Joung, J. K. & Sander, J. D. (2013) TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol*, 14(1), 49-55.

Jung, B., Doctolero, R. T., Tajima, A., Nguyen, A. K., Keku, T., Sandler, R. S. & Carethers, J. M. (2004) Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers. *Gastroenterology*, 126(3), 654-9.

Justice, M. J. & Dhillon, P. (2016) Using the mouse to model human disease: increasing validity and reproducibility. *Dis Model Mech*, 9(2), 101-3.

Kamps, R., Brandao, R. D., Bosch, B. J., Paulussen, A. D., Xanthoulea, S., Blok, M. J. & Romano, A. (2017) Next-Generation Sequencing in Oncology: Genetic Diagnosis, Risk Prediction and Cancer Classification. *Int J Mol Sci*, 18(2).

Kane, M. F., Loda, M., Gaida, G. M., Lipman, J., Mishra, R., Goldman, H., Jessup, J. M. & Kolodner, R. (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res*, 57(5), 808-11.

Kapitonov, V. V. & Jurka, J. (2008) A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet*, 9(5), 411-2; author reply 414.

Kararli, T. T. (1995) Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals. *Biopharm Drug Dispos*, 16(5), 351-80.

Karim, B. O. & Huso, D. L. (2013) Mouse models for colorectal cancer. *Am J Cancer Res*, 3(3), 240-50.

Karpf, A. R. & Matsui, S. (2005) Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. *Cancer Res*, 65(19), 8635-9.

Kawashima, E. H., Farinelli, L., Mayer, P. (1998) International patent no. WO1998044151A1, Method Of Nucleic Acid Amplification, https://www.lens.org/images/patent/WO/1998044151/A1/WO_%201998_044151_A1.pdf, retrieved 02.09.2018

Kawashima, E. H., Farinelli, L., Mayer, P. (1998) International patent no. WO1998044152A1, Method Of Nucleic Acid Sequencing; https://www.lens.org/images/patent/WO/1998044152/A1/WO_1998_044152_A1.pdf, retrieved 02.09.2018

Kent, O. A., Chivukula, R. R., Mullendore, M., Wentzel, E. A., Feldmann, G., Lee, K. H., Liu, S., Leach, S. D., Maitra, A. & Mendell, J. T. (2010) Repression of the miR-143/145 cluster by oncogenic Ras initiates a tumor-promoting feed-forward pathway. *Genes Dev*, 24(24), 2754-9.

Kikuchi, K., Ekwall, H., Tienthai, P., Kawai, Y., Noguchi, J., Kaneko, H. & Rodriguez-Martinez, H. (2002) Morphological features of lipid droplet transition during porcine oocyte fertilisation and early embryonic development to blastocyst in vivo and in vitro. *Zygote*, 10(4), 355-66.

Kim, B. C., Joo, J., Chang, H. J., Yeo, H. Y., Yoo, B. C., Park, B., Park, J. W., Sohn, D. K., Hong, C. W. & Han, K. S. (2014) A predictive model combining fecal calgranulin B and fecal occult blood tests can improve the diagnosis of colorectal cancer. *PLoS One*, 9(9), e106182.

Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H. R., Hwang, J., Kim, J. I. & Kim, J. S. (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods*, 12(3), 237-43, 1 p following 243.

Kim, E., Koo, T., Park, S. W., Kim, D., Kim, K., Cho, H. Y., Song, D. W., Lee, K. J., Jung, M. H., Kim, S., Kim, J. H., Kim, J. H. & Kim, J. S. (2017) In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nat Commun*, 8, 14500.

Kinzler, K. W., Nilbert, M. C., Su, L. K., Vogelstein, B., Bryan, T. M., Levy, D. B., Smith, K. J., Preisinger, A. C., Hedge, P., McKechnie, D. & et al. (1991) Identification of FAP locus genes from chromosome 5q21. *Science*, 253(5020), 661-5.

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z. & Joung, J. K. (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, 529(7587), 490-5.

Klose, R. J. & Bird, A. P. (2006) Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci*, 31(2), 89-97.

Kobayashi, M., Honma, T., Matsuda, Y., Suzuki, Y., Narisawa, R., Ajioka, Y. & Asakura, H. (2000) Nuclear translocation of beta-catenin in colorectal cancer. *Br J Cancer*, 82(10), 1689-93.

Kong, Q., Hai, T., Ma, J., Huang, T., Jiang, D., Xie, B., Wu, M., Wang, J., Song, Y., Wang, Y., He, Y., Sun, J., Hu, K., Guo, R., Wang, L., Zhou, Q., Mu, Y. & Liu, Z. (2014) Rosa26 locus supports tissue-specific promoter driving transgene expression specifically in pig. *PLoS One*, 9(9), e107945.

Korinek, V., Barker, N., Morin, P. J., van Wichen, D., de Weger, R., Kinzler, K. W., Vogelstein, B. & Clevers, H. (1997) Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science*, 275(5307), 1784-7.

Kucherlapati, R. S., Eves, E. M., Song, K. Y., Morse, B. S. & Smithies, O. (1984) Homologous recombination between plasmids in mammalian cells can be enhanced by treatment of input DNA. *Proc Natl Acad Sci U S A*, 81(10), 3153-7.

Kurome, M., Geistlinger, L., Kessler, B., Zakhartchenko, V., Klymiuk, N., Wuensch, A., Richter, A., Baehr, A., Kraehe, K., Burkhardt, K., Flisikowski, K., Flisikowska, T., Merkl, C., Landmann, M., Durkovic, M., Tschukes, A., Kraner, S., Schindelhauer, D., Petri, T., Kind, A., Nagashima, H., Schnieke, A., Zimmer, R. & Wolf, E. (2013) Factors influencing the efficiency of generating genetically engineered pigs by nuclear transfer: multi-factorial analysis of a large data set. *BMC Biotechnol*, 13, 43.

Kwon, D. N., Lee, K., Kang, M. J., Choi, Y. J., Park, C., Whyte, J. J., Brown, A. N., Kim, J. H., Samuel, M., Mao, J., Park, K. W., Murphy, C. N., Prather, R. S. & Kim, J. H. (2013) Production of biallelic CMP-Neu5Ac hydroxylase knock-out pigs. *Sci Rep*, 3, 1981.

Labianca, R., Nordlinger, B., Beretta, G. D., Mosconi, S., Mandala, M., Cervantes, A., Arnold, D. & Group, E. G. W. (2013) Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 24 Suppl 6, vi64-72.

LaDuca, H., Stuenkel, A. J., Dolinsky, J. S., Keiles, S., Tandy, S., Pesaran, T., Chen, E., Gau, C. L., Palmaer, E., Shoaepour, K., Shah, D., Speare, V., Gandomi, S. & Chao, E. (2014) Utilization of multigene panels in hereditary cancer predisposition testing: analysis of more than 2,000 patients. *Genet Med*, 16(11), 830-7.

Lahtz, C. & Pfeifer, G. P. (2011) Epigenetic changes of DNA repair genes in cancer. *J Mol Cell Biol*, 3(1), 51-8.

Lai, L., Kolber-Simonds, D., Park, K. W., Cheong, H. T., Greenstein, J. L., Im, G. S., Samuel, M., Bonk, A., Rieke, A., Day, B. N., Murphy, C. N., Carter, D. B., Hawley, R. J. & Prather, R. S. (2002) Production of alpha-1,3-galactosyltransferase knockout pigs by nuclear transfer cloning. *Science*, 295(5557), 1089-92.

Leary, R. J., Lin, J. C., Cummins, J., Boca, S., Wood, L. D., Parsons, D. W., Jones, S., Sjoblom, T., Park, B. H., Parsons, R., Willis, J., Dawson, D., Willson, J. K., Nikolskaya, T., Nikolsky, Y., Kopelovich, L., Papadopoulos, N., Pennacchio, L. A., Wang, T. L., Markowitz, S. D., Parmigiani, G., Kinzler, K. W., Vogelstein, B. & Velculescu, V. E. (2008) Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proc Natl Acad Sci U S A*, 105(42), 16224-9.

Ledford, H. (2011) Translational research: 4 ways to fix the clinical trial. *Nature*, 477(7366), 526-8.

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L. J., 3rd, Lohr, J. G., Harris, C. C., Ding, L., Wilson, R. K., Wheeler, D. A., Gibbs, R. A., Kucherlapati, R., Lee, C., Kharchenko, P. V., Park, P. J. & Cancer Genome Atlas Research, N. (2012) Landscape of somatic retrotransposition in human cancers. *Science*, 337(6097), 967-71.

Lee, I. H., Sohn, M., Lim, H. J., Yoon, S., Oh, H., Shin, S., Shin, J. H., Oh, S. H., Kim, J., Lee, D. K., Noh, D. Y., Bae, D. S., Seong, J. K. & Bae, Y. S. (2014) Ahnak functions as a tumor suppressor via modulation of TGFbeta/Smad signaling pathway. *Oncogene*, 33(38), 4675-84.

Lengauer, C., Kinzler, K. W. & Vogelstein, B. (1997) Genetic instability in colorectal cancers. *Nature*, 386(6625), 623-7.

Leuchs, S., Saalfrank, A., Merkl, C., Flisikowska, T., Edlinger, M., Durkovic, M., Rezaei, N., Kurome, M., Zakhartchenko, V., Kessler, B., Flisikowski, K., Kind, A., Wolf, E. & Schnieke, A. (2012) Inactivation and inducible oncogenic mutation of p53 in gene targeted pigs. *PLoS One*, 7(10), e43323.

Lewis, B. P., Burge, C. B. & Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120(1), 15-20.

Li, E., Bestor, T. H. & Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, 69(6), 915-26.

Li, J., Liang, H., Bai, M., Ning, T., Wang, C., Fan, Q., Wang, Y., Fu, Z., Wang, N., Liu, R., Zen, K., Zhang, C. Y., Chen, X. & Ba, Y. (2015a) miR-135b Promotes Cancer Progression by Targeting Transforming Growth Factor Beta Receptor II (TGFBR2) in Colorectal Cancer. *PLoS One*, 10(6), e0130194.

Li, Q., Zou, C., Zou, C., Han, Z., Xiao, H., Wei, H., Wang, W., Zhang, L., Zhang, X., Tang, Q., Zhang, C., Tao, J., Wang, X. & Gao, X. (2013) MicroRNA-25 functions as a potential tumor suppressor in colon cancer by targeting Smad7. *Cancer Lett*, 335(1), 168-74.

Li, S., Edlinger, M., Saalfrank, A., Flisikowski, K., Tschukes, A., Kurome, M., Zakhartchenko, V., Kessler, B., Saur, D., Kind, A., Wolf, E., Schnieke, A. & Flisikowska, T. (2015b) Viable pigs with a conditionally-activated oncogenic KRAS mutation. *Transgenic Res*, 24(3), 509-17.

Li, S., Flisikowska, T., Kurome, M., Zakhartchenko, V., Kessler, B., Saur, D., Kind, A., Wolf, E., Flisikowski, K. & Schnieke, A. (2014) Dual fluorescent reporter pig for Cre recombination: transgene placement at the ROSA26 locus. *PLoS One*, 9(7), e102455.

Li, T. & Chiang, J. Y. (2014) Bile acid signaling in metabolic disease and drug therapy. *Pharmacol Rev*, 66(4), 948-83.

Liao, Y., Smyth, G. K. & Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923-30.

Lillico, S. G., Proudfoot, C., Carlson, D. F., Stverakova, D., Neil, C., Blain, C., King, T. J., Ritchie, W. A., Tan, W., Mileham, A. J., McLaren, D. G., Fahrenkrug, S. C. & Whitelaw, C. B. (2013) Live pigs produced from genome edited zygotes. *Sci Rep*, 3, 2847.

Lim, S. H., Becker, T. M., Chua, W., Caixeiro, N. J., Ng, W. L., Kienzle, N., Tognela, A., Lumba, S., Rasko, J. E., de Souza, P. & Spring, K. J. (2013) Circulating tumour cells and circulating free nucleic acid as prognostic and predictive biomarkers in colorectal cancer. *Cancer Lett*.

Lin, S. L., Miller, J. D. & Ying, S. Y. (2006) Intronic microRNA (miRNA). *J Biomed Biotechnol*, 2006(4), 26818.

Liu, J., Nau, M. M., Zucman-Rossi, J., Powell, J. I., Allegra, C. J. & Wright, J. J. (1997) LINE-I element insertion at the t(11;22) translocation breakpoint of a desmoplastic small round cell tumor. *Genes Chromosomes Cancer*, 18(3), 232-9.

Liu, L., Nie, J., Chen, L., Dong, G., Du, X., Wu, X., Tang, Y. & Han, W. (2013) The oncogenic role of microRNA-130a/301a/454 in human colorectal cancer via targeting Smad4 expression. *PLoS One*, 8(2), e55532.

Liu, W., Dong, X., Mai, M., Seelan, R. S., Taniguchi, K., Krishnadath, K. K., Halling, K. C., Cunningham, J. M., Boardman, L. A., Qian, C., Christensen, E., Schmidt, S. S., Roche, P. C., Smith, D. I. & Thibodeau, S. N. (2000) Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. *Nat Genet*, 26(2), 146-7.

Livak, K. J. & Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25(4), 402-8.

Logan, J. S. & Martin, M. J. (1994) Transgenic swine as a recombinant production system for human hemoglobin. *Methods Enzymol*, 231, 435-45.

Love, M. I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 15(12), 550.

Lu, D., Yao, Q., Zhan, C., Le-Meng, Z., Liu, H., Cai, Y., Tu, C., Li, X., Zou, Y. & Zhang, S. (2017) MicroRNA-146a promote cell migration and invasion in human colorectal cancer via carboxypeptidase M/src-FAK pathway. *Oncotarget*, 8(14), 22674-22684.

Lu, Z., Ghosh, S., Wang, Z. & Hunter, T. (2003) Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion. *Cancer Cell*, 4(6), 499-515.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, 389(6648), 251-60.

Luo, Y., Li, J., Liu, Y., Lin, L., Du, Y., Li, S., Yang, H., Vajta, G., Callesen, H., Bolund, L. & Sorensen, C. B. (2011) High efficiency of BRCA1 knockout using rAAV-mediated gene targeting: developing a pig model for breast cancer. *Transgenic Res*, 20(5), 975-88.

Lv, J. H., Wang, F., Shen, M. H., Wang, X. & Zhou, X. J. (2016) SATB1 expression is correlated with beta-catenin associated epithelial-mesenchymal transition in colorectal cancer. *Cancer Biol Ther*, 17(3), 254-61.

Mak, I. W., Evaniew, N. & Ghert, M. (2014) Lost in translation: animal models and clinical trials in cancer treatment. *Am J Transl Res*, 6(2), 114-8.

Malapelle, U., Vigliar, E., Sgariglia, R., Bellevicine, C., Colarossi, L., Vitale, D., Pallante, P. & Troncone, G. (2015) Ion Torrent next-generation sequencing for routine identification of clinically relevant mutations in colorectal cancer patients. *J Clin Pathol*, 68(1), 64-8.

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L. & Church, G. M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*, 31(9), 833-8.

Malkhosyan, S., Rampino, N., Yamamoto, H. & Perucho, M. (1996) Frameshift mutator mutations. *Nature*, 382(6591), 499-500.

Mann, B., Gelos, M., Siedow, A., Hanski, M. L., Gratchev, A., Ilyas, M., Bodmer, W. F., Moyer, M. P., Riecken, E. O., Buhr, H. J. & Hanski, C. (1999) Target genes of beta-catenin-T cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. *Proc Natl Acad Sci U S A*, 96(4), 1603-8.

Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R. S., Zborowska, E., Kinzler, K. W., Vogelstein, B. & et al. (1995) Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science*, 268(5215), 1336-8.

McCalla-Martin, A. C., Chen, X., Linder, K. E., Estrada, J. L. & Piedrahita, J. A. (2010) Varying phenotypes in swine versus murine transgenic models constitutively expressing the same human Sonic hedgehog transcriptional activator, K5-HGLI2 Delta N. *Transgenic Res*, 19(5), 869-87.

McCarthy, D. J., Chen, Y. & Smyth, G. K. (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*, 40(10), 4288-97.

McCreath, K. J., Howcroft, J., Campbell, K. H., Colman, A., Schnieke, A. E. & Kind, A. J. (2000) Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. *Nature*, 405(6790), 1066-9.

Mehlen, P. & Fearon, E. R. (2004) Role of the dependence receptor DCC in colorectal cancer pathogenesis. *J Clin Oncol*, 22(16), 3420-8.

Mehrvarz Sarshekeh, A., Advani, S., Overman, M. J., Manyam, G., Kee, B. K., Fogelman, D. R., Dasari, A., Raghav, K., Vilar, E., Manuel, S., Shureiqi, I., Wolff, R. A., Patel, K. P., Luthra, R., Shaw, K., Eng, C., Maru, D. M., Routbort, M. J., Meric-Bernstam, F. & Kopetz, S. (2017) Association of SMAD4 mutation

with patient demographics, tumor characteristics, and clinical outcomes in colorectal cancer. *PLoS One*, 12(3), e0173345.

Meyer, M., de Angelis, M. H., Wurst, W. & Kuhn, R. (2010) Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci U S A*, 107(34), 15022-6.

Michael, M. Z., SM, O. C., van Holst Pellekaan, N. G., Young, G. P. & James, R. J. (2003) Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol Cancer Res*, 1(12), 882-91.

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., Kinzler, K. W., Vogelstein, B. & Nakamura, Y. (1992) Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res*, 52(3), 643-5.

Mir, R., Pradhan, S. J., Patil, P., Mulherkar, R. & Galande, S. (2016) Wnt/beta-catenin signaling regulated SATB1 promotes colorectal cancer tumorigenesis and progression. *Oncogene*, 35(13), 1679-91.

Misso, G., Di Martino, M. T., De Rosa, G., Farooqi, A. A., Lombardi, A., Campani, V., Zarone, M. R., Gulla, A., Tagliaferri, P., Tassone, P. & Caraglia, M. (2014) Mir-34: a new weapon against cancer? *Mol Ther Nucleic Acids*, 3, e194.

Miyaoka, Y., Berman, J. R., Cooper, S. B., Mayerl, S. J., Chan, A. H., Zhang, B., Karlin-Neumann, G. A. & Conklin, B. R. (2016) Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci Rep*, 6, 23549.

Miyoshi, Y., Nagase, H., Ando, H., Horii, A., Ichii, S., Nakatsuru, S., Aoki, T., Miki, Y., Mori, T. & Nakamura, Y. (1992) Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. *Hum Mol Genet*, 1(4), 229-33.

Molinari, F. & Frattini, M. (2013) Functions and Regulation of the PTEN Gene in Colorectal Cancer. *Front Oncol*, 3, 326.

Mongroo, P. S. & Rustgi, A. K. (2010) The role of the miR-200 family in epithelial-mesenchymal transition. *Cancer Biol Ther*, 10(3), 219-22.

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., Altshuler, D. & Groop, L. C. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*, 34(3), 267-73.

Moran, A., Ortega, P., de Juan, C., Fernandez-Marcelo, T., Frias, C., Sanchez-Pernaute, A., Torres, A. J., Diaz-Rubio, E., Iniesta, P. & Benito, M. (2010) Differential colorectal carcinogenesis: Molecular basis and clinical relevance. *World J Gastrointest Oncol*, 2(3), 151-8.

Morin, P. J., Sparks, A. B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B. & Kinzler, K. W. (1997) Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science*, 275(5307), 1787-90.

Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B. & Polakis, P. (1995) Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proc Natl Acad Sci U S A*, 92(7), 3046-50.

Mussolino, C. & Cathomen, T. (2012) TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol*, 23(5), 644-50.

Nagel, R., le Sage, C., Diosdado, B., van der Waal, M., Oude Vrielink, J. A., Bolijn, A., Meijer, G. A. & Agami, R. (2008) Regulation of the adenomatous polyposis coli gene by the miR-135 family in colorectal cancer. *Cancer Res*, 68(14), 5795-802.

National Center for the Replacement Refinement & Reduction of Animals in Research, https://www.nc3rs.org.uk/

Nickerson, D. A., Tobe, V. O. & Taylor, S. L. (1997) PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res*, 25(14), 2745-51.

Nishida, N., Yokobori, T., Mimori, K., Sudo, T., Tanaka, F., Shibata, K., Ishii, H., Doki, Y., Kuwano, H. & Mori, M. (2011) MicroRNA miR-125b is a prognostic marker in human colorectal cancer. *Int J Oncol*, 38(5), 1437-43.

Nishimura, J., Handa, R., Yamamoto, H., Tanaka, F., Shibata, K., Mimori, K., Takemasa, I., Mizushima, T., Ikeda, M., Sekimoto, M., Ishii, H., Doki, Y. & Mori, M. (2012) microRNA-181a is associated with poor prognosis of colorectal cancer. *Oncol Rep*, 28(6), 2221-6.

Nowak-Imialek, M. & Niemann, H. (2012) Pluripotent cells in farm animals: state of the art and future perspectives. *Reprod Fertil Dev*, 25(1), 103-28.

Nyiraneza, C., Sempoux, C., Detry, R., Kartheuser, A. & Dahan, K. (2012) Hypermethylation of the 5' CpG island of the p14ARF flanking exon 1beta in human colorectal cancer displaying a restricted pattern of p53 overexpression concomitant with increased MDM2 expression. *Clin Epigenetics*, 4(1), 9.

Nyren, P. & Lundin, A. (1985) Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. *Anal Biochem*, 151(2), 504-9.

O'Driscoll, L. (2007) Extracellular nucleic acids and their potential as diagnostic, prognostic and predictive biomarkers. *Anticancer Res*, 27(3A), 1257-65.

Okano, M., Bell, D. W., Haber, D. A. & Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99(3), 247-57.

Orom, U. A., Nielsen, F. C. & Lund, A. H. (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell*, 30(4), 460-71.

Ota, T., Doi, K., Fujimoto, T., Tanaka, Y., Ogawa, M., Matsuzaki, H., Kuroki, M., Miyamoto, S., Shirasawa, S. & Tsunoda, T. (2012) KRAS up-regulates the expression of miR-181a, miR-200c and miR-210 in a three-dimensional-specific manner in DLD-1 colorectal cancer cells. *Anticancer Res*, 32(6), 2271-5.

Palles, C., Cazier, J. B., Howarth, K. M., Domingo, E., Jones, A. M., Broderick, P., Kemp, Z., Spain, S. L., Guarino, E., Salguero, I., Sherborne, A., Chubb, D., Carvajal-Carmona, L. G., Ma, Y., Kaur, K., Dobbins, S., Barclay, E., Gorman, M., Martin, L., Kovac, M. B., Humphray, S., Consortium, C., Consortium, W. G. S., Lucassen, A., Holmes, C. C., Bentley, D., Donnelly, P., Taylor, J., Petridis, C., Roylance, R., Sawyer, E. J., Kerr, D. J., Clark, S., Grimes, J., Kearsey, S. E., Thomas, H. J., McVean, G., Houlston, R. S. & Tomlinson, I. (2013) Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet*, 45(2), 136-44.

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A. & Liu, D. R. (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*, 31(9), 839-43.

Peltier, H. J. & Latham, G. J. (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA*, 14(5), 844-52.

Pennisi, E. (1998) How a growth control path takes a wrong turn to cancer. *Science*, 281(5382), 1438-9, 1441.

Perleberg, C., Kind, A. & Schnieke, A. (2018) Genetically engineered pigs as models for human disease. *Dis Model Mech*, 11(1).

Pfaffeneder, T., Hackner, B., Truss, M., Munzel, M., Muller, M., Deiml, C. A., Hagemeier, C. & Carell, T. (2011) The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew Chem Int Ed Engl*, 50(31), 7008-12.

Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. (2017) Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods*, 14(7), 687-690.

Platt, R. J., Chen, S., Zhou, Y., Yim, M. J., Swiech, L., Kempton, H. R., Dahlman, J. E., Parnas, O., Eisenhaure, T. M., Jovanovic, M., Graham, D. B., Jhunjhunwala, S., Heidenreich, M., Xavier, R. J., Langer, R., Anderson, D. G., Hacohen, N., Regev, A., Feng, G., Sharp, P. A. & Zhang, F. (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell*, 159(2), 440-55.

Polakis, P. (1997) The adenomatous polyposis coli (APC) tumor suppressor. *Biochim Biophys Acta*, 1332(3), F127-47.

Quandt, K., Frech, K., Karas, H., Wingender, E. & Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res*, 23(23), 4878-84.

Ramaswami, G., Lin, W., Piskol, R., Tan, M. H., Davis, C. & Li, J. B. (2012) Accurate identification of human Alu and non-Alu RNA editing sites. *Nat Methods*, 9(6), 579-81.

Rampino, N., Yamamoto, H., Ionov, Y., Li, Y., Sawai, H., Reed, J. C. & Perucho, M. (1997) Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. *Science*, 275(5302), 967-9.

Roach, J. C., Boysen, C., Wang, K. & Hood, L. (1995) Pairwise end sequencing: a unified approach to genomic mapping and sequencing. *Genomics*, 26(2), 345-53.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. (2011) Integrative genomics viewer. *Nat Biotechnol*, 29(1), 24-6.

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139-40.

Rojas, A., Meherem, S., Kim, Y. H., Washington, M. K., Willis, J. E., Markowitz, S. D. & Grady, W. M. (2008) The aberrant methylation of TSP1 suppresses TGF-beta1 activation in colorectal cancer. *Int J Cancer*, 123(1), 14-21.

Rokavec, M., Li, H., Jiang, L. & Hermeking, H. (2014) The p53/miR-34 axis in development and disease. *J Mol Cell Biol*, 6(3), 214-30.

Ronaghi, M., Uhlen, M. & Nyren, P. (1998) A sequencing method based on real-time pyrophosphate. *Science*, 281(5375), 363, 365.

Roodink, I., Verrijp, K., Raats, J. & Leenders, W. P. (2009) Plexin D1 is ubiquitously expressed on tumor vessels and tumor cells in solid malignancies. *BMC Cancer*, 9, 297.

Roose, J. & Clevers, H. (1999) TCF transcription factors: molecular switches in carcinogenesis. *Biochim Biophys Acta*, 1424(2-3), M23-37.

Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. & Polakis, P. (1996) Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. *Science*, 272(5264), 1023-6.

Saalfrank, A., Janssen, K. P., Ravon, M., Flisikowski, K., Eser, S., Steiger, K., Flisikowska, T., Muller-Fliedner, P., Schulze, E., Bronner, C., Gnann, A., Kappe, E., Bohm, B., Schade, B., Certa, U., Saur, D., Esposito, I., Kind, A. & Schnieke, A. (2016) A porcine model of osteosarcoma. *Oncogenesis*, 5, e210.

Sachs, D. H. (1994) The pig as a potential xenograft donor. *Vet Immunol Immunopathol*, 43(1-3), 185-91.

Samaei, N. M., Yazdani, Y., Alizadeh-Navaei, R., Azadeh, H. & Farazmandfar, T. (2014) Promoter methylation analysis of WNT/beta-catenin pathway regulators and its association with expression of DNMT1 enzyme in colorectal cancer. *J Biomed Sci*, 21, 73.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74(12), 5463-7.

Sarver, A. L., French, A. J., Borralho, P. M., Thayanithy, V., Oberg, A. L., Silverstein, K. A., Morlan, B. W., Riska, S. M., Boardman, L. A., Cunningham, J. M., Subramanian, S., Wang, L., Smyrk, T. C.,

Rodrigues, C. M., Thibodeau, S. N. & Steer, C. J. (2009) Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states. *BMC Cancer*, 9, 401.

Sauer, B. (1996) Multiplex Cre/lox recombination permits selective site-specific DNA targeting to both a natural and an engineered site in the yeast genome. *Nucleic Acids Res*, 24(23), 4608-13.

Schmitz, K. J., Hey, S., Schinwald, A., Wohlschlaeger, J., Baba, H. A., Worm, K. & Schmid, K. W. (2009) Differential expression of microRNA 181b and microRNA 21 in hyperplastic polyps and sessile serrated adenomas of the colon. *Virchows Arch*, 455(1), 49-54.

Schnieke, A. E., Kind, A. J., Ritchie, W. A., Mycock, K., Scott, A. R., Ritchie, M., Wilmut, I., Colman, A. & Campbell, K. H. (1997) Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. *Science*, 278(5346), 2130-3.

Schonhuber, N., Seidler, B., Schuck, K., Veltkamp, C., Schachtler, C., Zukowska, M., Eser, S., Feyerabend, T. B., Paul, M. C., Eser, P., Klein, S., Lowy, A. M., Banerjee, R., Yang, F., Lee, C. L., Moding, E. J., Kirsch, D. G., Scheideler, A., Alessi, D. R., Varela, I., Bradley, A., Kind, A., Schnieke, A. E., Rodewald, H. R., Rad, R., Schmid, R. M., Schneider, G. & Saur, D. (2014) A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nat Med*, 20(11), 1340-7.

Schook, L. B., Collares, T. V., Hu, W., Liang, Y., Rodrigues, F. M., Rund, L. A., Schachtschneider, K. M., Seixas, F. K., Singh, K., Wells, K. D., Walters, E. M., Prather, R. S. & Counter, C. M. (2015) A Genetic Porcine Model of Cancer. *PLoS One*, 10(7), e0128864.

Schreuders, E. H., Grobbee, E. J., Spaander, M. C. & Kuipers, E. J. (2016) Advances in Fecal Tests for Colorectal Cancer Screening. *Curr Treat Options Gastroenterol*, 14(1), 152-62.

Schubert, R., Frank, F., Nagelmann, N., Liebsch, L., Schuldenzucker, V., Schramke, S., Wirsig, M., Johnson, H., Kim, E. Y., Ott, S., Holzner, E., Demokritov, S. O., Motlik, J., Faber, C. & Reilmann, R. (2016) Neuroimaging of a minipig model of Huntington's disease: Feasibility of volumetric, diffusion-weighted and spectroscopic assessments. *J Neurosci Methods*, 265, 46-55.

Segditsas, S., Sieber, O. M., Rowan, A., Setien, F., Neale, K., Phillips, R. K., Ward, R., Esteller, M. & Tomlinson, I. P. (2008) Promoter hypermethylation leads to decreased APC mRNA expression in familial polyposis and sporadic colorectal tumours, but does not substitute for truncating mutations. *Exp Mol Pathol*, 85(3), 201-6.

Serrati, S., De Summa, S., Pilato, B., Petriella, D., Lacalamita, R., Tommasi, S. & Pinto, R. (2016) Next-generation sequencing: advances and applications in cancer diagnosis. *Onco Targets Ther*, 9, 7355-7365.

Shalapour, S., Deiser, K., Kuhl, A. A., Glauben, R., Krug, S. M., Fischer, A., Sercan, O., Chappaz, S., Bereswill, S., Heimesaat, M. M., Loddenkemper, C., Fromm, M., Finke, D., Hammerling, G. J., Arnold, B., Siegmund, B. & Schuler, T. (2012) Interleukin-7 links T lymphocyte and intestinal epithelial cell homeostasis. *PLoS One*, 7(2), e31939.

Shen, L., Kondo, Y., Hamilton, S. R., Rashid, A. & Issa, J. P. (2003) P14 methylation in human colon cancer is associated with microsatellite instability and wild-type p53. *Gastroenterology*, 124(3), 626-33.

Shima, K., Nosho, K., Baba, Y., Cantor, M., Meyerhardt, J. A., Giovannucci, E. L., Fuchs, C. S. & Ogino, S. (2011) Prognostic significance of CDKN2A (p16) promoter methylation and loss of expression in 902 colorectal cancers: Cohort study and literature review. *Int J Cancer*, 128(5), 1080-94.

Shimizu, Y., Ikeda, S., Fujimori, M., Kodama, S., Nakahara, M., Okajima, M. & Asahara, T. (2002) Frequent alterations in the Wnt signaling pathway in colorectal cancer with microsatellite instability. *Genes Chromosomes Cancer*, 33(1), 73-81.

Shirafkan, N., Mansoori, B., Mohammadi, A., Shomali, N., Ghasbi, M. & Baradaran, B. (2018) MicroRNAs as novel biomarkers for colorectal cancer: New outlooks. *Biomed Pharmacother*, 97, 1319-1330.

Shussman, N. & Wexner, S. D. (2014) Colorectal polyps and polyposis syndromes. *Gastroenterol Rep (Oxf)*, 2(1), 1-15.

Sieren, J. C., Meyerholz, D. K., Wang, X. J., Davis, B. T., Newell, J. D., Jr., Hammond, E., Rohret, J. A., Rohret, F. A., Struzynski, J. T., Goeken, J. A., Naumann, P. W., Leidinger, M. R., Taghiyev, A., Van Rheeden, R., Hagen, J., Darbro, B. W., Quelle, D. E. & Rogers, C. S. (2014) Development and translational imaging of a TP53 porcine tumorigenesis model. *J Clin Invest*, 124(9), 4052-66.

Skarnes, W. C., Rosen, B., West, A. P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A. O., Thomas, M., Harrow, J., Cox, T., Jackson, D., Severin, J., Biggs, P., Fu, J., Nefedov, M., de Jong, P. J., Stewart, A. F. & Bradley, A. (2011) A conditional knockout resource for the genome-wide study of mouse gene function. *Nature*, 474(7351), 337-42.

Slaby, O., Svoboda, M., Fabian, P., Smerdova, T., Knoflickova, D., Bednarikova, M., Nenutil, R. & Vyzula, R. (2007) Altered Expression of miR-21, miR-31, miR-143 and miR-145 Is Related to Clinicopathologic Features of Colorectal Cancer. *Oncology*, 72(5-6), 397-402.

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X. & Zhang, F. (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science*, 351(6268), 84-8.

Smalley, M. J., Sara, E., Paterson, H., Naylor, S., Cook, D., Jayatilake, H., Fryer, L. G., Hutchinson, L., Fry, M. J. & Dale, T. C. (1999) Interaction of axin and Dvl-2 proteins regulates Dvl-2-stimulated TCF-dependent transcription. *EMBO J*, 18(10), 2823-35.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. & Hood, L. E. (1986) Fluorescence detection in automated DNA sequence analysis. *Nature*, 321(6071), 674-9.

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. (1985) Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature*, 317(6034), 230-4.

Solomon, E., Voss, R., Hall, V., Bodmer, W. F., Jass, J. R., Jeffreys, A. J., Lucibello, F. C., Patel, I. & Rider, S. H. (1987) Chromosome 5 allele loss in human colorectal carcinomas. *Nature*, 328(6131), 616-9.

Song, L. L. & Li, Y. M. (2016) Current noninvasive tests for colorectal cancer screening: An overview of colorectal cancer screening tests. *World J Gastrointest Oncol*, 8(11), 793-800.

Stachowiak, M., Flisikowska, T., Bauersachs, S., Perleberg, C., Pausch, H., Switonski, M., Kind, A., Saur, D., Schnieke, A. & Flisikowski, K. (2017) Altered microRNA profiles during early colon adenoma progression in a porcine model of familial adenomatous polyposis. *Oncotarget*, 8(56), 96154-96160.

Staden, R. (1979) A strategy of DNA sequencing employing computer programs. *Nucleic Acids Res*, 6(7), 2601-10.

Steines, B., Dickey, D. D., Bergen, J., Excoffon, K. J., Weinstein, J. R., Li, X., Yan, Z., Abou Alaiwa, M. H., Shah, V. S., Bouzek, D. C., Powers, L. S., Gansemer, N. D., Ostedgaard, L. S., Engelhardt, J. F., Stoltz, D. A., Welsh, M. J., Sinn, P. L., Schaffer, D. V. & Zabner, J. (2016) CFTR gene transfer with AAV improves early cystic fibrosis pig phenotypes. *JCI Insight*, 1(14), e88728.

Strillacci, A., Valerii, M. C., Sansone, P., Caggiano, C., Sgromo, A., Vittori, L., Fiorentino, M., Poggioli, G., Rizzello, F., Campieri, M. & Spisni, E. (2013) Loss of miR-101 expression promotes Wnt/beta-catenin signalling pathway activation and malignancy in colon cancer cells. *J Pathol*, 229(3), 379-89.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. & Mesirov, J. P. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, 102(43), 15545-50.

Suh, E. R., Ha, C. S., Rankin, E. B., Toyota, M. & Traber, P. G. (2002) DNA methylation down-regulates CDX1 gene expression in colorectal cancer cell lines. *J Biol Chem*, 277(39), 35795-800.

Sun, D., Wang, C., Long, S., Ma, Y., Guo, Y., Huang, Z., Chen, X., Zhang, C., Chen, J. & Zhang, J. (2015) C/EBP-beta-activated microRNA-223 promotes tumour growth through targeting RASA1 in human colorectal cancer. *Br J Cancer*, 112(9), 1491-500.

Susswein, L. R., Marshall, M. L., Nusbaum, R., Vogel Postula, K. J., Weissman, S. M., Yackowski, L., Vaccari, E. M., Bissonnette, J., Booker, J. K., Cremona, M. L., Gibellini, F., Murphy, P. D., Pineda-Alvarez, D. E., Pollevick, G. D., Xu, Z., Richard, G., Bale, S., Klein, R. T., Hruska, K. S. & Chung, W. K. (2016) Pathogenic and likely pathogenic variant prevalence among the first 10,000 patients referred for next-generation cancer panel testing. *Genet Med*, 18(8), 823-32.

Suzuki, H., Igarashi, S., Nojima, M., Maruyama, R., Yamamoto, E., Kai, M., Akashi, H., Watanabe, Y., Yamamoto, H., Sasaki, Y., Itoh, F., Imai, K., Sugai, T., Shen, L., Issa, J. P., Shinomura, Y., Tokino, T. & Toyota, M. (2010) IGFBP7 is a p53-responsive gene specifically silenced in colorectal cancer with CpG island methylator phenotype. *Carcinogenesis*, 31(3), 342-9.

Suzuki, H., Watkins, D. N., Jair, K. W., Schuebel, K. E., Markowitz, S. D., Chen, W. D., Pretlow, T. P., Yang, B., Akiyama, Y., Van Engeland, M., Toyota, M., Tokino, T., Hinoda, Y., Imai, K., Herman, J. G. & Baylin, S. B. (2004) Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nat Genet*, 36(4), 417-22.

Swindle, M. M., Smith, A. C. & Hepburn, B. J. (1988) Swine as models in experimental surgery. *J Invest Surg*, 1(1), 65-79.

Szpakowski, S., Sun, X., Lage, J. M., Dyer, A., Rubinstein, J., Kowalski, D., Sasaki, C., Costa, J. & Lizardi, P. M. (2009) Loss of epigenetic silencing in tumors preferentially affects primate-specific retroelements. *Gene*, 448(2), 151-67.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. & Rao, A. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324(5929), 930-5.

Takagi, H., Sasaki, S., Suzuki, H., Toyota, M., Maruyama, R., Nojima, M., Yamamoto, H., Omata, M., Tokino, T., Imai, K. & Shinomura, Y. (2008) Frequent epigenetic inactivation of SFRP genes in hepatocellular carcinoma. *J Gastroenterol*, 43(5), 378-89.

Takayama, T., Miyanishi, K., Hayashi, T., Sato, Y. & Niitsu, Y. (2006) Colorectal cancer: genetics of development and metastasis. *J Gastroenterol*, 41(3), 185-92.

Talseth-Palmer, B. A. (2017) The genetic basis of colonic adenomatous polyposis syndromes. *Hered Cancer Clin Pract*, 15, 5.

Tan, J., Lee, P. L., Li, Z., Jiang, X., Lim, Y. C., Hooi, S. C. & Yu, Q. (2010) B55beta-associated PP2A complex controls PDK1-directed myc signaling and modulates rapamycin sensitivity in colorectal cancer. *Cancer Cell*, 18(5), 459-71.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B. & Zhao, Y. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, 146(6), 1016-28.

Tan, W., Carlson, D. F., Lancto, C. A., Garbe, J. R., Webster, D. A., Hackett, P. B. & Fahrenkrug, S. C. (2013) Efficient nonmeiotic allele introgression in livestock using custom endonucleases. *Proc Natl Acad Sci U S A*.

Tetteh, P. W., Kretzschmar, K., Begthel, H., van den Born, M., Korving, J., Morsink, F., Farin, H., van Es, J. H., Offerhaus, G. J. & Clevers, H. (2016) Generation of an inducible colon-specific Cre enzyme mouse line for colon cancer research. *Proc Natl Acad Sci U S A*, 113(42), 11859-11864.

Thomas, K. R. & Capecchi, M. R. (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell*, 51(3), 503-12.

Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*, 14(2), 178-92.

Tomlinson, C., Wong, C., Au, H. J. & Schiller, D. (2012) Factors associated with delays to medical assessment and diagnosis for patients with colorectal cancer. *Can Fam Physician*, 58(9), e495-501.

Tomlinson, I., Webb, E., Carvajal-Carmona, L., Broderick, P., Kemp, Z., Spain, S., Penegar, S., Chandler, I., Gorman, M., Wood, W., Barclay, E., Lubbe, S., Martin, L., Sellick, G., Jaeger, E., Hubner, R., Wild, R., Rowan, A., Fielding, S., Howarth, K., Consortium, C., Silver, A., Atkin, W., Muir, K., Logan, R., Kerr, D., Johnstone, E., Sieber, O., Gray, R., Thomas, H., Peto, J., Cazier, J. B. & Houlston, R. (2007) A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet*, 39(8), 984-8.

Tomlinson, I. P., Webb, E., Carvajal-Carmona, L., Broderick, P., Howarth, K., Pittman, A. M., Spain, S., Lubbe, S., Walther, A., Sullivan, K., Jaeger, E., Fielding, S., Rowan, A., Vijayakrishnan, J., Domingo, E., Chandler, I., Kemp, Z., Qureshi, M., Farrington, S. M., Tenesa, A., Prendergast, J. G., Barnetson, R. A., Penegar, S., Barclay, E., Wood, W., Martin, L., Gorman, M., Thomas, H., Peto, J., Bishop, D. T., Gray, R., Maher, E. R., Lucassen, A., Kerr, D., Evans, D. G., Consortium, C., Schafmayer, C., Buch, S., Volzke, H., Hampe, J., Schreiber, S., John, U., Koessler, T., Pharoah, P., van Wezel, T., Morreau, H., Wijnen, J. T., Hopper, J. L., Southey, M. C., Giles, G. G., Severi, G., Castellvi-Bel, S., Ruiz-Ponte, C., Carracedo, A., Castells, A., Consortium, E., Forsti, A., Hemminki, K., Vodicka, P., Naccarati, A., Lipton, L., Ho, J. W., Cheng, K. K., Sham, P. C., Luk, J., Agundez, J. A., Ladero, J. M., de la Hoya, M., Caldes, T., Niittymaki, I., Tuupanen, S., Karhu, A., Aaltonen, L., Cazier, J. B., Campbell, H., Dunlop, M. G. & Houlston, R. S. (2008) A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet*, 40(5), 623-30.

Toyota, M., Suzuki, H., Sasaki, Y., Maruyama, R., Imai, K., Shinomura, Y. & Tokino, T. (2008) Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. *Cancer Res*, 68(11), 4123-32.

Truong, D. J., Kuhner, K., Kuhn, R., Werfel, S., Engelhardt, S., Wurst, W. & Ortiz, O. (2015) Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res*, 43(13), 6450-8.

Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A. J., Le, L. P., Aryee, M. J. & Joung, J. K. (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*, 33(2), 187-197.

Tse, J. W. T., Jenkins, L. J., Chionh, F. & Mariadason, J. M. (2017) Aberrant DNA Methylation in Colorectal Cancer: What Should We Target? *Trends Cancer*, 3(10), 698-712.

Tseng, C. H., Murray, K. D., Jou, M. F., Hsu, S. M., Cheng, H. J. & Huang, P. H. (2011) Sema3E/plexin-D1 mediated epithelial-to-mesenchymal transition in ovarian endometrioid cancer. *PLoS One*, 6(4), e19396.

Tsunoda, T., Takashima, Y., Yoshida, Y., Doi, K., Tanaka, Y., Fujimoto, T., Machida, T., Ota, T., Koyanagi, M., Kuroki, M., Sasazuki, T. & Shirasawa, S. (2011) Oncogenic KRAS regulates miR-200c and miR-221/222 in a 3D-specific manner in colorectal cancer cells. *Anticancer Res*, 31(7), 2453-9.

Tuupanen, S., Niittymaki, I., Nousiainen, K., Vanharanta, S., Mecklin, J. P., Nuorva, K., Jarvinen, H., Hautaniemi, S., Karhu, A. & Aaltonen, L. A. (2008) Allelic imbalance at rs6983267 suggests selection of the risk allele in somatic colorectal tumor evolution. *Cancer Res*, 68(1), 14-7.

Uchida, M., Shimatsu, Y., Onoe, K., Matsuyama, N., Niki, R., Ikeda, J. E. & Imai, H. (2001) Production of transgenic miniature pigs by pronuclear microinjection. *Transgenic Res*, 10(6), 577-82.

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. (2010) Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*, 11(9), 636-46.

Valeri, N., Gasparini, P., Fabbri, M., Braconi, C., Veronese, A., Lovat, F., Adair, B., Vannini, I., Fanini, F., Bottoni, A., Costinean, S., Sandhu, S. K., Nuovo, G. J., Alder, H., Gafa, R., Calore, F., Ferracin, M., Lanza, G., Volinia, S., Negrini, M., McIlhatton, M. A., Amadori, D., Fishel, R. & Croce, C. M. (2010)

Modulation of mismatch repair and genomic stability by miR-155. *Proc Natl Acad Sci U S A*, 107(15), 6982-7.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*, 3(7), RESEARCH0034.

Veeck, J., Geisler, C., Noetzel, E., Alkaya, S., Hartmann, A., Knuchel, R. & Dahl, E. (2008) Epigenetic inactivation of the secreted frizzled-related protein-5 (SFRP5) gene in human breast cancer is associated with unfavorable prognosis. *Carcinogenesis*, 29(5), 991-8.

Veigl, M. L., Kasturi, L., Olechnowicz, J., Ma, A. H., Lutterbaugh, J. D., Periyasamy, S., Li, G. M., Drummond, J., Modrich, P. L., Sedwick, W. D. & Markowitz, S. D. (1998) Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proc Natl Acad Sci U S A*, 95(15), 8698-702.

Venter, J. C. (2003) A part of the human genome sequence. *Science*, 299(5610), 1183-4.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., et al (2001) The sequence of the human genome. *Science*, 291(5507), 1304-51.

Vlachos, I. S., Paraskevopoulou, M. D., Karagkouni, D., Georgakilas, G., Vergoulis, T., Kanellos, I., Anastasopoulos, I. L., Maniou, S., Karathanou, K., Kalfakakou, D., Fevgas, A., Dalamagas, T. & Hatzigeorgiou, A. G. (2015a) DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Res*, 43(Database issue), D153-9.

Vlachos, I. S., Zagganas, K., Paraskevopoulou, M. D., Georgakilas, G., Karagkouni, D., Vergoulis, T., Dalamagas, T. & Hatzigeorgiou, A. G. (2015b) DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucleic Acids Res*, 43(W1), W460-6.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr. & Kinzler, K. W. (2013) Cancer genome landscapes. *Science*, 339(6127), 1546-58.

Vogt, M., Munding, J., Gruner, M., Liffers, S. T., Verdoodt, B., Hauk, J., Steinstraesser, L., Tannapfel, A. & Hermeking, H. (2011) Frequent concomitant inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal, pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft tissue sarcomas. *Virchows Arch*, 458(3), 313-22.

Volinia, S., Calin, G. A., Liu, C. G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R. L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C. C. & Croce, C. M. (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A*, 103(7), 2257-61.

Vychytilova-Faltejskova, P., Merhautova, J., Machackova, T., Gutierrez-Garcia, I., Garcia-Solano, J., Radova, L., Brchnelova, D., Slaba, K., Svoboda, M., Halamkova, J., Demlova, R., Kiss, I., Vyzula, R., Conesa-Zamora, P. & Slaby, O. (2017) MiR-215-5p is a tumor suppressor in colorectal cancer targeting EGFR ligand epiregulin and its transcriptional inducer HOXB9. *Oncogenesis*, 6(11), 399.

Wang, K., Jin, Q., Ruan, D., Yang, Y., Liu, Q., Wu, H., Zhou, Z., Ouyang, Z., Liu, Z., Zhao, Y., Zhao, B., Zhang, Q., Peng, J., Lai, C., Fan, N., Liang, Y., Lan, T., Li, N., Wang, X., Wang, X., Fan, Y., Doevendans, P.

A., Sluijter, J. P. G., Liu, P., Li, X. & Lai, L. (2017) Cre-dependent Cas9-expressing pigs enable efficient in vivo genome editing. *Genome Res*, 27(12), 2061-2071.

Wang, Y., Du, Y., Shen, B., Zhou, X., Li, J., Liu, Y., Wang, J., Zhou, J., Hu, B., Kang, N., Gao, J., Yu, L., Huang, X. & Wei, H. (2015) Efficient generation of gene-modified pigs via injection of zygote with Cas9/sgRNA. *Sci Rep*, 5, 8256.

Wei, C., Liu, J., Yu, Z., Zhang, B., Gao, G. & Jiao, R. (2013) TALEN or Cas9 - Rapid, Efficient and Specific Choices for Genome Modifications. *J Genet Genomics*, 40(6), 281-9.

Weinberg, R. A. (2007) *The biology of cancer*, 1 vols. New York: Garland Science.

Wertheim, B. C., Smith, J. W., Fang, C., Alberts, D. S., Lance, P. & Thompson, P. A. (2012) Risk modification of colorectal adenoma by CYP7A1 polymorphisms and the role of bile acid metabolism in carcinogenesis. *Cancer Prev Res (Phila)*, 5(2), 197-204.

Whiffin, N., Hosking, F. J., Farrington, S. M., Palles, C., Dobbins, S. E., Zgaga, L., Lloyd, A., Kinnersley, B., Gorman, M., Tenesa, A., Broderick, P., Wang, Y., Barclay, E., Hayward, C., Martin, L., Buchanan, D. D., Win, A. K., Hopper, J., Jenkins, M., Lindor, N. M., Newcomb, P. A., Gallinger, S., Conti, D., Schumacher, F., Casey, G., Liu, T., Swedish Low-Risk Colorectal Cancer Study, G., Campbell, H., Lindblom, A., Houlston, R. S., Tomlinson, I. P. & Dunlop, M. G. (2014) Identification of susceptibility loci for colorectal cancer in a genome-wide meta-analysis. *Hum Mol Genet*, 23(17), 4729-37.

Whitelaw, C. B., Radcliffe, P. A., Ritchie, W. A., Carlisle, A., Ellard, F. M., Pena, R. N., Rowe, J., Clark, A. J., King, T. J. & Mitrophanous, K. A. (2004) Efficient generation of transgenic pigs using equine infectious anaemia virus (EIAV) derived vector. *FEBS Lett*, 571(1-3), 233-6.

Whitworth, K. M., Lee, K., Benne, J. A., Beaton, B. P., Spate, L. D., Murphy, S. L., Samuel, M. S., Mao, J., O'Gorman, C., Walters, E. M., Murphy, C. N., Driver, J., Mileham, A., McLaren, D., Wells, K. D. & Prather, R. S. (2014) Use of the CRISPR/Cas9 system to produce genetically engineered pigs from in vitro-derived oocytes and embryos. *Biol Reprod*, 91(3), 78.

Willert, K., Shibamoto, S. & Nusse, R. (1999) Wnt-induced dephosphorylation of axin releases beta-catenin from the axin complex. *Genes Dev*, 13(14), 1768-73.

Williams, C. R., Baccarella, A., Parrish, J. Z. & Kim, C. C. (2016) Trimming of sequence reads alters RNA-Seq gene expression estimates. *BMC Bioinformatics*, 17, 103.

Wolff, E. M., Byun, H. M., Han, H. F., Sharma, S., Nichols, P. W., Siegmund, K. D., Yang, A. S., Jones, P. A. & Liang, G. (2010) Hypomethylation of a LINE-1 promoter activates an alternate transcript of the MET oncogene in bladders with cancer. *PLoS Genet*, 6(4), e1000917.

Wong, N. A., Britton, M. P., Choi, G. S., Stanton, T. K., Bicknell, D. C., Wilding, J. L. & Bodmer, W. F. (2004) Loss of CDX1 expression in colorectal carcinoma: promoter methylation, mutation, and loss of heterozygosity analyses of 37 cell lines. *Proc Natl Acad Sci U S A*, 101(2), 574-9.

Wood, L. D., Parsons, D. W., Jones, S., Lin, J., Sjoblom, T., Leary, R. J., Shen, D., Boca, S. M., Barber, T., Ptak, J., Silliman, N., Szabo, S., Dezso, Z., Ustyanksky, V., Nikolskaya, T., Nikolsky, Y., Karchin, R., Wilson, P. A., Kaminker, J. S., Zhang, Z., Croshaw, R., Willis, J., Dawson, D., Shipitsin, M., Willson, J. K., Sukumar, S., Polyak, K., Park, B. H., Pethiyagoda, C. L., Pant, P. V., Ballinger, D. G., Sparks, A. B., Hartigan, J., Smith, D. R., Suh, E., Papadopoulos, N., Buckhaults, P., Markowitz, S. D., Parmigiani, G., Kinzler, K. W., Velculescu, V. E. & Vogelstein, B. (2007) The genomic landscapes of human breast and colorectal cancers. *Science*, 318(5853), 1108-13.

World health statistics 2017: Monitoring health for the SDGs, Sustainable Development Goals. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO; http://apps.who.int/iris/bitstream/handle/10665/255336/9789241565486-eng.pdf;jsessionid=73C27896EAA6A4A17FDF776E162EE0FA?sequence=1, retrieved 02.09.2018

Wu, M., Wei, C., Lian, Z., Liu, R., Zhu, C., Wang, H., Cao, J., Shen, Y., Zhao, F., Zhang, L., Mu, Z., Wang, Y., Wang, X., Du, L. & Wang, C. (2016) Rosa26-targeted sheep gene knock-in via CRISPR-Cas9 system. *Sci Rep*, 6, 24360.

Wu, W., Yang, J., Feng, X., Wang, H., Ye, S., Yang, P., Tan, W., Wei, G. & Zhou, Y. (2013) MicroRNA-32 (miR-32) regulates phosphatase and tensin homologue (PTEN) expression and promotes growth, migration, and invasion in colorectal carcinoma cells. *Mol Cancer*, 12, 30.

Wu, W. K., Law, P. T., Lee, C. W., Cho, C. H., Fan, D., Wu, K., Yu, J. & Sung, J. J. (2011) MicroRNA in colorectal cancer: from benchtop to bedside. *Carcinogenesis*, 32(3), 247-53.

Xu, X., Chen, R., Li, Z., Huang, N., Wu, X., Li, S., Li, Y. & Wu, S. (2015) MicroRNA-490-3p inhibits colorectal cancer metastasis by targeting TGFbetaR1. *BMC Cancer*, 15, 1023.

Yamada, N., Noguchi, S., Mori, T., Naoe, T., Maruo, K. & Akao, Y. (2013) Tumor-suppressive microRNA-145 targets catenin delta-1 to regulate Wnt/beta-catenin signaling in human colon cancer cells. *Cancer Lett*, 335(2), 332-42.

Yamakawa, H., Nagai, T., Harasawa, R., Yamagami, T., Takahashi, J., Ishikawa, K., Nomura, N. & Nagashima, H. (1999) Production of transgenic pig carrying MMTV/v-Ha-ras. *Journal of Reproduction and Development*, 45, 111-118.

Yamakuchi, M., Ferlito, M. & Lowenstein, C. J. (2008) miR-34a repression of SIRT1 regulates apoptosis. *Proc Natl Acad Sci U S A*, 105(36), 13421-6.

Yamamoto, H., Gil, J., Schwartz, S., Jr. & Perucho, M. (2000) Frameshift mutations in Fas, Apaf-1, and Bcl-10 in gastro-intestinal cancer of the microsatellite mutator phenotype. *Cell Death Differ*, 7(2), 238-9.

Yamamoto, H., Kishida, S., Kishida, M., Ikeda, S., Takada, S. & Kikuchi, A. (1999) Phosphorylation of axin, a Wnt signal negative regulator, by glycogen synthase kinase-3beta regulates its stability. *J Biol Chem*, 274(16), 10681-4.

Yang, S., Farraye, F. A., Mack, C., Posnik, O. & O'Brien, M. J. (2004) BRAF and KRAS Mutations in hyperplastic polyps and serrated adenomas of the colorectum: relationship to histology and CpG island methylation status. *Am J Surg Pathol*, 28(11), 1452-9.

Yang, X., Zou, J., Cai, H., Huang, X., Yang, X., Guo, D. & Cao, Y. (2017) Ginsenoside Rg3 inhibits colorectal tumor growth via down-regulation of C/EBPbeta/NF-kappaB signaling. *Biomed Pharmacother*, 96, 1240-1245.

Yu, H., Lee, H., Herrmann, A., Buettner, R. & Jove, R. (2014) Revisiting STAT3 signalling in cancer: new and unexpected biological functions. *Nat Rev Cancer*, 14(11), 736-46.

Yu, H. H., Zhao, H., Qing, Y. B., Pan, W. R., Jia, B. Y., Zhao, H. Y., Huang, X. X. & Wei, H. J. (2016) Porcine Zygote Injection with Cas9/sgRNA Results in DMD-Modified Pig with Muscle Dystrophy. *Int J Mol Sci*, 17(10).

Yu, Y., Kanwar, S. S., Patel, B. B., Oh, P. S., Nautiyal, J., Sarkar, F. H. & Majumdar, A. P. (2012) MicroRNA-21 induces stemness by downregulating transforming growth factor beta receptor 2 (TGFbetaR2) in colon cancer cells. *Carcinogenesis*, 33(1), 68-76.

Zambrowicz, B. P., Imamoto, A., Fiering, S., Herzenberg, L. A., Kerr, W. G. & Soriano, P. (1997) Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A*, 94(8), 3789-94.

Zang, Y., Wang, T., Pan, J. & Gao, F. (2017) miR-215 promotes cell migration and invasion of gastric cancer cell lines by targeting FOXO1. *Neoplasma*, 64(4), 579-587.

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V. & Zhang, F. (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 163(3), 759-71.

Zhang, G., Zhou, H., Xiao, H., Liu, Z., Tian, H. & Zhou, T. (2014a) MicroRNA-92a functions as an oncogene in colorectal cancer by targeting PTEN. *Dig Dis Sci*, 59(1), 98-107.

Zhang, W., Zhang, T., Jin, R., Zhao, H., Hu, J., Feng, B., Zang, L., Zheng, M. & Wang, M. (2014b) MicroRNA-301a promotes migration and invasion by targeting TGFBR2 in human colorectal cancer. *J Exp Clin Cancer Res*, 33, 113.

Zhang, X., Li, X., Tan, F., Yu, N. & Pei, H. (2017) STAT1 Inhibits MiR-181a Expression to Suppress Colorectal Cancer Cell Proliferation Through PTEN/Akt. *J Cell Biochem*, 118(10), 3435-3443.

Zhang, Y., Lin, C., Liao, G., Liu, S., Ding, J., Tang, F., Wang, Z., Liang, X., Li, B., Wei, Y., Huang, Q., Li, X. & Tang, B. (2015) MicroRNA-506 suppresses tumor proliferation and metastasis in colon cancer by directly targeting the oncogene EZH2. *Oncotarget*, 6(32), 32586-601.

Zhang, Y., Tian, X., Ji, H., Guan, X., Xu, W., Dong, B., Zhao, M., Wei, M., Ye, C., Sun, Y., Yuan, X., Yang, C. & Hao, C. (2014c) Expression of SATB1 promotes the growth and metastasis of colorectal cancer. *PLoS One*, 9(6), e100413.

Zhou, X., Wang, L., Du, Y., Xie, F., Li, L., Liu, Y., Liu, C., Wang, S., Zhang, S., Huang, X., Wang, Y. & Wei, H. (2016) Efficient Generation of Gene-Modified Pigs Harboring Precise Orthologous Human Mutation via CRISPR/Cas9-Induced Homology-Directed Repair in Zygotes. *Hum Mutat*, 37(1), 110-8.

Zhu, J., Chen, L., Zou, L., Yang, P., Wu, R., Mao, Y., Zhou, H., Li, R., Wang, K., Wang, W., Hua, D. & Zhang, X. (2014) MiR-20b, -21, and -130b inhibit PTEN expression resulting in B7-H1 over-expression in advanced colorectal cancer. *Hum Immunol*, 75(4), 348-53.

10. Appendix

10.1 Wizard® SV Gel and PCR Clean-Up System

Wizard® SV Gel and PCR Clean-Up System

INSTRUCTIONS FOR USE OF PRODUCTS A9280, A9281, A9282, AND A9285.

Quick PROTOCOL

DNA Purification by Centrifugation

Gel Slice and PCR Product Preparation

A. Dissolving the Gel Slice

1. Following electrophoresis, excise DNA band from gel and place gel slice in a 1.5ml microcentrifuge tube.
2. Add 10µl Membrane Binding Solution per 10mg of gel slice. Vortex and incubate at 50–65°C until gel slice is completely dissolved.

B. Processing PCR Amplifications

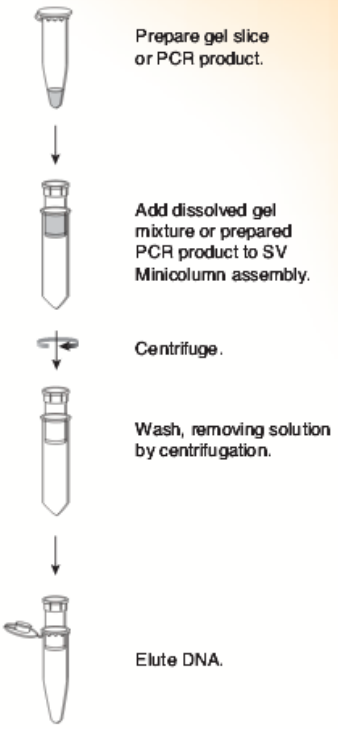
1. Add an equal volume of Membrane Binding Solution to the PCR amplification.

Binding of DNA

1. Insert SV Minicolumn into Collection Tube.
2. Transfer dissolved gel mixture or prepared PCR product to the Minicolumn assembly. Incubate at room temperature for 1 minute.
3. Centrifuge at 16,000 × *g* for 1 minute. Discard flowthrough and reinsert Minicolumn into Collection Tube.

Washing

4. Add 700µl Membrane Wash Solution (ethanol added). Centrifuge at 16,000 × *g* for 1 minute. Discard flowthrough and reinsert Minicolumn into Collection Tube.
5. Repeat Step 4 with 500µl Membrane Wash Solution. Centrifuge at 16,000 × *g* for ~~5 minutes.~~ 2 min
6. Empty the Collection Tube and recentrifuge the column assembly for ~~1 minute~~ 3 min with the microcentrifuge lid open (or off) to allow evaporation of any residual ethanol.



Elution

7. Carefully transfer Minicolumn to a clean 1.5ml microcentrifuge tube.
8. Add ~~50µl~~ of Nuclease-Free Water to the Minicolumn. Incubate at room temperature for 1 minute. Centrifuge at 16,000 × *g* for 1 minute. 30 µl of 65 °C warm water
9. Discard Minicolumn and store DNA at 4°C or –20°C.

3760MA07_2A

Additional protocol information is available in Technical Bulletin #TB308, available online at:
www.promega.com

ORDERING/TECHNICAL INFORMATION:
www.promega.com • Phone 608-274-4330 or 800-356-9526 • Fax 608-277-2601

Promega

Printed in USA. Revised 11/09
Part #9FB072

10.2 Unbiased miRNA target analysis

miRNA	hsa-miR-215-5p	hsa-miR-215-5p	hsa-miR-215-5p	hsa-miR-194-5p	hsa-miR-27a-3p	hsa-miR-23a-3p	hsa-miR-192-5p	hsa-miR-146a-5p	hsa-let-7d-5p	hsa-miR-375-3p	hsa-miR-139-5p	hsa-miR-182-5p	hsa-miR-214-3p	hsa-miR-582-3p	hsa-miR-92b-3p
database	Tarbase	Targetscan	microT	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase	Tarbase
higher (+) or lower(-) expression in HP	+	+	+	+	-	-	+	-	-	+	+	+	-	+	-
		PLXNB2	CTNNBIP1	BMI1	TERF2	TERF2	FZD7	CALU	CALU	CALU	BMI1	TMSB10	GPAT2	CALU	RTN4
		ALS2CR11	ESR1	TBC1D20	CALU	CCNT2	EFNB2	SLK	TGDS	EFNB2	RTN4	FAM102B	PLXNA2	EPB41L4A	ARPC5
		TTL	ZEB2	LAMB2	BMI1	FAM102B	CTNNBIP1	EPB41L4A	CCNT2	SLK	SLK	GOLPH3L	CCNG1	SCLT1	ZNF503
		FGD5	ZMAT3	CCNG1	TMSB10	ABCD4	WWC2	PELI1	ZDHHC18	WWC2	PELI1	ZDHHC18	NEK7	NEK7	DDI2
		DIEXF	FRMD4B	IGFBP5	ZKSCAN5	RPRD2	C10orf6 (hsa)	ATF6	SLK	CELF2	VPS13C	RTN4	ZEB2	PAQR3	IRS2
		MED14	C10orf126	TMED8	RNF41	EPB41L4A	COPS7A	C16orf52	ESPL1	TBC1D20	IRS2	COPS7A	FAM120A	ZEB2	RNF11
		ARL2BP	ALS2CR11	HBEGF	CCNT2	YTHDC2	B2M	BRK1	EPB41L4A	PELI1	KLHL28	ESR1	FST	WIPF2	BRAF
		PTPN4	BLCAP	MFAP3	POLR2B	ESR1	HSS00104143 (hsa)	NCAPG2	ESR1	CAMTA1	CAMK2D	DDI2	SCAMP2	GSK3B	FHL2
		C16orf46	TTL	UBXN7	GOLPH3L	E2F7	PPP1CA	(PTPRG) (hsa)	PPP1CA	SLC44A2	POU2F1	RAB2B	STAT3	PRKCA	POU2F1
		PREPL	GPR22	SCAMP3	RTN4	DDI2	TRIB3	SCARF2	TRIB3	ABCB10	ZNF544	LARP6	HPCAL1	ITPRIP	ANP32E
		SPATA13	DICER1	BRMS1L	VAPB	RAB2B	ABCB10	SLC38A1	SLC44A2	RPA1	IQCJ-SCHIP1	TRIB3	ASB1	PEAK1	TNFRSF11B
		DDX6	ADCY7	ARL6IP5	SLK	CAMTA1	RPA1	TLR2	ABCB10	ZBTB40	SEC16A	CAMTA1	PNPLA2	RREB1	DUSP4
		RAB2A	DIEXF	TNPO1	RPRD2	RPA1	VPS13C	NEK7	PLXNA2	GABBR1	PDLIM5	RGS17	ARID3A	TNPO1	MTO1
		CRB1	PHAX	NAA35	ESPL1	VPS13C	SRRT	GSK3B	RGS17	KLHL28	TRA2B	FBXO5	NDST1	CDH2	TMEM50A
		CXCL2	BRIX1	ARFGEF1	YTHDC2	KLHL28	C16orf52	SOX4	PFN1	CEP70	CTNND2	PFN1	ATP2C1	G3BP1	BRI3BP
		CCDC121	RIC8B	CDH2	SEMA6A	FHL2	FBXO5	FBXL3	IRS2	RPGR	ZNF786	PTPDC1	YAP1	TBC1D13	RAPGEF3
		B3GALNT1	MED14	G3BP1	NDUFS2	FASN	BUD13	OTUD1	RBM14	TTC7B	TSC1	COMMD5	RNF38	TCEB3	SERTAD3
		LIMS1	PRR23D1	ZSCAN25	CELF2	ZNF544	POU2F1	DAZAP2	BUD13	NACA	PTPN3	ZBTB40	TM9SF1	OTUD3	PPP1R37
		LPAR4	PRR23D2	MMS19	PPP1CA	CCNG1	UBASH3B	NFIX	BRAF	PRLR	MYBL1	SLC25A44	NUCKS1	FEM1B	PCGF6

		PNPT1	KCNQ5	TGFBR1	E2F7	DCLRE1A	ACTB	APMAP	POU2F1	PPP1R2	MRFAP1L1	IQCK	AMFR	RAP2B	JARID2
		SCN3A	ARL2BP	DSC2	DDI2	TANC1	EAF1	BRMS1L	SLC41A2	PAQR3	CACNG8	FASN	NRAS	TROVE2	TCF3
		DNAH9	PTPN4	OSBP	ABCB10	SLC38A1	HSPA2	RPS3	CDC6	SPAG16	SAR1A	DCLRE1A	CRKL	UBP1	GSK3B
		TYMS	DBT	VDAC1	PLXNA2	SNAP23	PRLR	ETAA1	ACTB	SEC16A	TRIM68	EAF1	TM9SF2	CDT1	RNF44
		PKP4	MCM4	BBS2	BRK1	PRLR	AF086032 (hsa)	RAG1	NACA	NETO2	CDH2	APTX	SLC47A1	ITGB1	FAM120A
		RPAP2	SOAT1	ZDHHC21	PTPDC1	SRPX	FBXO25	TNPO1	HDX	CCDC88 (hsa)	AP4B1	EIF3L	CLIC4	CCNB1	SOX4
		CNGB3	WNK1	RAB35	TMPRSS2	TMEM50A	DCP1A	ARGLU1	BRI3BP	CNOT6	RPS23	SLC38A1	KIF1B	EZR	PDPR
		SLC19A2	C16orf46	ZFAND5	C12orf73	BRI3BP	EGR1	G3BP1	MAGI2	TMED8	TUFT1	PHLPP2	BAIAP2	ZNF621	MCOLN2
		DYRK3	CRISP1	YAP1	KLHL28	NEK7	HSS00330832 (hsa)	TMTC2	MED29	ITGA (hsa)	CDS1	PTK7	MAGEF1	SSR1	BRMS1L
		IGJ	SPATA13	WASL	CAMK2D	MSI2	FOS	ADAM10	EGR1	CXADR	NFKB1	PRLR	ZNF710	ANKRD11	SPARC
		NCAPG	FAM229B	FUS	BRAF	PDLIM5	RAB8A	CAPN2	SFSWAP	SOX4	OAZ1	MPHOSPH8	ZMYM4	CCT2	RREB1
		MCPH1	UMODL1	ITGB8	FHL2	JARID2	Contig43200_RC (hsa)	ISG20L2	ZEB2	PRKCA	TROVE2	DCP1A	FBXW7	NUCKS1	CLPP
		L2HGDH	RAB2A	ATF2	HGSNAT	HS6ST2	ZEB2	IRF4	MLK4	IL1RAP	CHCHD2	RAB3GAP1	ENDOD1	C12orf49	TNPO1
		BHLHE22	SH3RF3	NUCKS1	ANP32E	HIPK2	MLK4	OAZ1	CASP8AP2	NUMB	ZBTB39	LRP10	COL12A1	RFX5	EFCAB14
		CDC7	CRNKL1	ITGA9	LAMB2	TLN2	ZMAT3	SLC4A1AP	HIPK2	POC1B-GALNT4	UBR1	ACKR3	THBS1	RAD23B	UBE2Z
		IKZF2	TRAF5	CAMSAP1	FASN	RPL31	TROAP	PHF20L1	LIN52	DAZAP2	PLCXD1	ZMAT3	RUNX1	CAMSAP1	EID2B
		PRKAR1A	PDP1	EN2	ADAM9	CRTC3	HIPK2	CCNB1	ADIPOR1	CRLF1	PET112	NETO2	EIF4A3	TIMM21	ZSCAN25
		KIF20B	MSN	CXCR4	CCNG1	BICD2	BC002811 (hsa)	ARL8A	RNF44	MYBL1	TTC39C	NAF1	HMGA2	FBXO21	TGFBR1
		RNF8	C9orf153	TMEM33	TANC1	LHX4-AS1	EDRF1	PLA1A	C5orf30	ATG7	DDX3Y	DPP3	PTRF	EMC10	CHMP7
		ZBTB4	TAOK1	PTBP2	PHLPP2	FAM120A	hCT1770636.1 (hsa)	WASL	GPR63	UGCG	YAP1	HIPK2	CDC25B	CAMK4	COX8A
		ACPP	ALCAM	RNF122	PRLR	PDPR	C5orf30	DDHD1	PDPR	TNPO1	SSR1	CNOT6	HMGA1	SLC37A3	GGCT
		RNF141	CCDC121	MTHFD2	TMEM50A	CDIP1	CREBL2	TPD52	SIPA1	PDGFRA	ELF2	CRTC3	PAK2	LPHN2	TNFRSF13B
		KIF5B	B3GALNT1	TM9SF2	NRCAM	PYCR1	Contig11417_RC (hsa)	CDS2	PAGR1	EFCAB14	WDR60	GSK3B	ANKRD6	MTHFD2	PLP2
		EREG	ZXDC	YWHAE	NEK7	PKIA	HBEGF	AK022044 (hsa)	PKIA	ELL2	ZNF827	ADIPOR1	ANKRD52	CYB561D1	FEM1B
			LIMS1	CHD1	BSCL2	PRKCA	CHST12	LTA4H	C9orf40	TOB2	ZSCAN32	CTDSP1	WDR41	GPATCH8	UBXN4

			LPAR4	ZNF532	MSI2	TNFSF4	IL1RAP	KSR1	NUMB	NPDC1	ITGB8	FAM120A	MCM6	HOXA1	SLC39A8
			ITSN2	NUP153	SERTAD3	RIC8A	OTUD1	RFX5	PRSS22	OTX1	WDR12	SOX4	TUBA1B	CALM1	SPRYD4
			GGA3	SCP2	SEC16A	ZBED5	DTL	ZNF12	LRRC47	EMP1	CCT2	TRA2B	RNF103	NME4	UBR1
			DNAH9	SOCS2	DCP1A	LRRC47	KDM1B	PHB2	DUSP22	VASH2	MFSD6	PAGR1	IST1	PCED1A	MED13L
			PKP4	PCLO	IGFBP5	KDM1B	hCT1641762.3 (hsa)	DNM2	HOXA7	ERBB2	CLASP2	MCOLN2	CTU2	ARF3	CIC
			RPAP2	TMED7	EGR1	TAP2	PHF19	MBOAT1	SMG5	WDR1	CTNND1	RNF43	ZDHHC8	PCDHGA1	PSMD5
			AQPEP	COL12A1	LRP10	TSC1	NFKBIZ	LMO2	SLMAP	APEX1	ZNF791	OLFM1	ACTG1	SIK1	DDX3Y
			MSH6	WWP1	PDLIM5	PHF19	NOL12	ATXN1L	SCAMP3	METTL2A	USP25	ESYT1	PXN	STOM	PTGER4
			RAB27B	SIK1	PICALM	CLU	RBM27	ELK4	SDF2	SPPL3	ZC3H18	UBFD1	MARS	TMEM167A	TRIM23
			CNGB3	THBS1	ZMAT3	ATAT1	ENC1	CLIC4	PHF19	LEPROTL1	TRIM27	MFAP3	SNX30	KIAA1522	RNF38
			NEK1	HMGA2	PDK3	VWA9	TUBGCP3	SRPRB	APMAP	EXT1	CXCR4	NUMB	ANPEP	AKIRIN2	FHL1
			SLC19A2	MYL12B	HIPK2	MGAT2	TRIM68	BAIAP2	RBM27	FEM1B	MAST4	USMG5	SPRED1	SPATA2	GCLC
			GOLGA6L1	DHCR24	CNOT6	STIP1	B3GNT5	MEA1	VWA9	SMC1A	TMEM33	DUSP22	ETS1	EXTL2	IRAK1BP1
			KIDINS220	MEIS2	TMED8	RREB1	MAN2A2	HNRNPA3	ENC1	SLC6A6	TTC23	TSC1	NPC1	CAV1	ITGB8
			SRSF6	CAV1	EDRF1	ERBB2IP	NUFIP1	NUP153	TUBGCP3	FSTL3	HP1BP3	PHF19	ASXL1	MED21	FAM126B
			DYRK3	HOOK3	WIPF2	TNPO1	TNPO1	AP1G2	DVL3	TROVE2	YWHAE	CDK7	WNK1	DDX56	PWWP2A
			IGJ	FLOT1	RPL31	UBE2Z	hCT1648208.3 (hsa)	AF086511 (hsa)	PCNA	SIPA1L3	TRIM52	ENC1	HSPA5	IDS	BSDC1
			ZNF773	TSC22D1	CRTC3	FERMT2	KIAA1333 (hsa)	YOD1	UGCG	CBFA2T3	TSC22D3	UST	DCAF5	DTX4	FAM135A
			ZMYM1	LRRC8A	GSK3B	TOX4	ARFGEF1	EML1	B3GNT5	SLC4A1AP	CRIM1	UGCG	PRKRIR	YWHAG	NUCKS1
			NCAPG	UBQLN1	BRWD3	ARFGEF1	C8orf46	ARF3	MYO9B	CACNA1G	UFC1	LPPR2	CD151	PPP1CC	SMAD2
			GOLGA6L6	ATG2B	CXADR	ZC3HC1	ELL2	HOXB8	EPRS	MYB	KBTBD2	STAT3	FOXK1	MCL1	SH3PXD2A
			RP11-625H11.1	TBCK	ASB7	TOB2	KIAA1958	EIF4A3	RREB1	NDST1	UCK2	STIP1	SLC31A1	CCND2	MRPS35
			NTPCR	PURA	FAM120A	CPSF4	FZD5	VANGL1	SIGMAR1	SLC30A5	FBXO30	RAB9B	ATP6V1E1	ACTG1	USP36
			BHLHE22	IST1	RBM25	PFKP	G3BP1	CDC25B	UTP15	PHF20L1	ESRRG	CEBPG	MCM5	TM9SF3	DICER1
			ZC3HAV1	PVR	HID1	ZNF638	HSPA1A	HDAC7	DIP2A	ITGB1	ZNF10	TNPO1	STAG2	RC3H2	TMEM205
			NKAIN2	MCL1	PDPR	UBE2R2	AK095096 (hsa)	MYL12B	GLG1	ZDHHC21	BEND5	EFCAB14	UBE4A	PDE3A	RBM12
			CCDC152	CCND2	SLC15A4	SNX18	CTH	PWP2	CSNK2A2	HNRNPR	FBXW7	GALNT7	EPG5	PTPN14	PAM
			CDC7	ACTG1	FBXL3	ADAM10	NM_020962 (hsa)	CCR9	PIEZO2	GDAP1	COL12A1	PRKCQ	ZNF384	BACH1	MAST4

			SLC24A4	ATP6V1F	FBXO11	DNAJA1	TMTC2	CASK	TOB2	HELB	TRIM24	CHCHD7	SRGAP2C	PSME3	TMEM33
			DAZ4	TOMM20	HBEGF	GGCT	CDON	IFIT1	C2CD3	CCDC32 (hsa)	NSRP1	ZC2HC1A	DAD1	PKM	SLC37A3
			INTS6	TTC33	UBFD1	CHST3	KIAA1731	ALP (hsa)	G3BP1	RPN1	TRIAP1	TAB1	FBLN5	MAP4	HCN2
			ERCC4	HNRNPF	KYNU	FEM1B	VASH2	ING3	MCAT	YAP1	USP6NL	ARFGEF1	ACAP2	FRS2	LRCH1
			MED28	ARPP19	IL1RAP	RAP2B	TCEB3	NF2	PFKP	P4HA1	AGO3	SWAP70	ABCC10	ITCH	BCAT2
			DHFRL1	DYNC1H1	OAS3	SNX11	OTUD3	NR6A1	RAB3GAP2	SSR1	TBC1D9	ELL2	SERPINA7	FZD6	NRAS
			YY1	PSME3	PYGL	SMC1A	APEX1	METRNL	UBE2R2	ANKRD26	TRPC1	MIER1	CD164	ADAT2	CRKL
			IKZF2	BCR	URGCP	MET	ANKRD13A	GPANK1	ZSCAN25	RPL32	AKIRIN2	MSANTD3	MINK1	PPRC1	ATXN1L
			REV3L	UTP3	ZBED5	TROVE2	ACADSB	ZNF597	TBC1D13	LTA4H	OXNAD1	C2CD3	SPCS3	RCSD1	MRPS23
			BRD3	LCOR	UBXN7	LRBA	hCG_1790474 (hsa)	FLI1	TGFBR1	RB1CC1	ZNF431	E2F1	SQSTM1	SRCAP	REXO1
			USP14	LMO7	USMG5	SIPA1L3	LEPROTL1	P2RX5	YIPF3	IL17RA	ZNF131	RAB3GAP2	CTSC	PKD1	ELK4
			KCNA4	PPP3R1	DAZAP2	ZNF273	HYLS1	ATP11AUN	RTTN	ACSL3	NF2	RPS23	C9orf117	FBRS	CLIC4
			LIN7A	FTSJ1	SMG5	BLCAP	PCOLCE2	HNRNPF	ADAMTS5	THEM6	TSC22D1	MASTL	ETF1	PTBP1	KBTBD2
			LRRFIP1	SRCAP	SCAMP3	CNOT11	S100A16	HERPUD1	DNAJA1	NA	YIPF6	TBC1D13	ASPH	ERLIN2	ATP5F1
			SERF1A	CDH11	FST	SSH2	ATP6V0E1	ARPP19	FBXO32	KCNQ2	TGFA	MORF4L1	WASF3	SELT	CLDN11
			CATSPER2	ARL14EP	TSC1	USP53	KIF14	CARD8	COX8A	BDH2	CHD8	ADAM10	MAP3K13	TRAF4	EXOC5
			GABPB2	C7orf60	CLU	ITGB1	TROVE2	SKA2	FAM189B	NELFCD	PURA	DNAJA1	FAM91A1	TUBB	DENND4B
			ZBTB4	CDKN1B	APMAP	CCNB1	SLC39A8	KLF9	PDHX	FAM126B	40787	HIST1H2AJ	PSAP	PPTC7	GPATCH8
			CBLN4	MACF1	RPA3OS	HNRNPR	GALNTL6	HM13	ARID3A	LIG3	ATAD1	ISG20L2	B3GALT5	STAG2	DCAF8
			CST9	WIPF1	PEAK1	ARL8A	C14orf1	BACH1	UTP14A	COG2	NAP1L1	OSBP	DCAF10	IRF2BP2	ERO1L
			SRSF3	UBL3	ARL6IP5	DCTD	ZNF273	LRRC8D	WASF1	RAPGEF1	GCLM	LEPROTL1	COPB1	DEPDC1B	ZMYM4
			PTPRT	DCUN1D4	PSD4	MED13L	TBL1X	ITCH	LRBA	NUCKS1	CGN	FAM168A	HERC1	KLHL21	TCF4
			SERF1B	MDFIC	SPARC	CUTC	MYB	LCOR	BLCAP	RDH11	SGPL1	NFKB1	MCM7	RAB21	KIAA2022
			CLSTN1	RAB2A	UGCG	RAB35	HNF1B	ZNF24	ADCY1	SHROOM3	MYOCD	PPP2R5E	SMARCD1	SFT2D2	RRN3
			XPO1	ANKRD17	EPRS	RPN1	ID2	OLFML2A	CSK	COL6A5	ELF3	PDHX	MSN	EFNA3	HMGCR
			DIXDC1	UTP20	MAN2A2	ZBED4	BORA	PPRC1	CDT1	KCNN4	MCL1	FEM1B	NDUFC2	SHCBP1	CRK
			TM2D1	LPPR1	RREB1	ADRBK1	ITGB1	SLC17A5	PPP4C	KRT8	CCND2	SNX11	PRPF8	PABPC4	FBXW7
			RNF141	CYTIP	SFXN4	DDHD1	KRT18P8 (hsa)	SERPINB9	BBS2	ALMS1	ADAT1	ERI2	MLH1	GADD45A	SSH1
			AMER1	TNKS2	NAGA	TRAPPC3	MED13L	ISG15	CCNB1	ZNF611	ZDHHC8	MET	AMMECR1L	ZFHX3	COL12A1
			APLN	SP3	CLPP	ELF2	PCSK9	NA	ZBTB39	SKOR1	TSPAN3	SLC6A6	CAPRIN1	PSKH1	ARMC8
			GMEB1	CDCA5	TNPO1	ANGEL2	STOX2	KLHL20	ZNF691	DICER1	HNRNPF	TROVE2	PFDN4	ERCC6	ARRDC4

	ATP8B4	TECPR2	GALNT7	RNF38	PET112	SMARCAD1	GRAMD1B	ARHGAP11A	PERP	WASF1	IER3	SNN	SIK1
	NIN	RNF217	UBE2Z	C2orf48	POLR1C	IER2	ARL8A	SDC1	ARPP19	BLCAP	GIGYF1	NOP56	IPO9
	MCM10	SF3A1	ARFGEF1	ZBTB38	NUCB2	VWCE	UBR1	RBM12	PDE3A	CNOT11	ATP2B1	WWTR1	MTMR6
	MYLK	WEE1	EEF2K	ZNF93	NDFIP2	SYNJ1	MED13L	C15orf41	PTTG1IP	POPDC3	NSD1	MAP3K1	HMGA2
	EHMT1	SETD5	TOB2	IL17RA	DDHD1	RHPN2	PLCXD1	CRLF3	ROBO2	UBP1	TUBA3C	ETF1	VANGL1
	KIF5B	GPR158	E2F1	ZNF594	IL1R1	HLA-C	BCL9	VPS37D	ZBTB24	C20orf24	FAM222B	EXT2	TOR1AIP2
	MUM1L1	IGF1R	TMEM204	PEX19	ZDHHC23	PHF23	RAB35	ETV6	LCOR	CSK	TMEM147	FEM1C	PHACTR2
	EREG	NEMF	RCC1	LSM3	TRIM23	ZNF620	SNX19	LIMD2	REV1	TTYH3	TP53	SEC61A1	LGR4
	NCR3LG1	DDX17	MASTL	MCM9	GATM	MED13	HELB	FGFR3	YDJC	ITGB1	TNRC6B	HN1L	ZNF267
	ZBTB8B	PSME4	ASB1	WDR7	TOMM40	PPP2R4	FGD6	PRKAA2	ZFP36	ASZ1	SSRP1	UBE2J1	PIP5K1C
		EVI5	TGFBR1	LIG3	Contig37200_RC (hsa)	ATP11B	CIC	SLC25A15	ADAT2	TBCB	HSPG2	QSER1	KIF20A
		ZFHX3	MSI1	TM9SF1	Contig37920_RC (hsa)	PAPOLA	STOX2	REPS1	UBE2E2	GPNMB	NPR1	HOXB3	ATMIN
		GNAI3	PDE4DIP	ZNF208	ZBTB38	DCUN1D4	RPN1	NLRP2	WAC	WARS2	COL6A1	SLC7A5	CDKAL1
		ZNF195	CXCL16	SMS	RB1CC1	MDFIC	CEP97	KCNE3	ZNF562	MED13L	RBMS1	OLR1	CASK
		TMOD3	WDR1	CXCL8	hCT18819827.1 (hsa)	CSDE1	YAP1	HN1	HSPA5	TIMM9	HECTD1	PES1	C11orf30
		VAPA	TCEB3	GZF1	AK095149 (hsa)	TOMM70A	ZBED4	MTHFD2	RAF1	USP39	CHP1	CDK6	TMEM184B
		LUC7L3	APEX1	FLNC	WDR47	MOB1B	NDFIP2	WHSC1L1	LONP2	FNTB	CD81	IPO7	SETD7
		KCNJ1	GGCT	NUCKS1	THEM6	PPP2R5D	PLA1A	TNNI3	FRK	CYFIP2	AGPAT1	DDX3X	EXTL2
		GM2A	DSC2	BLM	MCM9	CCDC117	TPD52	KCNIP4	SOD1	RPF1	TAGLN2	IGF2BP1	GNPDA1
		IKZF3	EPHB2	KCTD12	SAYSD1	PSMD3	IL1R1	PHAX	RRM2B	C7orf25	COPB2	DNAJC9	SEL1L
		FEM1C	FBXO46	ATN1	TCERG1	SFT2D2	SSR1	SLC7A6	AKAP1	STOX2	CBLB	USP15	TBC1D10B
		CD2AP	NFKB1	FAM83G	MAPK8IP2	PHF3	ELF2	CRIM1	ZNF706	ZFAND5	NUP214	SMARCD1	ZNF431
		COPS2	FAM189B	CTNND1	FAM126B	AGFG1	C1orf43	CHD1	C7orf60	YAP1	LARP1	SIRPA	CPEB4
		GOLM1	IRF4	KDM5C	PWWP2A	BTBD3	PACS1	KIF1B	CDKN1B	NUMBL	FH	BUB3	UBIAD1
		HIC2	UNKL	MRPS35	RHOBTB3	EGFR	GATM	ATP5F1	TP63	TTC39B	PRDX4	CTC-260F20.3	SUGP2
		NISCH	PDHX	LINGO1	CMPK2	TMEM41B	TOMM40	RAB3D	LIN28B	NDFIP2	MAPKBP1	RPS20	USP31
		IPO7	EXT1	ACSS1	NM_020467 (hsa)	IER5L	CSNK1G2	CCNA2	EPG5	WASL	CNIH1	NCOA5	IL10RB
		SERINC1	FEM1B	CBL	DDX50	LY75	RB1CC1	DR1	LPPR1	DDHD1	COL2A1	LYSMD1	USP1
		INIP	RAP2B	SCRN1	IGFN1	IFITM3	PKN3	NT5C2	ZNF302	TPD52	DBP	KLF5	UBQLN1
		RRAS2	EIF3F	DICER1	NUCKS1	SQSTM1	PEX19	LPCAT1	GOLT1B	EIF2AK1	NAV1	RHOT2	WNT5A
		FNBP1	PTPRU	CTR9	GPR137B	HEATR1	FUS	CALM3	KIF18A	ZNF621	NAV2	DST	CTBP1

				MSN	SNX11	CASP7	BLM	PTMS	THEM6	DCTN5	SETD5	ZBTB44	SLIT2	CAPZA1	PHGDH
				ZCCHC2	SMC1A	RCN1	RP11-11C5.2 (hsa)	DHX40	ITGB8	NUP153	SFT2D2	SLC35F1	BTG2	CHM	RNF103
				SNIP1	TOP3A	RBM12	RDH11	HAS2	TCERG1	FAM193A	IGF1R	SSR1	FBN1	NDUFA12	KLF2
				ZNF770	MET	CRLF3	ENST00000296531 (hsa)	ZNF768	HELZ2	ZNFX1	ZNF69	ST3GAL3	GP2	ZNF680	HTT
				APP	UTP14A	PDK4	AF086173 (hsa)	SEC61A1	RHOBTB3	GTF2H1	ZNF480	SH3BGRL	MSH6	RICTOR	DUSP12
				LBR	MFSD1	SLC7A6OS	KCTD12	SF1	WDR7	NME4	PSME4	C1orf43	NR3C1	ASH1L	PWP1
				TAOK1	TROVE2	PRKAA2	ATN1	EVC	UBE2T	PCLO	ZFHX3	ATXN7L3	SSFA2	SDHC	S100A4
				FADS1	UBXN4	DIEXF	U79257 (hsa)	SNRNP27	CDK4	IP6K2	CLTC	ZIC2	PIGC	AMMECR1	SZRD1
				CAPZA1	SRSF9	SCAF4	hCT2328509 (hsa)	TMEM164	GPHN	YOD1	RASGEF1B	KAT6A	AXL	EEF1D	LMBR1L
				ABHD2	SIPA1L3	C12orf76	Contig30437_RC (hsa)	PDP1	FLNC	EBPL	MAP3K4	ANGEL2	TPD52L2	ZNF207	MCL1
				FAM60A	CBFA2T3	SORL1	ESCO2	HNRNPU	LDB2	HMGCR	XBP1	RNF38	COL4A2	DYRK2	DUSP6
				FJX1	BLCAP	STK4	EIF2AK3	TAB2	NUCKS1	VPS13D	ETF1	UBXN6	SMARCA5	TEX19	TGS1
				MYSM1	CNOT11	LPIN1	ALMS1	RPS6	BLM	PSMA6	PMP22	PAX6	KPNA5	CCNT1	HLA-E
				TMEM123	MEGF9	YWHAE	ZNF611	HSPH1	IARS	TRIM37	FAM91A1	MRS2	DNAJC13	ARL4C	ZDHHC5
				GIGYF1	SSH2	CRIM1	STAM	STMN3	KCTD12	ARMC8	FOXC1	RB1CC1	DNMT1	MTR	PSMD2
				CCT6A	RRAGD	FUT4	PXDC1	C1orf21	PREX1	TXNDC10 (hsa)	ZDHHC9	NUP85	ZDHHC4	GABRB3	SATB2
				PSMD13	ID2	DUSP11	TTF2	STAT5B	FAM83G	IPO9	VLDLR	MCM9	MECP2	SCAMP1	MARS
				ZNF680	CSK	MSMO1	RNASEL	IGF2BP1	SNX33	HSPA4	ZKSCAN8	ITGB8	ZFP36L1	XPO7	HOXA9
				RGPD3	VDAC1	KBTBD2	SH2B3	USP54	WIPI1	USP6NL	HEG1	FAM126B	GPBP1L1	TRAFD1	RC3H2
				INSIG1	TTYH3	CHD1	DICER1	RER1	KDM5C	VANGL1	KIF23	RHOBTB3	ARID1A	ZFYVE27	SLC25A32
				RALGAPB	ITGB1	SKIDA1	Contig54491_RC (hsa)	ATP6V1H	SMAD2	TOR1AIP2	NISCH	CDK4	PCBP1	PEG10	STK38
				DHODH	PPP4C	VAMP4	ARHGAP11A	MTHFD1	MRPS35	NIPAL3	VPS37B	FIG4	MTPN	MLEC	U2SURP
				SPTY2D1	RPL17	SPPL2A	C19orf55	MID1	GARS	ZMYND11	DDX3X	ZCCHC9	COL1A1	ZNF143	ZNF551
				THAP10	TOR1A	RFWD2	OVOL1	PRPF8	USP36	AGO3	TCP1	BSDC1	XPR1	ZNF217	RECQL5
				PRRC2C	TTL	ATE1	CASP7	DBF4B	PRSS21	PHF14	PRKCE	DDX50	CENPB	ITGAV	ZNF148
				SEC61A2	ZBTB39	FAM78A	RBM12	CHMP4B	USP38	NCOA7	ZNF438	USP9Y	EIF4G1	DMWD	OTUD6B
				RICTOR	GPNMB	HNRNPD	SMUG1	BCHE	CASP4	TBC1D9	PPIA	NUCKS1	RAB11A	PPL	SYPL1
				ACBD3	WARS2	BTBD1	ZBTB10	RGS13	SH2B3	CEP57L1	PML	POLR2C	SHISA5	CNIH1	LINS
				KLHL11	NUP62	RIC8B	ERICH5	EFNA5	DICER1	CUL2	FAM162A	RDH11	FLNA	POP7	CBX4
				TMEM132B	UBR1	EXOC5	DNAJC19	KBTBD6	CTR9	CLCN3	ZNF107	MGST1	PABPC1	SP4	CASP1

		TRIP12	MED13L	DENND4B	CAMK4	PDS5A	TMEM205	TPR	DEK	GTF2F2	ZNF426	RAB11FIP2	SNRK
		BCAT1	RAB35	TTC4	TCFL5	RFX7	C19orf55	PIK3IP1	C18orf8	ATN1	TEAD1	ZFX	HAND1
		ATP9A	SRR	ERO1L	CCNO	DDX60	KBTBD8	REST	MAT2B	KIAA1324	TMEM185B	NFAT5	GOLGA8B
		GYG1	SNX19	RUVBL1	NM_024084 (hsa)	MTUS2	E2F2	SETD7	CKAP5	RAD23B	PHC2	SLC5A6	KLF12
		CYTH3	RPF1	ANKHD1	SLC37A3	GIGYF1	TRIM27	LMNB2	TAB3	EIF2AK3	NRP1	CMIP	RAD54B
		HECTD1	FGD6	FAM107B	CRLF3	NSD1	CXCR4	CEBPZ	TP53TG3	TIMM21	ERN1	AGO1	MKRN2
		ASAP1	CIC	ZNFX1	PTBP2	NAPG	RBM12	CEP128	PLAGL1	ALMS1	FIGNL1	MAPK9	LRRCC1
		GREB1L	STOX2	ZNF724P	ADCY7	SLC16A3	MAST4	OXNAD1	OSGEPL1	USP25	SRRM2	EEF2	ZNF566
		LAPTM4A	RPN1	GOSR1	DNAJC3	DDX21	EMC10	PSD-95 (hsa)	ADIPOR2	ARL2	TNFRSF21	RAB3IP	ZNF121
		ACVR2B	GNA12	YOD1	ADRBK2	ZNF197	C15orf41	IFIT1	DCTN2	GARS	PEBP1	PHC3	LCOR
		EDEM3	PET112	IFI30	UGGT1	ACBD3	PTBP2	CAPN8	WDR5	BRD2	HKR1	ORMDL1	FZD6
		ALG5	TACC2	PPAT	PRKAA2	SRSF1	HCN2	USP31	NSD1	CBL	TAPT1	FBN1	GPR180
		CHP1	EFS	TAF9B	DIEXF	ZDHHC12	CBFA2T2	SPTLC2	FAM222B	PRSS21	HIST1H3H	ZNF711	ZNF24
		CD81	ZFAND5	PDDC1	RPL27	NMI	NRAS	DPYSL3	ZNF324	DICER1	SURF4	SYBU	TOB1
		RAB5A	TTC39C	GUCY1A3	DLG7 (hsa)	THAP5	LIMD2	DIDO1	ZNF680	FRMD4A	ZNF518B	PPP1R15B	WAC
		FAM69B	YAP1	SSH1	CXorf57	STC1	H2AFX	VSIG10	RPS6KA5	KBTBD8	PLIN5	CLINT1	PPRC1
		SLC2A3	ADRBK1	TRIM37	TMEM194A	AZIN1	DIEXF	USP1	SIX1	CBX2	RBPMS	RNF4	RPL37
		PTPN13	NDFIP2	UBR5	ARHGAP19	GLTP	DUSP2	AP2B1	RHOT1	DNM2	ZNF451	BABAM1	SETD1B
		ETNK1	FOXN4	ARF3	SASS6	OPN (hsa)	IP6K1	CHD8	PRRC2C	CMTM3	TSPAN17	FAM208B	ARID1B
		CBLB	ZNF621	THBS1	WHSC1L1	ZNF490	TMEM194A	PURA	ASH1L	MLF2	AHSA1	FNDC3B	IRGQ
		ZNF217	ZBTB44	DBNDD2	THUMPD3	HSPA1A	ATXN2L	MYO5B	TYW3	TPD52L1	TNKS1BP1	ZFP37	HLA-DRA
		TOE1	SSR1	RBM18	ATXN1L	SLFN11	ARHGAP19	ODF2	TOP2B	CXCR4	RPS6KA3	PPP3CB	ITGA5
		DCAF6	ELF2	USP6NL	STK4	PAPD5	C6orf136	PAFAH1B1	ZSWIM6	RBM12	PRKAR1A	PTBP3	SLC36A4
		TAF1	AREL1	HMGA2	PIWIL3	ZDHHC16	HN1	IDS	EDEM1	ZBTB10	MYADM	VTI1B	PRKRIR
		EBF3	C1orf43	PTRF	PRRC1	KIAA1524	SASS6	MREG	SRSF1	ERICH5	CDK17	STRBP	CCDC28A
		RAB11FIP2	GATM	TOR1AIP2	RSE_00000676604 (hsa)	BG547557 (hsa)	MTHFD2	PITPNA	PURB	FAM32A	ATXN7L3B	ABCC9	APPL2
		DDIT4	ZIC2	UBTD2	PIGX	CD81	ATXN1L	HIST1H1E	KLHL11	TMEM33	ELMO1	COPS5	LIMCH1
		RBM39	KAT6A	PIK3CB	FEN1	AKT2	ZBTB2	ARL2BP	AMMECR1	MGAT5	BRPF1	MSL1	TIAM2
		SOST	ANGEL2	DAB2	ELK4	SLC2A3	VPS18	REEP3	EEF1D	BCAT2	PCDH7	E2F3	CCT5
		PCNX	RNF38	HDAC7	NRG4	DKC1	YWHAE	DTX4	RFC1	MMGT1	GUCY1A2	RGMB	BCL11A
		BTG2	ZBTB38	ZMYND11	SAMHD1	CECR2	TSC22D3	YWHAG	STC1	NRAS	GNAI2	MTRNR2L1	PEX5

				PDLIM1	LTA4H	SBF2	CDC14A	ETNK1	CD80	PDZD2	AZIN1	PDK4	IGFBP3	MYC	AFAP1
				ZNF395	ZNF827	ZNF267	HSS00018228 (hsa)	TMEM67	UFC1	SEC23A	ZNF343	STX5	ZBTB33	UBE3C	SELT
				TSC2	RB1CC1	AGO3	PIF1	GART	INPPL1	MCL1	DCAF7	LTN1	C5orf24	BMF	HIST2H4B
				JUN	DNAJC6	PHF14	RAB3D	TRIM22	REXO1	PTPN4	ZNF490	LPHN2	ENPP1	CSTF1	TMED5
				CCT3	IL17RA	MYL12B	AF131846 (hsa)	GIMAP4	ELK4	DUSP6	RYK	PRKAA2	CHCHD6	ANKRD12	OSER1
				CLINT1	ACSL3	DHCR24	RMI1	KLHL42	FYCO1	TGS1	CCNT1	DIEXF	TPP1	VEZT	MYH9
				CCND1	ALPK3	TBC1D9	FAM111B	ITGBL1	CLIC4	KIAA1191	EZH1	IAPP	XPO4	MTCH1	STYX
				GRPEL2	KSR1	ZNF681	COMMD7	TRAK2	PRR3	SCLY	ZBTB46	SLC25A15	FN1	SRSF6	DDX6
				WDR82	LSM3	CUL2	RIC8B	DNAPTP6 (hsa)	KBTBD2	CCND2	RNF10	ADORA2B	WDFY1	SPRYD3	DTX2
				SLC35D1	MCM9	CLCN3	PITPNC1	KIAA1432	CHD1	SECISBP2	KIF1A	RNF122	SGTA	DEGS1	PPTC7
				SPEN	TRAPPC11	SLC19A1	Contig51252 (hsa)	YTHDF2	ABT1	FAM35A	LAPTM4A	STX10	GJA1	CPNE3	HIPK3
				RTDR1	PLEC	C11orf30	GPATCH8	TET3	TMED7-TICAM2	NUDT21	ACVR2B	IP6K1	TMEM30A	SLC16A1	DCBLD2
				PVRL4	FAM126B	MEIS2	UCK2	IFI44	ZFYVE1	HSP90AA1	ZNF721	SH3KBP1	MED20	TFRC	PAPOLA
				AXIN2	LAMB1	ANKRD52	DCAF8	EPSTI1	CCNA2	DISP2	PTK2	HIST1H1C	ATP5L	FLNB	CAND1
				KDM4A	SAR1B	PRRC2B	KLHL8	CLIP1	AGAP3	MKL2	DKC1	NDRG3	COL8A1	ABHD4	MRPL37
				MYLIP	TOR2A	MTUS1	GNAS	ZNF26	MBTPS1	EEA1	STK38L	C6orf136	CCDC80	TMED10	AFF4
				CTTNBP2	ATF2	LMNB2	CIB2	BTG2	DENND4B	HNRNPDL	ZBTB37	CD59	KIF21A	ATRX	IBTK
				INPP5K	LDB2	HOOK3	TMED7	C1RL	SCARA3	RAB6A	PNPT1	THUMPD3	TUBA4A	CBX6	EPG5
				ZNF264	NUCKS1	LRRC8B	PPAT	JUN	GPATCH8	HOXA9	ETNK1	ATXN1L	KIFC3	ZMYM1	CSDE1
				NUFIP2	BLM	USP37	EML1	ABHD17B	WDR55	TOMM20	LCORL	KLF15	RCOR3	XPR1	NA
				POGK	HIST1H1D	CELSR1	ENTPD5	BABAM1	ZNF710	TTN	ZNF217	CNOT4	PTX3	ZMYM2	MARCH9
				TMEM87A	KCTD12	B4GALT5	APC	PCDHGB6	ERO1L	FHOD3	LARP4B	SLC25A46	FAT1	SMURF1	SLC7A11
				HSDL2	PREX1	GNPDA1	Contig40516_RC (hsa)	SMAD4	NUP153	LMBR1	LARP1	MAP7D3	PFKFB2	SLC38A2	MEF2D
				NUP107	ZNF12	ZNF431	TAF9B	ZNF678	HIP1	MTMR4	ARHGEF11	SLC7A6	CHRNA10	NCOR1	MTMR9
				SACS	CCDC92	CPEB4	CRK	TBC1D24	NME4	MAPKAP1	SHBG	RBM45	BNIP3L	RNF19A	TNKS2
				ZDHHC6	CTNND1	ZNF131	CUZD1	ANXA7	YOD1	GMFB	DMWD	PRRC1	LIMA1	EIF4G1	SP3
				LIMK2	CHCHD1	MPDZ	ENDOD1	SPEN	GUCY1A3	FAM82C (hsa)	TBX1	YWHAE	AHNAK	COPZ1	MAN2A1
				CENPF	SH3PXD2A	DPYSL3	ARF3	PTBP3	SSH1	HERPUD1	CSE1L	HDAC3	NRIP1	GLUL	C19orf12
				MKLN1	BNIP3	FTH1	C9orf164 (hsa)	AK026784 (hsa)	TRIM37	ARPP19	TSKU	TSC22D3	LEPR	RBM26	TECPR2
				CECR5	MRPS35	IL10RB	ZNF114	FASTK	COL12A1	PGM2	CDH1	BRIX1	TMEM127	AAMP	YWHAB
				SNX1	PKP1	MTX3	SRD5A1	INTU	ORC2	CARD8	MRPS18B	FYCO1	PPIF	SLC39A9	RNF217

				GLI3	KRT8	CA2	BC041380 (hsa)	OASL	ARF3	AKAP9	TET3	UPF3A	USP11	NCOA2	PRDX5
				ITPRIPL2	ZNF70	LRRC8A	SIRT1	ZNF264	SRD5A1	SPRED1	CMIP	COPS8	NCKAP1	MEPCE	WEE1
				CLDND1	ALMS1	ATG2B	Contig46232 (hsa)	IFIT3	SIK1	SKA2	KLHDC10	BTRC	ALDOA	BAZ2A	USF2
				DENND5B	USP25	NR6A1	SLC35G1	NUFIP2	RUNX1	MYH10	MMP2	MMS22L	LAMP1	COX4I1	ZNF41
				BIRC6	GARS	HHIP	VANGL1	LRRC58	TMEM167A	CMAS	SCAPER	PRR3	TAF1B	ZNF592	MTFMT
				SLC25A24	BRD2	SORT1	GTF3C4	FAM8A1	IPO9	NRBP1	SLC5A3	AP3S1	SRD5A3	PHLPP1	PPCS
				TTC3	CBL	RSF1	ZMYND11	CNNM3	C1S	EFTUD2	ATP5G3	CHD1	PBX2	CHAC1	SH3D19
				SMURF1	C11orf95	E2F8	SUPT4H1	THC2247947 (hsa)	ZNF521	TCF12	TRERF1	KIF1B	ITPR3	KAT2A	SNRPE
				GIPC1	SH2B3	CSNK1G3	ZNF385B	TRAF6	RBM18	RAB28	SLIT2	ACER3	VEGFA	KLHL15	LSM14A
				RABAC1	BCL3	FBXW8	ZNF267	ZNF493	HMGA2	FOLR1	PCNX	ABT1	PTEN	SLC12A2	DSE
				KANK2	HNRNPA1	MYO5B	CYC1	LNPEP	CPSF7	CBX4	CLIP1	TGFBRAP1	FGFR1	CUX1	SSR2
				DDB1	DICER1	ZNF275	NPHP3-ACAD11	SUPT16H	C5orf45	YPEL2	EEF2	RFWD2	DUSP16	WDFY3	AGO4
				RAD51C	CTR9	CYP51A1	NA	SRPK1	TMEM101	STX3	NOTCH1	RIC8B	ERRFI1	GLCE	GADD45A
				CERK	ARHGAP11A	40787	ROPN1L	DEDD	ITGA8	USP16	PLCG1	PPM1K	CLIC1	SNX2	ZFHX3
				JMJD1C	CBX2	SLA	AK057167 (hsa)	LMTK2	C17orf75	SERINC5	BTG2	B3GALT6	NFIC	STC2	GNAI3
				TLE4	CAMK2G	RPS15A	PRIMPOL	IFI44L	ATF6B	CCDC12	ZNF37A	RAI1	ATP13A3	PAF1	HSPD1
				FOSL2	DRG1	ATAD1	ZNF816	FBXO3	SBF2	MKRN2	ZNF395	LZIC	MAPK1	ANKFY1	ARF1
				H2AFZ	DNM2	LANCL1	MEX3B	MKLN1	ZNF385B	CAMSAP2	JUN	GPATCH8	C22orf29	ATP6V1A	GSTM3
				DARS	CSRP2BP	HBS1L	40422	HLA-DPA1	AGO3	UXS1	FBXO28	STXBP6	HECTD4	CDC42	CLTC
				CYB5R4	EN2	RNF166	DHCR24	PAFAH2	PIP5K1C	ZNF121	XPO6	SIMC1	FAM84B	ADAMTS1	ZNF800
				PATZ1	TRIM27	SLTM	ARMCX4	ACO1	SALL2	ANKRD50	TXLNA	UCK2	HOXA10	SKIV2L	VCL
				CFL2	RBM12	NAP1L1	PRRG4	TP53INP1	KIAA1522	KIAA0152 (hsa)	IMPAD1	RUVBL1	USP22	TMEM106B	AHRR
				MRPL10	ZBTB10	GCLM	X97444 (hsa)	ELMSAN1	ETV3	CBWD2	PFN2	DCTN5	PCM1	SDE2	ARHGEF10
				TCF20	PTBP2	ZNF43	KIF20A	WDR36	DHCR24	DNAJC10	JUNB	LRRC61	TAX1BP3	TRIO	ZNF224
				BAZ2A	CBFA2T2	HTT	RECQL	FLJ10986 (hsa)	KANSL2	SAMD9	CTNNB1	CALM1	RPL4	PSMD1	ARFGEF2
				SATB1	LAMB3	YTHDF1	ATAD2	ITPRIPL2	PRRG4	HIST1H4C	UNK	L3MBTL2	SLC12A7	PIK3C2A	ASPH
				TGIF1	MMGT1	YWHAG	TRPC1	GTPBP3	TTLL12	INPP5F	ZNF252P	ZMYM4	JUP	RPL28	FEM1C
				PPM1L	NRAS	SEC23A	ANKRD6	MAPRE3	SMIM8	WAC	VTI1B	KLHL8	PIAS1	ZNF518B	GRAMD3
				ZNF280D	PDK4	PPP1CC	NM_053048 (hsa)	CPNE3	SLC19A1	IQGAP1	MSH6	ZNFX1	TBL1XR1	PAWR	SF1

		CHAC1	SMIM21	MCL1	AK024556 (hsa)	MTPN	PCSK5	DNMT3B	PABPC3	YOD1	SLC1A5	GEMIN7	CDK5R1
		FXR1	UGGT1	PROK2	MEIS2	DHRS1	EDEM2	GAN	NUFIP2	WARS	GRSF1	BTAF1	ZNF330
		PPIL4	LTN1	HLA-E	NM_018071 (hsa)	REL	FAT3	TRIP11	TSR1	TMED7	TECR	LRRC45	PMP22
		SENP1	CTB-50L17.10	NUP133	Contig44186_RC (hsa)	ACTN1	ANKRD52	WNK1	PM20D2	BRPF3	RAB14	ADNP	APPL1
		SLC26A2	ADORA2B	GNA13	hCT1831218 (hsa)	POLR2E	MAP2K7	GPATCH2	NDUFS1	PPAT	JAK1	FIGN	CD2AP
		ASCC3	HIST1H1C	SECISBP2	PITHD1	CBX6	PVRL2	NFIL3	RGMB	APC	CD63	CDK13	COPS2
		RAC1	ARMCX1	NUDT21	EXTL2	BAZ1A	SPATA2	SEMA3C	VPS37A	PPME1	MDM2	UBAP2	QSER1
		ZSWIM4	ATXN2L	ZNF639	LMNB2	IFI27	MTUS1	ZFYVE26	ZNF254	SOX11	PGK1	ALX1	HNRNPU
		ANKFY1	HN1	PCK2	XK	RAB18	PPP2CA	PREPL	TRAF6	SRP54	SLC39A14	PDCD11	TAB2
		DNMT3A	MAP1S	RALGAPA2	Contig1111_RC (hsa)	RPL11	PITHD1	TRPS1	KTN1	PDDC1	DDR2	MAT2A	VLDLR
		KIAA0226	WHSC1L1	IFNB1	NM_018059 (hsa)	RBL1	TPP2	RAF1	DNAJC13	VPS13D	ATF4	PGRMC1	KRAS
		DZIP1	ATXN1L	PXN	CEP128	RBM26	TMCO6	FNBP1L	LNPEP	PSMA6	AJUBA	PSMC3IP	CDK6
		KIAA0247	STK4	EID1	HOOK3	ZNF175	UBIAD1	B3GNT7	CPQ	DHX29	JAG2	NDUFA8	FOXN2
		BTAF1	DGKE	TM9SF3	CGNL1	PTGS2	EXTL3	BPTF	MKLN1	GUCY1A3	CDC42BPB	FBXW4	DDX3X
		MED28	MAP2K2	TTN	GNPDA1	MR1	FTH1	RPUSD4	SMIM15	NFKBIB	SERPINA1	TGFBR3	FLAD1
		OSTM1	VPS18	RC3H2	BCL2L13	DHX8	PLAGL2	APPL2	USP46	TRIM37	MAP1B	HDLBP	SETD2
		SMIM13	TPCN1	FBXL4	HSS00303498 (hsa)	PCMTD2	ACVR1B	TMEM64	WSB1	CFL1	NACC2	CBLL1	COPB1
		CDC23	PPP2R3A	CFDP1	GAPDH	TCF20	KIF2C	LIMCH1	SERPINI1	HIST1H2BK	TIMP1	MGEA5	PHF10
		ARRDC3	ZNF124	NUP210	UBIAD1	AEN	LRRC8A	NUP93	KCNIP2	TRIM24	PALM2-AKAP2	MAPK12	PRKCE
		NDUFB2	YWHAE	PTGS1	HSS00102527 (hsa)	SREK1	NR6A1	SYT4	PTP4A2	ARF3	WBP1	CDK12	ASF1B
		RNF157	TSC22D3	SPOCK1	MCM6	IFIT5	AP2B1	RTN2	ITPRIPL2	WWP1		CNOT1	WDR26
		TSPAN14	CRIM1	MTMR4	PLAGL2	RORA	TBCK	BCL10	DEGS1	SHC1		TREM1	TRAM2
		TGFBR3	FEN1	ZCCHC12	CADPS2	FBXL19	PURA	BCL2	YARS	RNF123		CREB3L2	ATOX1
		CYP20A1	ELK4	U2SURP	ZNF346	STIM1	IST1	ECHDC1	MSH2	TRIAP1		CLMN	ZNF770
		FAM160B1	TMEM59	FAM198B	HSS00138557 (hsa)	KLHL15	XRCC6	TMED5	DYNLT1	AFTPH		TMA16	RABEP1
		TLK1	MORC2	ARPP19	USP1	FXR1	MRPL51	SRGAP1	TUBB2A	SRD5A1		DDX39B	MARCH7
		MTRR	BID	ANPEP	hCT10607.2 (hsa)	SLC26A2	CD37	MED13	BIRC6	SIRT1		PABPN1	LBR
		PRDM2	CLIC4	ZNF148	KIF15	STXBP2	PAFAH1B1	ZDHHC11	SLC25A23	HOXB8		DHDDS	SIK2

				ST13	HOXB4	SKA2	CDKN3	ZSWIM4	OGT	MACF1	C3orf70	MED14		EIF3M	MYO6
				SLC44A1	TIMELESS	ARNT	RNF103	ZNF823	NAP1L1	DDX6	TMED10	THBS1		ASS1	ERGIC2
				ZBTB7A	PRR3	LMBRD2	CKLF	PRKAA1	GCLM	PTPN1	ZNF525	NLN		CYB5B	CENPL
				REV3L	KBTBD2	BRIP1	E2F8	LTB4R2	ARL2BP	PPTC7	CCDC47	IPO9		AURKA	ABI3BP
				BRD3	CHD1	PSME3	PURA	IQCE	REEP3	ANKRD36C	ARNTL2	EIF3H		NUP43	NCOA3
				TMEM50B	CCR3	TAF11	HSPBAP1	STX12	LMBR1L	CAND1	ME1	USP6NL		PLEKHA1	SKP1
				ZNF281	FAM195B	MAP4	XRCC6	HSPA4L	NEDD4L	BOD1L1	SLC25A17	HMGA2		FAM102A	CCT6A
				BCL7A	KLHDC2	SNRK	PAFAH1B1	ZBTB22	YWHAG	DCUN1D4	C6orf106	TOR1AIP2		RTF1	NCL
				PARP1	SPPL2A	CTC1	ACSF3	CARD10	MCL1	TMX1	ANKLE2	GTF3C4		UBR7	PHTF1
				SEC24B	TMED7-TICAM2	BCL9L	ATAD1	C16orf72	LRRC20	AFF4	GPR56	ZNF780B		SKAP2	USP28
				GRAMD1A	TMEM39A	GOLGA8B	STXBP4	NLGN1	POMT2	HNRNPA0	HOXB9	CEP85L		SEC22C	SOX21
				CD4	RCAN1	EML4	RAB40B	TXLNG	SECISBP2	CSDE1	SLC11A2	PHACTR2		SLBP	INSIG1
				RBBP7	CCNL2	COLEC12	AK057710 (hsa)	DYNLL2	DNAJB11	DPY19L1	HES4	PIK3CB		IPO8	FAM104A
				SMC2	CCNA2	MAP2K5	SOWAHC	CLASP1	TAOK2	TPRKB	EIF4G1	DAB2		CTCF	ZNF354B
				ELOVL7	GNS	MESDC1	HTT	NF1	ACTG1	ITIH4	SNRPF	ZMYND11		APOOL	RALGAPB
				RWDD2A	SEL1L3	ANKRD50	DCXR	RP11-111M22.2	HSP90AA1	HELQ	TMEM177	MAPK7		SYNE2	TNRC6B
				EIF3M	FAM78A	LCOR	ARL2BP	SLC50A1	TSPAN3	WDR20	KCMF1	LGR4		GTPBP10	THAP10
				RBX1	BTBD1	MCM4	ZNF200	PIGB	FBXL14	LMNB1	ARL6IP6	AGO3		HIAT1	RPP14
				FAM193B	ZNHIT2	SLC26A11	ARHGEF12	JADE2	ADSS	CDKN2B	PABPC1	CABLES2		TNRC6A	ASH1L
				C5orf24	PPM1K	LMO7	SAMD13	CDC73	RALGAPA2	MTMR9	BAZ2A	MSH3		ZMAT2	RBBP6
				ATP6V1C1	EXOC5	DYNLL1	FAM189A2	FANCF	ZNF445	BCAS1	CPEB2	KIAA1522		DISP1	VARS
				NGRN	LPCAT1	ZNF24	Contig37371_RC (hsa)	CDC23	WDR70	SP3	RORA	ETV3		RAC3	EDEM1
				TRDN	ANKRD30B	PPP3R1	NEDD4L	NFE2L1	ATP6V1F	MOB1B	HSP90B1	HMGA1		ZBTB18	FADD
				IL6ST	MAGEF1	RDX	PPP1CC	TSPAN14	EEA1	ATP1B1	CDC42SE2	TBC1D9		CSTF2T	HUS1B
				ZNF644	ZBTB17	HDHD3	CD83	SLC22A15	MARS	SERINC3	NAA30	MORC4		SMG7	USP45
				KMT2A	FHL3	ZNF117	Contig33888_RC (hsa)	PLIN2	TOMM20	DEPDC1B	ZKSCAN3	NPLOC4		FAT1	AMMECR1
				UBAP2L	ATP11C	CTBP2	CAPNS1	MIDN	GLI2	YWHAB	TMPO	CUL2		DONSON	TAS2R30
				SLC7A2	ERO1L	WAC	Contig55242_RC (hsa)	PPHLN1	PSMB7	THRA	TNK2	LENG8		PPM1H	ARIH1
				UBN2	CALM3	IQGAP1	SMARCA2	LINC00304	IFNLR1	AIM1	ASCC3	PAK2		DKK1	SLC25A38
				CTCF	DCTN5	GAN	MKL2	PLAG1	LMBR1	SFT2D2	WDFY3	ATMIN		ZBTB4	BCAT1
				LRIF1	LRRC61	PSME1	SATB2	REV3L	EHD4	PHF3	ZNF823	SMIM8		PNMA1	DHRS7

		ATP8B2	ZMYM4	ASXL1	EEA1	STON2	MAPKAP 1	AGFG1	WDSUB1	ATAD2		BTBD7	FAM20C
		TNRC6A	RBM48	WNK1	TM9SF3	TSPYL1	HNRNPF	RCOR1	SHOC2	PPIC		PHF13	AZIN1
		AVL9	NUP153	PLEKHF2	HSS00062828 (hsa)	TRMT112	ARPP19	U88048 (hsa)	KMT2D	EIF4E3		TMEM127	H2AFV
		UBR3	MIER2	GSE1	NM_017842 (hsa)	PABPN1	ZNF551	PPCS	KIAA0226	ANKRD52		TSN	EZH1
		PRKD3	HIP1	VPS29	DBT	KIAA0040	GDF11	MSL2	LDLRAD3	PRRC2B		ZDHHC17	HECTD1
		RRM2	KDM2B	PHF20	ORC1	TNIK	PDE3A	IGF1R	RNF219	PVRL2		PIEZO1	MAPK1IP1L
		LDLR	GNAS	B4GALT6	IFNLR1	ACTL6A	PELO	SH3D19	SDE2	SPATA2		TTC27	LAPTM4A
		PTX3	PCLO	SETD1B	HNRNPF	GNAI2	CMAS	CENPQ	C16orf72	MTUS1		NKRF	EDEM3
		FAT1	IP6K2	PREPL	PERP	ZBTB33	STIM2	U69195 (hsa)	RAD1	NKX2-5		IVNS1ABP	FAM46A
		NAA25	FAM115A	SPATA13	AK023367 (hsa)	ZNF257	ETS1	IDNK	DYNLL2	ELK3		TADA1	ACAA1
		GORASP2	GOSR1	GPSM1	XM_170418 (hsa)	MVD	KLF9	PSME4	ZC3H11A	TMEM184B		PDZD8	GRAMD4
		PTPN12	MKL1	MGA	NPTXR	AK057704 (hsa)	HM13	UGP2	ADNP	PPP2CA		ATP6V0D1	MTR
		BTBD7	WARS	FAM229B	hCT1642490.1 (hsa)	SCAF8	DYNC1H1	ANKRD1	ZNF430	CREB5		USP3	AKT2
		FAM63B	TMED7	BPTF	CEP78	KIF22	ARHGAP17	AGO4	KIAA0100	XK		RAB11FIP1	GET4
		FAM214A	IWS1	TMEM64	PDLIM3	GRPEL1	ZFC3H1	GADD45A	KLC1	PDCD6IP		GTF2A1	DPY19L4
		PRKDC	SMAD9	RPL7L1	BRIP1	CMC2	BACH1	ZFHX3	UCHL3	SSU72		PHIP	ZNF721
		PAFAH1B2	LAMA5	KCNMA1	ZNF45	NDC1	TMEM63A	WWC1	MAT2A	HOOK3		PTEN	SERP1
		SRGN	BRPF3	ZNF664	NOD2	TULP4	PSME3	COX7A2L	ZNF451	TRIM56		NUB1	VPS4B
		UBE2W	PPME1	SYT4	LRRC8D	C12orf4	LRRC8D	SERPINE2	USP42	FLOT1		EWSR1	NPM1
		KRR1	BRD8	PTDSS1	HUWE1	STARD7	GNG12	CLTC	HRAS	SMARCA4		SLC35D3	NLK
		JADE1	HMGCR	NT5E	KIAA0907	UBN2	HUWE1	KPNA4	CYP3A5	B4GALT5		FOXO1	PTGES2
		SUFU	SOX11	ZNF507	RAD54B	BCL2A1	GLMN	CHPT1	USP24	BCL2L13		EMC1	TES
		CLCN5	CRK	GFOD1	LRRCC1	GATAD2B	SNRK	CHPF	RPS6KA3	UBIAD1		FNDC3A	ZNF217
		RNF6	BEND5	SMC3	CHSY3	RNASET2	PCTP	GLS	NFE2L1	ADAR		ABL2	DCAF6
		MOAP1	CUZD1	FOXK1	ERCC3	RAD54L2	STX3	PTPRF	ATP5E	SUGP2		H1F0	KLHL42
		SALL1	VPS13D	HLA-C	SAP30	AP3S2	HAND1	CTSC	SCOC	USP31		HECTD4	TAF1
		IVNS1ABP	FBXW7	MYH9	C19orf25	BAG6	KLF12	AMOT	TGFBR3	AFF1		ESRP1	ABCF2
		TADA1	DHX29	STYX	NMU	CDKN1A	ALG8	PA2G4	VMA21	COPS7B		AP3M1	KIF3B
		SGPP1	SSH1	MED13	Contig51328_RC (hsa)	ZNF92	RAD54B	LUC7L3	BTG3	MED21		ITSN1	CCAR1

				UBE2B	GATA3	CUL1	GPR180	UHMK1	RBFOX2	CNIH4	TMEM245	DCAKD		KIAA0895	C4orf29
				AKAP6	TRIM37	MACF1	XM_172878 (hsa)	IRAK2	KIF2A	LRRC6	ZCCHC14	PLAGL2		KATNAL1	CNIH1
				ZNF518A	COL12A1	SNTB2	SOAT1	OSBPL1A	ZBTB24	MAP3K1	CKB	TUBA1B		SLFN5	PHTF2
				KMT2E	UBR5	TMTC3	INPP5F	EDARADD	SAMM50	EMC3	AKT3	MTX3		PARD6B	RNF145
				XPO1	TRIM24	WIPF1	FAM101B	AAK1	NBEAL2	RHOA	37865	ZYG11B		SETD1A	KIAA1432
				SPTBN1	ARF3	HLA-B	IQGAP1	TNRC6A	CDCA8	PCDH10	USP14	MAST3		ZNF675	DDIT4
				CHD6	WSCD1	ATP11B	WNK1	SRSF11	ARHGAP31	LRIG1	ZZEF1	CDKN3		BRWD1	SFXN1
				AGL	SHC1	NUP98	RUFY3	OGFOD1	SLC6A8	ENSA	H3F3B	YIPF6		TBL1XR1	PPP1R12A
				ATL3	NDUFAF6	NRGN	PLEKHF2	AVL9	ZNF24	KIAA0513	STAMBP	SYK		GOLPH3	MBTD1
				SLAIN2	AFTPH	DCBLD2	GSE1	EIF4EBP2	RDX	FAM175B	UTP23	HOXD13		SLC39A1	DROSHA
				NUP205	ARMC8	APH1A	HOXB5	SESN3	IFT122	C1QBP	SMC2	RSF1		MTRNR2L8	KLHDC10
				C5orf15	ARRDC4	PAPOLA	DNAJB5	COPA	ZFP36	NCAM1	VPS33A	CHMP1B		PRRC2A	NECAP1
				USP33	NR2F6	UBL3	C16orf46	C17orf51	KIAA1549	RRP36	SPX	ZNF275		SLC30A6	BTG2
				IL10	SIRT1	ELTD1	B4GALT6	HNRNPA2B1	ZNF446	SLC25A42	GNAI2	XRCC6		SYVN1	DNAL1
				KMT2C	SIK1	MRPL37	IRX4	STAT1	WAC	SPG20	PIK3CA	BTBD10		SMG1	FBN1
				ITSN1	HOXB8	AGO2	IL7	HIPK1	C17orf49	PSAP	PFKM	ALDH3A2		KIF5B	JUN
				UGT8	MED14	OTUB1	ZFYVE26	YTHDF3	ASXL1	LYPD5	GALNT2	MAZ		SUMF2	PPP1R15B
				CASP14	THBS1	IL21R	SPATA13	PRKDC	WNK1	RPS6	ZNF407	HBS1L		MCF2L	RBMS2
				REEP5	STOM	ZSCAN29	IRGQ	RHBDD1	NME6	HSPH1	SCD	NAP1L1		MIER3	FBXO28
				TMUB1	NLN	ELP2	GPSM1	ZNF367	MNT	VLDLR	ZNF704	GCLM		DHX33	RNF4
				LYSMD3	RUNX1	ATPIF1	CDCA4	IFITM2	RUFY3	PLOD2	VEZF1	ZNF200		FAM134B	FANCM
				RANBP6	TMEM167A	MRPS16	SNX10	ZBTB32	TBC1D12	HOMER2	GOLGB1	SGPL1		PPP1CB	TXLNA
				C14orf159	IPO9	RRP7A	FRK	ADD1	GPATCH2	HEY2	ZBTB48	YTHDF1		TRIM28	GOLGA8A
				TMEM165	JAM3	PHF2	Contig52708_RC (hsa)	NUPL2	PBX1	IPO7	VPS37C	NEDD4L		ZBTB8A	N4BP2
				CBX5	ZNF521	CSDE1	C5orf34	ULK1	B4GALT6	NQO1	ABLIM1	YWHAG		EDIL3	FNDC3B
				RPE	RBM18	IRF2BP2	HSS00303609 (hsa)	ALDOA	ZFYVE26	SETD2	ZNF696	MCL1		RMND5A	NRXN3
				SLAIN1	USP6NL	TPRKB	CCT4	PTAR1	KIAA1671	CREB3	BCLAF1	VWDE		CYFIP1	CTNNB1
				PIAS1	VANGL1	RBM22	NHLRC3	MOB1A	TMEM41A	CORO1C	UHMK1	GNA13		MTRNR2L9	EFR3A
				IREB2	VPS8	MCFD2	NBN	MX2	SETD1B	FKBP1A	ZBTB14	ZDHHC5		CCL3	ZNF678
				SLC1A5	CENPM	LMNB1	LIMCH1	FAM111A	ARID1B	CHD4	NAPEPLD	CCND2		LIX1L	PTBP3

				PCBD1	NUP62CL	HMGB2	SUMO1	SAMD9L	PREPL	SMARCD1	HIAT1	RAB7A		ZFP91	PMM2
				RIMS2	GTF3C4	ALDH1L2	HSS00087671 (hsa)	SLC2A12	IRGQ	ZBTB41	AAK1	CELSR2			PLEKHB2
				PGK1	ZNF780B	CXCL12	CIT	VEGFA	SERPINB 9	MTMR12	ARF6	ADAT1			NR3C1
				BMPR2	PHACTR2	TOMM70 A	FBXO4	IRF7	MGA	SRSF2	TTC21B	SECISBP2			SSFA2
				TMBIM6	PIK3CB	GPATCH2 L	KIFAP3	GOPC	WNK2	MORF4L2	USP9X	BAX			PHLDB2
				ARHGAP 12	CDC25B	MACROD 2	ENST00000270 999 (hsa)	TRAM1	SMAD3	WDR26	MBNL1	CAPNS1			GIGYF2
				TOR1AIP 1	TRABD	ZNF761	SLC25A51	DNAJB4	PALD1	CCDC132	PDXDC1	NUDT21			KCTD2
				KIF5B	HDAC7	TJP2	PTBP1	ERRFI1	FRK	CCDC71L	ZNF813	PPIL1			MYLIP
				METAP2	ZMYND11	ZNF292	GFOD1	HNRNPUL1	BPTF	PRPF8	CHAMP1	ALDH1A3			ASPM
				HLA-DQA1	SRSF5	TNFAIP3	LIN9	LATS1	FBXO45	KLF5	ZBTB18	FOXM1			AXL
				GPR156	PYCR2	SLC25A5	BCL2	PPP1R1C	BACH2	MAT2B	SMG7	SMARCA2			NUFIP2
				KCTD20	AGO3	IPPK	CPSF6	PLEKHM1	TMEM13 8	TAOK1	HNRNPA2 B1	ADSS			CLTA
				MAP1B	DBN1	RAB21	MPZL1	IRAK1	LIPH	SH3BGRL3	MEMO1	RABGAP1			NCK2
				CD47	PHF14	WEE1	UBA52	DCAF16	TMEM43	CDR2	ZBTB4	ZNF260			LRRC58
				SYNCRIP	SASH3	SETD5	IDH3B	MAP3K7	NHLRC3	VASN	HMGCS1	NHP2			SLC4A7
				SPRY2	P4HA2	ZNF736	NAP1L4	SDCBP	MAP4K3	F3	ZNF606	SATB2			ADM
				CUL4B	NCOA7	CEP85	ZNF320	NUPL1	SMARCA D1	NKX6-1	ERGIC1	EEA1			POGZ
					TBC1D9	RCOR1	NM_052864 (hsa)	RAD54L	GMNN	ADIPOR2	CACNA2D 1	HNRNPDL			SACS
					PAXIP1	SHCBP1	Contig58637_R C (hsa)	METTL7A	ATPAF1	RFX7	YTHDF3	RAB6A			GDAP2
					IMPDH1	IVD	MED13	CD84	OCLN	HK2	ZBTB34	HOXA9			DMXL1
					GLRB	BLMH	TUBB	ITGB2	PTBP1	VPS13A	TMEM12 7	TM9SF3			DNAJB9
					AMPD3	CD164	PPP2R4	SESTD1	ERLIN2	ABI2	ZNF367	DBT			TRIP13
					SMIM8	WRB	MACF1	TMPPE	KIAA089 5L	AGPAT5	ZDHHC17	RC3H2			DNAJC13
					ATAD2	DDX17	hCT1640801.2 (hsa)	ZHX1	ARID3B	ANK2	ZNF3	FHOD3			DPY30
					LYAR	ZNF627	XPOT	KMT2C	LDB1	TMEM145	ZNF106	CFDP1			LNPEP
					TMEM248	KARS	TMTC3	ATP5G2	DBF4	ATP2B1	FAM134A	AUTS2			FOXN3
					PRKCI	BTBD3	EME1	CPNE8	TNFRSF1 1A	FAM96A	ABCC5	ZZZ3			IRS4
					C11orf30	EGFR	MAP1LC3B	SLC25A33	SELT	RCN2	DYRK1A	SNX30			MYC

					EDEM2	ASPHD2	Contig39983_RC (hsa)	FNBP4	TMED5	NCL	ABHD17C	STK38			QKI
					HLTF	TSHZ3	PTPN1	TFF3	RNF111	PROCR	SPIN1	KAT6B			UQCRC2
					C10orf54	UGP2	HIPK3	RAB20	PHB	Bim (hsa)	CARF	MTMR4			UBE3C
					KIAA1211L	TMEM41B	DCBLD2	RSAD2	NDST2	KLF10	XRCC5	GMFB			IKZF5
					AKIRIN2	ZFHX3	DLG5	SNTG2	OSTF1	SNPH	B3GNT2	PMEPA1			HNRNPC
					TPR	GNAI3	UBL3	STK40	CEP120	MPP5	UNC13B	ROCK2			IQGAP2
					ANKRD52	ZNF195	CCDC14	BRWD1	ZNF516	INSIG1	RBM33	ARPP19			CSTF1
					PRRC2B	HSPD1	NM_015393 (hsa)	RAD9A	MED13	ZNF571	ZNF652	ZNF551			PLK2
					NSUN5	SULF2	AGO2	ZNF410	TUBB	TNRC6B	TRIM25	PDE3A			VEZT
					SAMD8	LTBP3	RAB2A	HYOU1	MACF1	NLE1	ACADM	CARD8			MTMR1
					PVRL2	NACC1	C1orf86	IREB2	BCAR1	FYTTD1	ZNF841	OTUD6B			DEPDC1
					TNXB	SERPINE2	AXIN1	SLC1A5	CCNF	ZNF540	TRMT6	ARPC3			INSIG2
					THBS2	CLTC	TMEM180	TFG	TMTC3	B3GALNT1	ATP13A3	ETS1			C21orf91
					SPATA2	KPNA4	MCFD2	MAML1	DDX6	LITAF	CHD9	NRBP1			TMEM9B
					MTUS1	ZNF580	MIB1	ZC3H10	NUP98	DDX21	ABL2	DPP8			COL5A1
					REST	NOTCH2	GTPBP1	TNFAIP8	SLC16A10	PRRC2C	CXCL10	ZNF45			GPBP1L1
					ELK3	MT2A	SLC7A11	PRRC2A	KIF13A	LPAR4	RPL27A	CPSF2			KLF3
					UAP1	GLS	LMNB1	RAPH1	APH1A	EGR3	TET2	PSME3			CPNE3
					TRAPPC2P1	LUC7L3	hCT1971266.1 (hsa)	MDM2	PAPOLA	PANK3	MGAT4A	PKM			CTSZ
					SETD7	NOP56	SP3	RAPGEF5	NR4A2	CEP19	HOXA10	LRRC8D			SKIL
					PITHD1	PARP14	WNT3	ABL1	CAND1	RBBP6	ITSN1	HUWE1			ZAK
					LMNB2	MAP3K1	ZNF384	TMBIM6	FAM127B	SRSF1	KATNAL1	TCF12			IFI16
					TPP2	NHLRC2	MIPOL1	EIF4G2	DCUN1D4	MYNN	WNT7B	FRS2			CCDC47
					CAV1	ZNF224	BRCA2	MAP3K5	KLHL5	AMMECR1	ZNRF3	GNL3L			SRPR
					CEP128	HAS2	Contig41377_RC (hsa)	CARHSP1	LIN28B	DACT1	ZNF623	BCL9L			ARNTL2
					LRRC8B	ARFGEF2	Contig40026_RC (hsa)	DNPH1	IMP4	HDHD1	ZCCHC24	TOM1L1			ATRX
					GTF3C3	TIMP2	KCNH1	TRIM26	KIAA2018	ZMIZ2	TOP1	STX3			GLO1
					B4GALT5	ETF1	POLD1	SYT12	STAG2	MTRNR2L4	RBM4	HAND1			PRPSAP1
					SEL1L	RNF138	KIF18A	NACC2	ZHX1-C8orf76	ZFP82	PIAS1	GOLGA8B			FOPNL
					ZNF431	PCDH10	TJP2	CDC25A	PRKAR1B	SLC35F5	DHX36	KLF12			BTK
					KCNN3	SEC61A1	AF085923 (hsa)	OCN (hsa)	HNRNPA0	ITSN2	HYOU1	GLE1			SYT1

					CPEB4	LRIG1	RAB21	KIAA1586	GRHL2	TRIP12	PTPRS	CAMSAP2			AARS
					GAPDH	CACUL1	CDK1	ABRACL	CSDE1	CHGB	GRSF1	ERCC3			SMURF1
					ADAR	MIEF1	FAM71E1	KDM5B	TTC17	DCAF7	EIF1	DIP2B			APLP2
					SUGP2	TRAF5	UVRAG	ZDHHC13	LPAR2	CDH10	RPL15	LCOR			GIPC1
					USP31	ANXA2	AGFG1	TXNIP	SLC38A7	CCSAP	RAPH1	MCM4			CNNM4
					DPF1	CHD7	NFYA	HIST1H4B	MIB1	PRPF3	ZNF629	LAMTOR3			SLC11A2
					SPTLC2	RPP30	TNFSF13B	CPNE1	PRDM11	HECTD1	SCML1	ACLY			SLC38A2
					AFF1	TJP1	IGF1R	RASSF5	MEF2D	KIAA1524	STK24	GOT1			MIA3
					MCM6	ZNF765	TPGS2		HAUS1	MAPK1IP1L	ATF4	SLC6A8			CERK
					PLAGL2	SPG20	CD164		ALDH1L2	ACVR2B	ZNF141	FAM69A			XYLT2
					ZDBF2	C18orf25	TIFA		GOLT1B	MTR	EIF4G2	FZD6			JMJD1C
					DIDO1	CD2AP	Contig51846_RC (hsa)		ADRB2	ZBED1	AP2M1	GPR180			NOL4L
					ZNF346	NARS	PABPC4		RNF217	PCGF5	XIAP	PPP3R1			GSTP1
					WFS1	HEBP1	EEF1A2		ZNF292	FIS1	NPTX1	RDX			SMAD5
					MTX3	PRSS23	XM_210054 (hsa)		THRA	SERP1	TMF1	IFT122			CLK1
					PIK3R4	PSAP	AB002443 (hsa)		ZNF587	C9orf41	CD69	TOB1			TRIM33
					TSC22D1	GOLM1	LTBP3		NOD1	PRRT3	NACC2	ADAL			MMD
					NR6A1	ENDOV	GEN1		ABCC10	STK38L	XPO5	ZFP36			ITPR1
					C10orf10	MTMR2	RAB23		WEE1	EIF2AK2	C2orf69	ZNF117			MCMBP
					HHIP	ZNF85	TARDBP		PCNXL3	FAH	ZCCHC3	HINFP			CFL2
					DLG1	QSER1	ZNF226		INPP4B	DIP2C	TXNIP	ENTPD6			MEPCE
					PHGDH	STT3B	Contig25707_RC (hsa)		NUP160	LARP4B	TYW1B	WAC			HYPK
					RSF1	HNRNPU	CLTC		SHCBP1	SH3BP4	ZNF2	IQGAP1			AP1S3
					TAF2	THAP6	SPCS3		ZBTB26	HNRNPUL2	TUBGCP5	UNC45A			BAZ2A
					E2F8	SPRED2	NOTCH2		EPN2	MAL2	ZNF792	DNMT3B			PRDM15
					CHD8	KCNK3	ZNF519		EIF4EBP1	PHTF2	GTF2I	GAN			CPEB2
					PURA	VLDLR	SERAC1		IGF1R	CCDC43	ZNF28	PSME1			TMCC1
					GCA	MAGED1	CHPT1		ASB13	OSTC		ARHGAP32			RORA
					ODF2	GXYLT1	SLC9A5		PIK3C2B	ELAVL4		ASXL1			GLCCI1
					ZNF275	CD9	HSS00386249 (hsa)		KARS	KIAA0286 (hsa)		WNK1			PRKAR2A
					TMEM110	CNOT8	hCT1788782.2 (hsa)		GMPPA	FZD4		PARK7			IPMK
					IST1	EZH2	ARHGEF10		TLE1	SFXN1		GPATCH2			NAA30
					ELN	CDK6	AL133577 (hsa)		PSME4	CUL5		PPRC1			METTL10
					40787	EPDR1	ERCC6		KPNA2	AEBP2		PBX1			KLHL15

					MBOAT7	HEG1	NOXO1		TMEM41B	RPL35A		CRIPAK			FXR1
					MRPL51	IER3IP1	NOP56		ZFHX3	ZBED6		ZFYVE26			METTL21A
					PAFAH1B1	MAP3K9	PPP2R2D		ARF1	MIRH1 (hsa)		SLC6A10P			METTL6
					OGT	NISCH	BMP6		FAM174A	TYMS		SETD1B			LYST
					IDS	ACAA2	AF131784 (hsa)		SERPINH1	AGO1		MRPS33			EIF5B
					RPS15A	GRTP1	NHLRC2		SPCS3	IFT46		METTL4			CUX1
					IFITM10	FYN	ASPH		GLS	SGMS2		PAIP1			PTPRJ
					PRPF40A	WBP2	ATAD5		VCL	DLG4		ZNF562			XRN1
					ZNF597	DDX3X	FEM1C		CTSC	RAB3IP		SRCAP			H3F3C
					NDUFS4	PDIA6	LRIG1		GNB1	POLR2A		TRPS1			C6orf62
					NAP1L1	SETD2	CDKN2A		AMOT	DNAJB1		FNBP1L			SRC
					SOWAHC	OXSR1	RTKN2		VAPA	CASP3		SLC36A4			ANKRD13C
					HTT	SERINC1	HSS00075325 (hsa)		GNL1	TMTC4		SNRNP48			STC2
					CGN	CNNM2	TAF6		ERCC6	JUN		CCDC28A			NOLC1
					SGPL1	RPIA	CHD7		TXNDC11	ZNF711		FBXO45			FGF2
					SZRD1	ZNF189	HADH		MAP3K1	DCP2		NET1			POP1
					TRAF7	CORO1C	PGM3		CXCL5	SLC24A3		BACH2			TMEM222
					FAM46C	MAP3K11	UBE2J1		MFNG	FAM208B		SMEK1			ZNF672
					APOF	MCM7	APPL1		TLR4	TOP2A		EXOSC10			MCUR1
					HYI	SLC16A9	CCDC7		ABHD14B	TIMM17A		NHLRC3			SNAPC1
					NEDD4L	PHF10	SPG20		XBP1	PBXIP1		ZNF764			LUM
					YWHAG	FBL	TMEM164		EIF3B	JAK2		TMEM64			PIK3C2A
					PPP1CC	TCP1	MAP6D1		FIZ1	GRPEL2		NBN			RDH10
					MCL1	HACL1	RAB1A		B3GALNT2	LHFPL2		LAMA1			DSTN
					DUSP6	CALM2	PRSS23		PARG	PPARA		RPL7L1			PUS7
					ZDHHC20	CHD4	HSS00309444 (hsa)		PRDM5	EPHB4		FAM110B			FAM199X
					MAPK14	MTMR12	HIC2		ZNF768	PNMAL1		MAP4K3			TXLNG
					DUSP15	ASF1B	GAS2L3		MGLL	PUM1		ZNF664			TMEM2
					KIAA1191	ZCCHC2	ZNF85		CACUL1	ACTR3		GMNN			ADCK3
					NUP133	RPS20	HSS00009331 (hsa)		ZNF526	ANXA7		BCL11A			SSX2IP
					GNA13	ZNF107	KDM4D		GTF2B	SPEN		GOLGA5			PSMD12
					CELSR2	EPC1	HESX1		MIEF1	RAE1		URB2			PPP1R12C

		TNIP1	ADD3	COX5A		CHD7	CTNNA1		BCL2			CALD1
		TYK2	TANK	XM_210341 (hsa)		H6PD	FAM89A		CAT			ZC3H11A
		BAX	PRPF8	GLI4		TJP1	SEC24C		APPBP2			LCLAT1
		TBX3	CELF1	ZBTB6		KDM2A	PHLDB2		PDXK			CDK13
		ZNF639	POU5F1B	GXYLT1		PGM3	STRBP		TTF1			SUV420H1
		HSP90AA1	VHL	GID8		PMP22	XRCC6BP1		CDKN1B			HIVEP1
		TRMU	KLF5	RET		UBE2J1	NCOA6		SRGAP1			COL1A2
		SMARCA2	EPAS1	EPDR1		DDX28	NONO		RNF111			RBPJ
		ADSS	MARCH7	KIF23		ZEB1	ASPM		FOXK1			FNIP1
		RALGAPA2	LBR	ATP6V1G1		SLC30A7	CBWD3		SYNJ1			NPTN
		MKL2	CKAP5	SLC25A40		GALK1	CDV3		RNF144B			RASSF3
		SATB2	TAOK1	DDX3X		STT3B	MXI1		PHB			ARRDC3
		PXN	ABHD2	SETD2		TAB2	AXL		GNPDA2			IKZF2
		MARS	CANX	USP54		HSDL1	SAMD4A		OSTF1			LAMC1
		TM9SF3	SIKE1	SNRPD1		PPP1R18	NUFIP2		HBP1			FOXJ3
		PNPLA6	ADIPOR2	HMGN1		SPRED2	GCC2		TRAF4			FAM160B1
		RC3H2	CLPTM1L	AL049354 (hsa)		PNPLA8	WBP1L		IMMT			ARL5B
		FBXL4	CAPRIN1	INIP		EP400	CRABP2		PRR26			ZCCHC14
		ORC1	KIAA1109	ANKRD32		VEPH1	PAICS		NADK2			ZBTB7A
		CFDP1	RFX7	COX16		GXYLT1	HIF1A		SNTB2			ZNF780A
		TASP1	ABI2	PHF10		WDR6	PPP1R10		ATP11B			REV3L
		ZZZ3	NCOA3	C1orf112		SPRYD7	MT1X		PHACTR4			CELSR3
		EHD4	NFX1	CXCL2		PLOD2	ANP32B		ANKIB1			FBXW2
		MTMR4	REXO4	DENND1A		IGDCC3	UBE3A		PTPN1			GOLGA1
		HNRNPF	ZNF622	Contig43667_RC (hsa)		COL3A1	KTN1		HIPK3			TMEM50B
		GMFB	LPP	TMEM106A		HOMER2	RIT1		RPL38			YIPF5
		ROCK2	IFNAR2	FNBP1		FOXN2	CGGBP1		APH1A			BMPR1A
		FAM198B	YWHAQ	ZBTB41		FYN	ZFHX2		PAPOLA			CEP135
		ARPP19	RASSF2	MTMR12		DDX3X	MKRN1		UBL3			OSGIN2
		NABP1	PIP4K2A	MSN		NAA40	STAU1		ATP6V1D			CREB3L2
		VGF	QSOX2	BARHL1		PARP9	DLC1		AAGAB			ATXN7
		KATNBL1	FAM222B	HRH1		IGF2BP1	KCTD10		CAND1			TSPYL1
		ZNF148	SAFB	DEK		MDH2	PDE11A		GALNT11			VPS41
		SKA2	SNX14	SREBF1		FLAD1	MYC		FAM127B			RABGEF1
		PTTG1IP	USP28	XM_212432 (hsa)		PARS2	QKI		OTUB1			H3F3B
		NRBP1	KLF10	BC031660 (hsa)		SLC35A4	LIMK2		MCM5			C11orf24

					CEP78	SLC25A1	Contig66615_RC (hsa)		COPB1	PAQR8		LIN28B			G3BP2
					LMBRD2	ZNF680	ATG4A		OXSR1	NPPB		DPCD			ANO10
					SH3BP2	MPP5	SLCO3A1		SERINC1	CENPF		IL21R			ZNF730
					PTPN14	KIF21B	CELF1		FSTL1	COX20		GLUD1			CCDC88C
					DYNC1H1	TMEM181	DBF4B		MARK3	TMEM178B		AFF4			DDX19A
					ADCY3	NHP2L1	ALG10		HERC1	FAM83D		NT5DC3			SNX13
					VARS2	INSIG1	NREP		SHMT2	PLK2		ATPIF1			CDC27
					HS2ST1	TNRC6B	TDG		CORO1C	VEZT		UTP20			LONRF2
					ARHGAP17	SPTY2D1	FADS1		SLC16A9	TBX19		MRPS16			SLC6A9
					PKM	B3GALNT1	ANG		TWIST2	PAFAH2		HNRNPA0			ACSL4
					GNG12	RTCB	DST		FBL	RPGRIP1L		NAMPTL			SDC2
					EFTUD2	LEMD3	hCT14363.3 (hsa)		MID1	KHK		TTC17			PAPD7
					MAP4	NUAK1	HSS00061034 (hsa)		AP1M1	USP46		RBM22			SREK1IP1
					SF3B4	DDX21	CANX		FKBP1A	DNAH11		FBXL2			PCMTD1
					AGRN	UBB	DPP9		PREP	INSIG2		USP13			AURKA
					FRS2	DCTN4	HSS00209806 (hsa)		SMARCD1	SRSF6		MCFD2			LEPREL2
					SNRK	ASH1L	Contig54503_RC (hsa)		FZD3	SLC39A11		PIGH			DNAJB12
					CTC1	TOP2B	SPIN4		CPD	C12orf29		LMNB1			NUP43
					NPAS3	IMPA1	U29115 (hsa)		MSN	HOXA3		ZNF302			KMT2A
					ITCH	HIST1H4F	MTERFD1		ZCCHC2	SUB1		MEF2D			SCAF1
					GOLGA8B	PANK3	RFX7		ACTA2	PTP4A2		HAUS1			NSF
					ALG8	VARS	AL137310 (hsa)		CCDC71L	EIF3A		MTMR9			UBAP2L
					STXBP5	SRSF1	NCOA3		TANK	POC1B		BCAS1			MED19
					RBFOX2	SPECC1L	NUDT15		NAA15	C4orf3		TNKS2			ALDH9A1
					KIF2A	PURB	Contig38566_RC (hsa)		PRPF8	PLEKHA3		SRM			CTTNBP2NL
					LOXL2	KLHL11	HAUS3		CELF1	SMAP1		MRPL48			RLF
					MESDC1	SNX8	CHM		FBXL7	GOLGA4		MOB1B			MECOM
					LTBP2	HIF1AN	AGPAT5		HS6ST3	GPBP1L1		SERINC3			ZNRF2
					ZNF121	SENP6	Contig57398_RC (hsa)		MARCH7	CDCA7L		C19orf12			TULP4
					ANKRD50	DCBLD1	GIGYF1		XPNPEP1	MNS1		CDCA5			GATAD2B
					EVA1C	SDHAF2	ANLN		LBR	CLDND1		TRMT61A			CTCF
					MDFI	EEF1D	RHCE		TUBG1	PRDX1		POLD1			COG3
					LCOR	ZNF266	NSD1		TAOK1	STK35		KIF18A			RNF182

					RBM23	COX19	LRRFIP2		SLC35A2	DOK7		CCDC117			KIAA0232
					SLC25A25	TRIP12	PHTF1		ERGIC2	CYP2U1		IPPK			SYNE2
					ACLY	BCAT1	NAB1		FADS1	FBXO8		CCSER2			CLSPN
					ARHGAP31	SIPA1L1	MT1F		C4orf46	MTPN		RAB21			ITM2A
					SLC6A8	AKAP8L	AK021796 (hsa)		DST	CPM		CDK1			DOCK9
					KIAA0319L	AZIN1	UNG		AMMECR1L	SLC35A3		GON4L			RBM10
					PAPOLG	ZNF253	ALCAM		HES1	SKIL		SFT2D2			USP9X
					ZNF24	SORBS3	KLF10		PLAGL1	TFRC		LRP12			MBNL1
					SOAT1	COL6A1	CCDC121		CDR2	BIRC6		NFXL1			ZFYVE21
					SAMD9	FLII	MPP5		RALY	FLNB		CEP85			NFIA
					TOB1	FRAT2	SMARCB1		KBTBD6	VIM		RCOR1			RP2
					ZFP36	MED4	USP12		DPP9	BMP2K		EPN2			ATP2B4
					KIAA1549	CCSAP	INSIG1		PDS5A	ABHD4		BLMH			CNEP1R1
					PLEKHJ1	SLFN11	PPP1R26		INTS12	SPAG9		CTDNEP1			MRPL44
					ARHGAP32	FAM122B	SOCS6		MICB	SUPT6H		MSL2			ZBTB18
					WNK1	PRR14L	PRADC1		NCOA3	FOXA1		PLXNA1			GFPT2
					MNT	CHPF2	GCH1		NUPR1	SETD8		IGF1R			MBNL3
					RUFY3	HECTD1	TNRC6B		LIMD1	REL		CD164			SESN3
					ZWINT	PRTG	EMC7		MTUS2	CEP350		OSBPL3			SLX4
					GSE1	ATG14	EXOC4		DPM2	MBD2		LSM14A			HNRNPA2B1
					GPATCH2	NCK1	B3GALNT1		EIF3K	BRAP		CENPQ			EIF1AD
					PHF20	NUMA1	ALDH5A1		SLC41A1	ARNTL2		MFN2			PER2
					PBX1	HDDC2	hCT1957917.1 (hsa)		GIGYF1	ATRX		DDX17			NUP155
					CSNK2B-LY6G5B--991	APAF1	GLTSCR1L		LRRC42	EIF2S3L		RAB12			TPT1
					KDM3A	PAPD5	ZXDC		ATP2B1	CBX6		MIPEP			LRP3
					SPATA13	ASAP1	PHF6		PLEKHA8	TXNDC4 (hsa)		ZNF627			NDUFS5
					TPST1	ZDHHC16	FAM178A		TSPAN11	TTC3		ARMC6			TUB
					SNAI2	MAPK1IP1L	PRRC2C		RPS6KB2	XPR1		PABPC4			PHF13
					METTL4	ACVR2B	LIMS1		ATP5B	LRAT		KARS			FAM214A
					IRGQ	MT1E	HSS00321855 (hsa)		FAM222B	HOXB6		DOCK11			ATP2A2
					PAIP1	DPY19L4	RICTOR		TMEM147	ZNF614		BTBD3			PCBP2
					ZNF562	SLC7A1	ASH1L		CHEK1	DUT		C6orf211			ZBTB34

					SRCAP	ZNF721	TYW3		MYCBP2	SNRPA1		MTHFS			AP1G1
					TRPS1	KRIT1	IMPA1		ALCAM	TMEM131		ODC1			LEPREL1
					MGA	ARHGAP35	CEP19		SLC16A3	SET		CSTB			PAPSS2
					CDH11	PTK2	HS3ST3B1		KLF10	TTC1		ZYX			RAB8B
					FNBP1L	ST3GAL5	C10orf35		ARHGEF1	SMURF1		GADD45A			UBE2W
					DCAF5	ARID2	PURB		CPSF1	WASF2		EVI5			RAN
					B3GNT7	SLC9A3R1	USP45		ACP1	GDE1		ZFHX3			HERPUD2
					LONP2	LMLN	AMMECR1		ZNF57	N-cadherin (hsa)		HSPD1			DUSP5
					SOD1	IDI1	NMI		C15orf39	HSP90AB1		C12orf43			JADE1
					BPTF	CNN2	TMEM16F (hsa)		ALOX5	ARHGDIA		FOXQ1			OSBPL8
					FBXO45	SOGA1	IQCC		FAM104A	C11orf58		MSANTD4			SUFU
					NET1	MAVS	Contig659_RC (hsa)		DFFA	BAZ1A		FAM64A			ANAPC13
					C5orf34	FURIN	BC033326 (hsa)		MAML3	NCOR1		CLTC			ULK1
					NHLRC3	TES	C4orf32		TP53	AHCYL1		VAMP1			DYRK1A
					TMEM64	TICRR	RPAP3		TNRC6B	EIF4G1		SQSTM1			ZCCHC11
					CIT	STK38L	MIS18BP1		SPTY2D1	CEP44		CHMP2A			SNX25
					EYA4	ETNK1	RALGPS2		LRIG2	BATF2		VCL			LEPROT
					C7orf43	IFIT2	TMEM66		ADAMTS12	PDK1		SFMBT1			PTAR1
					CCT5	WSB2	TBC1D9B		GPR4	RNMT		DHX40			STEAP3
					MAP4K4	GPR125	SS18		NUAK1	CHTOP		VAPA			CAMK1D
					SYT4	ZNF217	TRPV1		PPP1R11	GBAS		ZNF75D			ITGA6
					SLC25A51	ITGAV	SEC63		PRRC2C	FOSL2		PA2G4			MAP2K4
					C18orf32	HSPA12A	POLA2		C9orf69	CMTM4		MAP3K1			SAMD9L
					PTBP1	DCAF6	CCSAP		ZSWIM6	SFPQ		EMC3			RPS6KA4
					ERLIN2	FAF2	ARL4C		GAA	NCOA2		XBP1			XRCC5
					GFOD1	NUP214	ANKRD40		RBBP6	PABPC1		RNF138			DKK3
					PEX5	PSD3	Contig398_RC (hsa)		ACSS2	MR1		ITM2B			RSBN1
					LIN9	NAA50	NM_030941 (hsa)		FADD	CFL2		NDRG1			GTF2A1
					CAT	PRDX4	PAPD5		GMFG	DHX8		ASPH			TANC2
					AFAP1	SDF2L1	hCT1971393.1 (hsa)		ZDHHC12	KLHL7		B3GALNT2			CPEB3
					APPBP2	NCAM2	ACVR2B		HIF1AN	RNPEPL1		SPATS2L			MTF2
					PDXK	KIAA1147	CHP1		PRDM1	ERP29		BCKDHB			FDPS

					SULF1	TRIM59	RAB5A		SLC35F5	CHIC1		WASF3			SGK3
					LRP11	ZFYVE16	LAMTOR5		DNAH1	SKI		CACUL1			ATL3
					U2AF2	MAL2	PCGF5		ARIH1	DHX38		ENSA			LIN54
					NNMT	DOCK10	KRIT1		SDHAF2	RBBP8		MIEF1			LATS1
					WNT2B	CSE1L	MPHOSPH1 (hsa)		ZNF432	ARPC1A		KIAA0513			NFIC
					ECHDC1	ZNF676	MYO1D		DPF2	GAS8		FAM175B			BMP7
					TMED5	MYD88	CKAP2L		GSR	YWHAZ		ANXA2			TRMT6
					MAP2K3	PHTF2	SLAMF7		C4orf32	YLPM1		FDFT1			RAD17
					RNF111	NDE1	SLC25A30		ITSN2	AKAP7		TJP1			C16orf62
					SYNJ1	CDH1	LMLN		TMEM132B	NXT2		ZNF330			RNF141
					RNF144B	GPR37	IDI1		SLC16A11	PRKX		ANAPC1			TRIB1
					PHB	KEAP1	PRICKLE1		VAV2	SLC25A36		ZNF280C			MAPK1
					CEP120	FZD4	Contig50189_RC (hsa)		POSTN	MOSPD3		CD2AP			MAP3K7
					STYX	PDIA4	NLK		DYRK2	ZC3HAV1		MORN4			ELP3
					TPBG	PPP1R12A	CCZ1B		H2AFV	PPP2R3C		SLC30A7			YPEL5
					AP4E1	LIPT2	LIPA		TEX2	FIGNL1		KIAA0196			LATS2
					MED13	AEBP2	TICRR		SBF1	RAP2C		PRSS23			CREBBP
					MAP7D2	TET3	IL15		GYG1	SRRM2		QSER1			CHD9
					BIRC5	CMIP	NA		TBC1D9B	TMEM74B		TAB2			PRDX3
					CUL1	SLC5A3	PISD		ZNF436	ZNF404		MRPS14			TRAF3
					PPP2R4	RIF1	TMEM67		TSG101	PEX11B		ALDH18A1			USP33
					MACF1	SF3B3	CYP4V2		FBRSL1	SENP1		HSDL1			ZHX1
					XPOT	AGO1	XM_211811 (hsa)		ZNF589	JAZF1		COQ9			RAD21
					SNTB2	SND1	Contig50129_RC (hsa)		RBMS1	DDX5		RAB31			PCGF3
					TMTC3	TBC1D4	NCEH1		UTY	STAP2		RCAN2			SLC40A1
					PAPPA	MAPK9	ENST00000282683 (hsa)		YME1L1	RIMKLB		VLDLR			DNAJC27
					DDX6	EIF2A	TPM4		SZT2	XRN1		GPX7			KMT2C
					ANKIB1	PCNX	BRCA1		DAP3	IRS1		SPRYD7			CCNG2
					PTPN1	C17orf89	ITGAV		TRIM41	FAM3C		CDK6			GOLGA3
					PPTC7	KHSRP	AJ227908 (hsa)		SLCO5A1	NIF3L1		HEG1			VAT1
					HIPK3	POLR2A	HSS00388495 (hsa)		EZH1	TMEM170 (hsa)		CNTNAP2			MPP3
					KIF13A	DNAJB1	EHBP1		PRR14L	PEBP1		RNF135			SLFN5
					RPL38	CASP3	NAA20		C7orf73	NOLC1		TFAP2A			PDHB
					PAPOLA	POLR2M	SH3BP4		CHPF2	SHOC2		ACAA2			KLF6
					PPP6C	ORMDL1	OBSCN		ARL8B	CDC42		COL3A1			PFKFB4

							HECTD1	PMPCB		FOXN2			SLMO2
							PRTG	KMT2D		FYN			KIAA1279
							PNO1	ADAMTS1		STAT5B			ZNF470
							APAF1	TMEM106 B		DDX3X			CBX5
							ASAP1	CCDC152		NAA40			KLF4
							CRAMP1 L	LDHB		INO80C			STARD4
							RNF10	SESN1		IGF2BP1			TTPAL
							GREB1L	IQCE		SETD2			JMY
							EDEM3	SNAPC1		OXSR1			TBL1XR1
							OAS1	GSC		LIPT1			WDR37
							ELF4	ANKRD46		DNAJC9			ZNF91
							OXA1L	RHOQ		CNNM2			BA22B
							FAM69B	ZRF1 (hsa)		MARK3			DUSP1
							UBA6	EBAG9		RPIA			SPOPL
							SLC7A1	TMED4		WBP5			ASXL2
							HMGXB3	TSNAX		ANKRD32			SLC1A5
							SLC2A3	KIAA1212 (hsa)		SHMT2			IRF1
							SLC9A3R 1	TSC22D2		CORO1C			ADRB1
							MRPS7	ELAVL1		MCM7			PSMB6
							C8orf58	DYNLL2		PHF10			MDM2
							FOXP1	AP2A1		TCP1			PRR11
							NLK	HSPA8		GSTO1			KIF24
							SOGA1	USP8		FKBP1A			TMC7
							XPO7	KDSR		PRKCE			FANCI
							CEBPB	EIF2S1		SMARCD1			VPS54
							EIF2S3	CASC5		FZD3			CNDP2
							ZBTB37	ZFP1		FNBP1			PNISR
							IL15	RARS		MTMR12			EIF4G2
							LYPD3	ARID5B		MSN			BCL2L11
							PEG10	EMC2		FARSA			CLOCK
							PPP1R16 B	CLDN1		SLC22A5			ZNF703
							ARHGAP 18	PIGB		SLC35C2			KIF5B
							PISD	KLHDC8B		MORF4L2			MTF1
		UBL3	ARFGAP2	NAA50									
		CCDC14	NSUN7	KLHL42									
		BOD1L1	STX16	RSRC1									
		DCUN1D4	ZNF711	STK17B									
		DNAJB6	PIK3R3	CCDC86									
		AGO2	MCM3AP	PRPS2									
		WDR46	ZDHHC2	TMEM17									
		GLUD1	WDR33	LYPD6									
		AFF4	ABHD17B	TRIM59									
		IBTK	DCP2	CENPK									
		UBE4A	RNF4	TMX4									
		ADH1B	FANCM	DDX24									
		UTP20	GOLGA8 A	hCT1831253.2 (hsa)									
		AGAP1	FAM208B	TRAK2									
		ITPKB	TOP2A	NM_018590 (hsa)									
		MRPS16	CCND1	CRTC2									
		PRKCD	IMPAD1	PHTF2									
		RRP7A	ASTE1	NDE1									
		N4BP2L2	TSSC1	hCT1971392.1 (hsa)									
		CSDE1	PFN2	AK092985 (hsa)									
		IRF2BP2	MGP	SP4									
		RASIP1	SMAD4	RAB11FIP2									
		PRUNE2	FMNL3	GPR37									
		MYO1C	FNDC3B	CENPA									
		SLC7A11	ICMT	PDHA1									
		LMNB1	CTNNB1	FZD4									
		RHOBTB1	ZNF678	ENST00000295 064 (hsa)									
		HMGB2	PUM1	SMC4									
		ALDH1L2	LPAR1	ZFX									
		TNKS2	B3GNT1	CUL5									
		TOMM70A	PTBP3	RAP1GAP									
		PRPF19	VTI1B	ZFYVE28									
		MOB1B	PPT1	PPP1R12A									

					GOLT1B	VRK3	SLC35F3		SRRD	JAG1		WDR26			HOXC8
					MAN2A1	MSH6	APOLD1		SF3B5	FAM98A		TRAM2			SOBP
					YWHAB	CTNNA1	SLC5A3		CADM1	RBPJ		EPC1			KCTD20
					TMEM68	GFPT1	AF086431 (hsa)		POLR2D	C11orf54		ZNF770			NPTX1
					RNF217	PLEKHB2	SF3B3		YARS2	WDR89		C18orf8			ATR
					CCDC117	SEC24C	Contig47985_RC (hsa)		MN1	UCHL3		PRPF8			CTGF
					ZNF292	HECTD3	AP1S2		ITGAV	EXOC6		ZNF322			MIER3
					TNFAIP3	FARSB	MAPK9		FAF2	ZNF451		HPS3			CD69
					IPPK	NRD1	SGMS2		GGA3	DDEF2 (hsa)		CELF1			TTC9
					THRA	SBNO1	VGLL3		SALL3	YBX1		PDS5B			BSN
					ACAP2	NCOA6	KCNS3		NUP214	BBS10		VHL			FNIP2
					WEE1	NONO	DLG4		NAA20	STX6		ZNF337			SYNCRIP
					SETD5	RBM17	PKP4		RASA1	ILF2		EPAS1			ASAP2
					SFT2D2	MAN1B1	STAMBPL1		KLHL42	D15Wsu75e (hsa)		MED12L			DSEL
					PHF3	ARL6IP1	LARP1B		MLLT1	PEX13		LBR			MRPS21
					NUP160	USP5	ERMP1		HNRNPUL2	USP24		TUBG1			NEFL
					RCOR1	IFFO2	POLR2A		KIAA1147	FOXJ3		TAOK1			HNF1B
					PPP2R2B	PNP	Contig29802_RC (hsa)		DMWD	ERC1		HIST1H2BD			TAF15
					EIF4EBP1	COL18A1	CENPI		PGM2L1	LMAN1		DST			
					PLXNA1	SAMD4A	TMEFF1		MAL2	TGFB2		KIAA2026			
					LONRF1	NUFIP2	RPAP2		LRRC59	CYP20A1		AMMECR1L			
					IL24	TSR1	DNAL1		PTCHD1	MTRR		ZNRD1			
					MINK1	SEC23B	FBN1		DMD	METTL2B		PLAGL1			
					ASB13	CLTA	TMTC4		CNIH1	PROSER1		FKBP9			
					PIK3C2B	LRRC58	HSS00305022 (hsa)		THG1L	POLR3G		CANX			
					MFN2	COL4A2	SLC16A14		RNF145	MYT1L		CDR2			
					DDX17	SLC4A7	GPC4		CERS2	MGEA5		CENPL			
					RAB12	HIF1A	B4GALT1		URB1	CHSY1		DPP9			
					DCUN1D2	E2F3	RDM1		MRPS18B	SRL		TIPARP			
					KARS	POGZ	ZNF227		PITPNM3	ZBTB7A		CWC27			
					DNAJC5	SACS	ALG10B		TGFB1I1	S100PBP		PDS5A			
					AFG3L2	TMEM97	ANAPC10		KIAA1432	PRPF4		ALDH6A1			
					C6orf211	ZNF254	ZDHHC2		HDGF	RSBN1L		TMEM38B			

					DOC2B	BTN3A2	GPR19		NFAT5	SNX29		MTRF1L			
					UGP2	ANP32B	HSS00264804 (hsa)		TFAP4	MRE11A		GALNT4			
					KPNA2	SEC14L1	KIAA1143		PPP1R12A	FBXW2		MTERFD1			
					CCDC149	UBE3A	PPP1R15B		TET3	GOLGA1		SLC25A43			
					TMEM41B	AP1S1	ATRNL1		SLC5A6	PDGFC		ABI2			
					ZFHX3	ZNF493	PTPRE		CCSER1	TGM2		NCOA3			
					HSPD1	CDK9	ZNHIT3		THRAP3	PCNP		LMTK3			
					TPX2	FOSB	FANCM		DNAH7	WDR3		LPP			
					WWC1	UQCR10	HAS3		CRYBG3	BCL7A		SLC41A1			
					NACC1	STS	RBM4B		ARNT2	SYTL2		SKP1			
					TARDBP	SLC29A1	N4BP2		GNB2	PARP1		SP2			
					SERPINH1	F2R	TNFAIP2		PCNX	TNFSF13		ANLN			
					SERPINE2	FAM134C	WDR82		SMG9	TBCE		NSD1			
					CLTC	ACTR2	FNDC3B		KHSRP	CNN3		C1orf174			
					KPNA4	SUPT16H	ICMT		EEF2	TMEM150C		LRRFIP2			
					SPCS3	MYC	SBDSP (hsa)		AR	SOD2		TMEM184C			
					MAD2L1BP	THOC7	ELOVL1		PLCG1	CPOX		PLEKHA8			
					NOTCH2	NDUFA3	ANXA7		RAB3IP	PAPD4		CETN3			
					SQSTM1	LIMK2	WDR44		POLR2A	TRAPPC6B		CCR7			
					GABARAP	PBRM1	SNX27		DNAJB1	ZFR		NCL			
					ZNF800	PCDH1	MRPL13		ARHGEF2	KIAA0040		RPS6KB2			
					GLS	HMCN1	Contig36447 (hsa)		BTG2	TNS3		PRCC			
					ID4	COX20	MSH6		CASP3	SUZ12		FAM222B			
					VCL	SLC45A1	XM_210560 (hsa)		FBN1	PLCE1		CMTR2			
					USP47	SF3B2	RAB27B		ENO1	CCDC88C		MYCBP2			
					WDFY2	CD46	WDR76		NCOR2	ZNF609		RBM38			
					GNB1	ZNF626	BICD1		ANKRD49	IGFBP3		KLF10			
					RASGEF1B	MKLN1	ZADH2		HOMEZ	EIF3M		PSMD13			
					ZNF75D	TXNRD1	BTNL9		ARHGEF40	LGALS8		MRFAP1			
					PA2G4	PPM1A	NCOA6		HNRNPLL	RAB3B		CSRNP2			
					GNL1	EARS2	MYLIP		PPP1R15B	ACSL4		SMARCB1			
					SNN	MYO9A	NA		PTPN6	NAT10		NHP2L1			
					LUC7L3	STRN4	NONO		CCT3	BAK1		G6PD			
					MINPP1	VEZT	ASPM		FBN2	GLIPR1		C15orf39			

			MAP3K4	VCAN	PNP		ATP6AP2	CSTF2		FAM104A			
			WWTR1	USP46	C2orf44		XPO6	CYB5B		DFFA			
			GM2A	INSIG2	Contig47441_RC (hsa)		FBXO17	AQP11		CERCAM			
			CXXC1	SRSF6	Contig40673_RC (hsa)		TXLNA	ZNF644		MFHAS1			
			SPAG17	TMCC3	MXI1		CATSPERB	ZNF704		TP53			
			ARFGEF2	ELMSAN1	TMEM19		TMEM173	RPN2		BAG4			
			XBP1	C12orf29	AK090904 (hsa)		MARS2	LRRC1		SPTY2D1			
			PPP2R5C	LZTFL1	Contig26388_RC (hsa)		TOP2A	BZW1		EXOC4			
			COA6	LIFR	H2AFY		CCND1	UBAP2L		TUBA1A			
			EIF3B	POLR1B	AXL		MAN2B1	RTF1		FYTTD1			
			ASPH	PTP4A2	WWC3		GRPEL2	RLF		WDR49			
			FIZ1	CHERP	ALPI		FAM19A2	FOXO3		MITF			
			RHOA	RAB32	WBP1L		ASTE1	DERL1		NTMT1			
			B3GALNT2	PRKCSH	hCT2318908 (hsa)		RBMX	SLC7A2		NUAK1			
			PCDH10	ZFP36L1	AF075045 (hsa)		ST6GAL1	STARD7		THAP10			
			TMEM199	KLF3	SEC23B		C2CD5	WDFY1		ZXDC			
			SEC61A1	FAM49B	NEK1		JUNB	GATAD2B		LIMS1			
			MGLL	EPS15	SLC19A2		AP3B1	ABHD3		MTURN			
			CACUL1	DEGS1	LRRC58		SOWAHA	RAB22A		SEC61A2			
			SF1	TULP3	C1orf198		ANXA7	USP7		MRPL18			
			RTKN2	PCBP1	Contig55770_RC (hsa)		SPEN	EIF4G3		ANXA5			
			CHD7	TBC1D14	Contig20748_RC (hsa)		PTBP3	CDKN1A		RICTOR			
			HADH	CPNE3	HEL2		RAE1	NAGPA		ASH1L			
			NGFRAP1	STX17	MXD4		CRY2	UHMK1		TOP2B			
			VPS26A	INTS7	AF075069 (hsa)		MSH6	BIRC3		PANK3			
			KDM2A	ZNF460	MT1X		GPD2	KIAA0232		METTL23			
			PGM3	TFRC	ZNF443		RNF25	SYNE2		SSRP1			
			ZNF330	ZMIZ1	RP9P		ASNS	GTPBP10		RBBP6			
			UBE2J1	BIRC6	HJURP		SNRPA	DOCK9		RERE			
			SNRNP27	NEK6	NA		NEDD8	TNRC6A		PURB			
			NAAA	TMED10	EFCAB11		CCNE2	MAP3K2		CCDC97			
			APPL1	TCF24	NFATC2		TRIM65	ARF6		USP45			

					C18orf25	DLL1	KPNA5		SLC25A13	KCNH7		AMMECR1			
					TMEM164	KPNA1	AK055807 (hsa)		TAPBP	USP9X		ZMIZ2			
					CD2AP	REL	DNAJB9		RFTN1	WDR43		IFT74			
					GALK1	SLC25A24	HERC2		E2F5	NFIA		RECK			
					PRSS23	CEP350	CYP24A1		CTHRC1	GFM2		CTTN			
					PSAP	OXCT1	CREBRF		SSFA2	OGFOD1		GABPB1			
					PDP1	SLC43A2	RSL24D1		SEC24C	UBR3		SRA1			
					GOLM1	GLO1	TRIP13		PHLDB2	ZCCHC7		MED1			
					HOXC11	VCP	CA13		SLC8B1	PTPMT1		RNF40			
					HIC2	LPIN2	BC031864 (hsa)		SBNO1	CHAMP1		ITSN2			
					QSER1	PINLYP	DLC1		UBXN8	MRPL42		RPAP3			
					HNRNPU	TTC3	ACTR2		CDV3	EIF4EBP2		TRIP12			
					MRPS14	KCTD3	MLF1		PNP	LSM12		MRPL55			
					CYLD	MOGS	IRS4		CDC45	SMG7		SIPA1L1			
					HSDL1	ZMYM2	FAM120AOS		NPEPPS	VGLL4		MIS18BP1			
					PNPLA8	APLP2	MYC		SKP2	CKAP2		AZIN1			
					HSPH1	WASF2	CCDC18		PBX3	SESN3		RIMS4			
					RCAN2	PRNP	TEX30		SH2D3C	OTULIN		RALGPS2			
					VLDLR	DNAJC16	Contig43135_RC (hsa)		FAR1	TMEM25		DYRK2			
					SLC7A5	FBXO10	UQCRC2		COL18A1	HNRNPA2B1		TEX2			
					OLR1	SLC38A2	RFC4		AXL	HRSP12		DCAF7			
					GXYLT1	BAZ1A	PCDH1		POGK	GTPBP4		THAP1			
					WDR6	HIC1	CENPF		UFL1	ATXN3		CREB1			
					WIZ	RNF19A	EPM2A		TSR1	PARD3		TBC1D9B			
					SNX15	ZNF100	RAB6C		CLTA	FIP1L1		TDRD3			
					ICAM2	NIPA2	CSTF1		SLC19A2	HMGCS1		MORN2			
					EZH2	TMEM87B	FBXO3		IGFBP4	PTPN12		FAM122B			
					CDK6	C17orf58	FAM83D		TMEM243	DDX1		HECTD1			
					EPDR1	KDM3B	AL137347 (hsa)		DPYSL5	SHANK2		AGPAT2			
					HEG1	ZNF737	MANF		COX14	CCNI		PEX14			
					CNTNAP2	NOL4L	PASK		SLC4A7	BTBD7		ASAP1			
					MAP3K9	HNRNPM	PPM1A		TRAK1	OPA1		GOSR2			
					RHEB	RBM26	PARP16		CISH	SEC11C		ZDHHC16			
					DCAF10	SMAD5	COPS3		LRIG3	SLC25A27		RNF10			

					IPO7	ZNF138	AK098664 (hsa)		CNNM3	FAM63B		FAM217B			
					GATA2	ELOVL5	DEPDC1		HIF1A	TMEM189		ELF4			
					DDX3X	CREBZF	VCAN		CLN3	MAP1A		RAB5A			
					MDH2	STAM2	RPGRIP1L		E2F3	SP110		PKDCC			
					SETD2	SLC39A9	TUFM		POGZ	RASSF8		MT1E			
					RER1	DENR	TP53INP1		TBL2	RIMS3		AGPAT1			
					TGFB1	CDC42BPA	LRP8		ITM2C	PRKDC		ATP6V1B2			
					FSTL1	NCOA2	hCT1835226.2 (hsa)		GDAP2	CNBP		KRIT1			
					CNNM2	ZNF263	AL137333 (hsa)		CEP41	DGKD		EIF5			
					MARK3	ITPR1	ST3GAL6		ACER2	IFT80		FBXO34			
					INIP	EI24	C21orf91		MAPK8	RAN		MRPS7			
					RPIA	CFL2	PRKCSH		KPNA5	C1GALT1		PTPN13			
					MVB12B	THAP11	CLIP4		SLC35D2	PPIF		C7orf26			
					TCEB1	SEC24A	KLF3		NOP2	TRRAP		PRICKLE1			
					MCM7	ATIC	MNS1		CSNK1A1	ZFP36L2		FOXP1			
					SLC16A9	PIK3R1	FAM49B		CREBRF	OSBPL8		SOGA1			
					PHF10	TCF20	DEGS1		CCR5	ZNF106		ZSWIM1			
					MID1	RB1	FAM126A		TRIP13	GPR157		CAPZB			
					SLC35A5	PNRC2	DYNLT1		TRMT1L	AGPAT3		DKC1			
					NOL9	BCAP29	UBAP1		MDGA1	CRCP		TAGLN2			
					GSTO1	BAZ2A	Contig23423_RC (hsa)		ZFHX2	GATAD1		TRAFD1			
					CHD4	SLC35E1	ZNF460		CHCHD10	RNF6		MAVS			
					ARHGDIB	SATB1	hCT2297806 (hsa)		BLNK	FUT8		TICRR			
					MBD1	SMARCC1	PCNXL2		WDR48	SNX25		GNA11			
					FZD3	STRN3	DSN1		STAU1	AGTPBP1		EIF2AK2			
					CPD	TGIF1	AES		DLC1	SGOL2		BLVRA			
					FNBP1	CSNK2A1	HSS00300718 (hsa)		KCTD10	SLC35B4		CD48			
					ZBTB41	MARCKSL1	BMP2K		PPP1R15A	DOCK7		PEG10			
					BUB3	KDM5A	SPAG9		MYC	LPHN1		MLEC			
					MSN	NXN	TMED10		MFSD10	HEY1		SRRD			
					SLC35C2	USO1	STIL		TMC5	ACIN1		POLR2D			
					TPM3	DOCK4	FOXA1		DEDD	UBE2B		GDA			
					RPS20	PTPN11	NM_173466 (hsa)		IFI44L	MAP2K4		BRCA1			

			WDR26	POM121C	ZNF525		PBRM1	DNAJA2		ZNF217			
			TRAM2	BUD31	CCDC47		PARVA	TANC2		MLX			
			ZNF107	HSP90B1	PIM1		SON	PSIP1		LARP4B			
			EPC1	ARPC1A	IL17RB		HSPB7	EPHA4		DCAF6			
			API5	FARP2	ARNTL2		HCFC1	UNC13B		NUP214			
			SREBF1	YLPM1	ZC3H12C		TRIM3	SYT15		CROT			
			NDUFC2	YKT6	CBX6		SF3B2	MARCH6		PSD3			
			ADD3	DYNC2LI1	ZMYM1		RGL2	SPTBN1		RASA1			
			TANK	SLC25A36	HCN3		FBXO3	CHD6		LARP1			
			PRPF8	ZNF708	ECT2		PLK2	TMEM55A		SH3GL2			
			SLCO3A1	GDI1	AK022645 (hsa)		PIGP	HOXB13		KLHL42			
			CELF1	ZC3HAV1	Contig32087_RC (hsa)		VEZT	NFIC		PTCD1			
			PDS5B	DHTKD1	hCT2336241 (hsa)		TBX19	CAB39		FH			
			DBF4B	TMPO	RBP7		DEPDC1	LMAN2		MED24			
			EPAS1	SRRM2	FAM20B		TP53INP1	PSMF1		PIAS3			
			MED12L	PTPRG	HOXB9		ELMSAN1	TRIB1		NCAM2			
			LZTR1	TEX261	TBC1D31		RPTOR	NSFL1C		KIF3B			
			LBR	TOMM22	PRNP		MRPS2	AZI2		RNASEH1			
			CKAP5	KLHL15	LGMN		GPRC5A	NIPBL		DMWD			
			TAOK1	TFPI2	AL078636 (hsa)		DIAPH1	NUPL1		PRCP			
			SIK2	NUAK2	PARN		IGF2	CCDC59		MAL2			
			ALAS1	RCHY1	NHSL1		CCT8	NUS1		DDX24			
			TFPI	ZNF506	ZNF215		TK1	SRF		NUP35			
			ERGIC2	MID1IP1	TCAM1 (hsa)		SPRYD3	RRAGC		PTCHD1			
			ALG9	LARP4	FAM80A (hsa)		ORMDL2	ENOPH1		TRAK2			
			DST	PTPRJ	TRAF3IP3		ZFP36L1	PRDX3		CRTC2			
			CAPZA1	ZBTB43	hCT2255806 (hsa)		C4orf3	GREB1		KANSL1			
			ABHD2	TIPRL	C17orf58		HPN	TET2		TSKU			
			HES1	PIP5K1A	TIMP3		CDCA7L	ZNF33A		CERS2			
			PLAGL1	C6orf62	AL137535 (hsa)		ARID1A	VPS35		ANP32A			
			BMP8B	ANKRD13C	CLK1		PANK2	NPHP1		URB1			
			CANX	OTUD4	TTK		FAM49B	KMT2C		MICU1			
			CDR2	DPP7	C5orf51		MTRF1	HELLS		HS3ST3A1			

					F3	ZNF429	NA		DEGS1	SERTAD2		SMC4			
					SMG8	PRKAA1	ELOVL5		DENND5B	SLFN5		RPS26			
					ZNF746	MAK16	HIATL1		RASGRP4	FKBP1B		SCAMP5			
					SIKE1	ANKFY1	PIP4K2B		PCBP1	RAB10		TRMT5			
					ADIPOR2	PEBP1	CREBZF		BEND4	PCM1		PPP1R12A			
					KIAA1109	MESDC2	SLC39A9		RRAGA	PARD6B		TET3			
					ACVR1	DNMT3A	SFPQ		SOCS5	TBC1D5		SLC5A6			
					IER3	KMT2D	REEP4		E2F6	RRAGB		NDUFAF4			
					CAMKK2	ADAMTS1	hCT2283962 (hsa)		SKIL	ATP10D		MLLT11			
					ZBTB20	DNAJB14	EI24		BIRC6	SAMD11		SLC5A3			
					ABI2	FSCN1	CFL2		IKBKAP	TAF3		ATP5G3			
					NCOA3	PAK6	ID1		TMEM38A	ZNRF3		AGO1			
					HDAC9	UBE2K	NIPAL1		ADAM12	CCDC158		NECAP1			
					NFX1	NUP50	PANX1		AES	ENAH		EPT1			
					WDR5	IPO5	TBKBP1		XKR7	TNPO3		SND1			
					LIMD1	SIAH1	LRRC8C		TGM1	MRPL52		RPP25L			
					LPP	STX12	ANKRD27		KLHL36	DNAJC2		GGA2			
					IFNAR2	SESN2	NT5M		SLC25A39	C14orf159		ATG101			
					TMEM123	HSPA4L	Contig16329_RC (hsa)		CCDC47	FAM79A (hsa)		VGLL3			
					LRRC42	TRIO	STARD3NL		PIM1	NA		COPRS			
					ATP2B1	CITED2	TMEM169		SLC25A24	CBX5		PAN2			
					SKP1	TNFRSF10B	RCBTB2		FCHO2	KCNJ3		PCNX			
					YWHAQ	TSNAX	TFCP2L1		PDGFB	SETX		SMG9			
					LSM5	ZNF595	MTSS1		APOBEC3C	JMJD6		STAMBPL1			
					SUPT3H	COMMD2	RORA		COL1A1	SPOP		RFXAP			
					RCN2	HILPDA	MRPL43		PRKACA	BAZ2B		HNRNPAB			
					PCDH17	PIK3C2A	TMEM246		CHD3	GATA6		EEF2			
					PLEKHA8	YEATS2	HSS00250837 (hsa)		XPR1	TRMT10A		POLR2A			
					PRCC	C16orf72	C2orf49		MOGS	SOX12		DNAJB1			
					TMEM11	ZNF518B	NPNT		GPR56	EIF1		BTG2			
					FAM222B	ST7	PHC2		CNNM4	RAB14		POLR2M			
					SAFB	ELAVL1	AKAP7		SLC11A2	USP10		RPAP2			
					CMTR2	BTAF1	AL110163 (hsa)		WASF2	PTGES3		ZNF395			

					SNX14	TMEM2	GLCCI1		EGLN3	KLK10		TSC2			
					MYCBP2	ADCK3	ACVR2A		SMNDC1	JAK1		GPD1L			
					ALCAM	DNAJC12	KIAA0284		ZBTB5	EXO1		PNN			
					KLF10	DHX15	RNPC3		MYO5A	STX2		ZNF711			
					MPP5	CLASP1	ZNF708		LRP1	AKAP11		RBMS2			
					TMEM181	KCNK1	L2HGDH		COQ10B	MDM2		ABHD17B			
					NHP2L1	ZC3H11A	ZSWIM3		BAZ1A	GALM		FBXO28			
					INSIG1	ADNP	AF078844 (hsa)		HIC1	MTDH		DCP2			
					ATXN10	NF1	NM_015595 (hsa)		RNF19A	NIN		XPO6			
					ARFIP2	PSEN1	ZNF184		DDB1	PRR11		DHPS			
					FAM104A	FIGN	BC039371 (hsa)		EIF4G1	DNAJC8		BABAM1			
					RPS6KA5	ZNF700	TMPO		RAB11A	ABL1		NCAPD2			
					RALGAPB	ZNF430	ACTR1B		RBBP5	AP5M1		TOP2A			
					SOCS6	CDK13	SRRM2		SVIP	EHMT1		CCND1			
					CBFB	DSG2	CDK5RAP2		ZC3H13	HAT1		RBMX			
					MFHAS1	SUV420H1	AK057062 (hsa)		SHISA5	COIL		WDR82			
					SIDT2	ZFP1	KLHL15		HIST1H2BG	CDC123		TRDMT1			
					MAML3	SLC16A6	PRIM1		XYLT2	BMPR2		SMAD4			
					CUL4A	SP1	EOGT		KDM3B	KIAA1551		USP32			
					TP53	DAB2IP	Contig30392_RC (hsa)		JMJD1C	DMXL2		SLC35D1			
					BAG4	LURAP1L	CDK14		NOL4L	PNISR		FNDC3B			
					TNRC6B	NCSTN	SENP1		HNRNPM	BCL2L11		SEPHS1			
					SPTY2D1	BNC2	SLC26A2		RNMT	MED8		MEAF6			
					ALDH5A1	CD93	hCT1640064.3 (hsa)		SNX16	NR2F2		GIT1			
					CPPED1	AGK	C12orf65		TLE4	KIF5B		CTNNB1			
					MOCS3	WDR35	C1orf53		PDE12	MTF1		ZNF805			
					MITF	SSR3	METTL9		CLK1	LMO4		C9orf89			
					ADAMTSL3	OSTM1	DIS3L		C5orf51	MYO1B		EFR3A			
					LITAF	ACVR1C	KBTBD11		C19orf47	LIN7C		ACTR3			
					GLTSCR1L	ATP6V0A1	PRPF38A		FER	SIN3A		PPP3CB			
					DDX21	SMIM13	RNASE4		PSMC4	MLLT10		ANXA7			
					PRRC2C	EGLN2	SPRY4		ELOVL5	PHLDA1		THOC3			
					SEMA3F	PRKCB	C6orf62		TGOLN2	HOXA13		SNX27			
					RPL23	CDC73	KLHL23		MMD	MAP1B		THOC6			
					ANXA5	AKIRIN1	RIPK5 (hsa)		SLC39A9	EPB41		PTBP3			

			C9orf69	CDC23	ZNF823		ARIH2	APH1B		VTI1B			
			RICTOR	MAT2A	IFRD1		RNF20	TMF1		RAE1			
			DCTN4	DAPP1	ICAM3		ASNA1	CTGF		TUBA1C			
			ADAMTS2	ZNF451	Contig33235_RC (hsa)		SF3A2	IFI6		CTNNA1			
			TOP2B	RASSF3	CORO2B		GAD2	DHX33		WDR76			
			ZSWIM6	TSPYL4	ATF1		MSANTD2	UBLCP1		MAPRE2			
			EGR3	TNKS1BP1	CTPS2		TBC1D25	GKAP1		KDM4A			
			PANK3	ARHGAP29	PTP4A1		RPRD1B	ZKSCAN1		PLEKHB2			
			SSRP1	NFE2L2	ABR		CFL2	FKBP10		NR3C1			
			PTER	ZNF732	NOLC1		ACADVL	BSN		TAPBP			
			IMPA2	STX6	FGF2		DHX34	SYNCRIP		LYRM2			
			RBBP6	DOT1L	Contig35629_RC (hsa)		HOOK1	YTHDC1		PIGC			
			SRSF1	SLC20A1	Contig42328_RC (hsa)		KLHDC3	ZBTB8A		SAP30BP			
			SH3PXD2B	PPP6R1	FHDC1		GINS4	ESPNL		BAG1			
			PURB	LAMC1	FSCN1		NIPAL1	KDM5B		MYLIP			
			FADD	RAI14	PAK6		ZNF362	TXNIP		FEZ2			
			AMMECR1	NFE2L1	AK024927 (hsa)		ZNF426	GGNBP2		MPHOSPH6			
			EVL	FOXJ3	Contig39852_RC (hsa)		TCF20	PCDHB9		QARS			
			HIF1AN	MKI67	RNF219		PXK	RP5-850E9.3		HYAL3			
			SLC35F5	PAXBP1	STK3		AEN	C5orf22		SKP2			
			PNKD	TGFBR3	ENST00000296018 (hsa)		ERP29	NCOA4		ZNF264			
			GABPB1	UBE2V2	MMAB		SMARCC1	ANKRD34A		FAR1			
			EEF1D	APEX2	SERTAD4		NADK	GTF2I		H2AFY			
			DPF2	PRKAR1A	FBXO36		ANKRD27	SMEK2		SLC35C1			
			RNF40	ASCC2	ENST00000278622 (hsa)		CSNK2A1	TMEM189-UBE2V1		NUFIP2			
			TMEM170B	FAM160B1	HES6		MARCKSL1	PNLIPRP2		RAPGEF2			
			ZNF207	SQRDL	TTPA		KCTD17	ZFP91		PRMT6			
			ITSN2	MIDN	CDC7		WDTC1	CTDSP2		COPS5			
			UROD	ZNF714	TOR1B		ARL4D	WBP1		UCP2			
			RFC1	HDLBP	BMP2		PRPF38B			DUSP18			

				HNRNPK	DDAH1	KCNC4		CD97			RNPS1		
				COX19	GHITM	ZBTB22		CNKSR3			MSL1		
				TRIP12	TMEM245	PAWR		MRPL43			ADPGK		
				DHRS7	ARFIP1	SKA1		POM121C			MRPL30		
				GRN	XXbac-BPG32J3.20	RAD1		BUD31			TRAK1		
				SIPA1L1	FBXO7	FAM199X		CDC42SE2			C1orf198		
				AKAP8L	TP53BP1	TMEM2		FARP2			PAICS		
				AZIN1	CDK17	DHX15		B4GALT4			ABCE1		
				GLTP	S100PBP	KDELC2		COX4I1			HIF1A		
				VAV2	PUM2	CDC20		C11orf84			HELZ		
				DYRK2	ATXN7L3B	BUB1B		YIF1A			GIT2		
				TEX2	NFIB	DYNLL2		YLPM1			NDUFS1		
				RELT	AHR	BTC		GLCCI1			APITD1		
				G6PC3	SELPLG	ADNP		ACVR2A			MT1X		
				DCAF7	MDK	MAD2L1		NRP1			SACS		
				ZNF253	CHML	C6orf120		PRKAR2A			CTNNBL1		
				TMEM66	ATF7IP	FAIM		NAA30			TMEM97		
				CREB1	WDR3	LYRM7		MBD6			ZNF254		
				TSG101	DDX47	CCNY		TST			SLC35F2		
				POGLUT1	CNOT1	TOPBP1		PHLPP1			SEC14L1		
				PPP4R2	ZNF281	ERCC4		CTSL			MAPK8		
				CCSAP	ATG3	FZD1		WDR81			ICE2		
				HEATR3	ZZEF1	KIAA0100		STIM1			AKT1		
				DAP3	PHF12	hCT2285841 (hsa)		SEMA4C			TBC1D7		
				CCNT1	TOPORS	DHFRL1		ZC3HAV1			TLK2		
				ARL4C	KCNK5	MGRN1		FIGNL1			RQCD1		
				ANKRD40	TPRA1	CASC5		RAP2C			KTN1		
				ATP6V0B	ILDR1	HINT3		TNK2			PKMYT1		
				EZH1	CYB561	HIBADH		SRRM2			CGGBP1		
				FAM122B	PLK1S1	ZFP1		RNFT1			PPP2R2A		
				PRR14L	CKAP4	YY1		DLGAP5			RNASEK-C17orf49		
				ARL8B	POLQ	CCL3L1		GEMIN5			RNF24		
				HECTD1	H3F3B	SLC16A6		RPAP1			MKRN1		
				KLF13	SEC24B	HIVEP1		KLHL15			RAVER1		
				ZNF608	TMEM150C	COL1A2		EOGT			DESI2		

					ASAP1	UBE2D1	CCDC15		PPIL4			CEBPA			
					GALNT6	RNF167	KPNA3		NUAK2			KCTD10			
					FAM217B	CTAGE5	MDM4		TMEM259			ACTR2			
					LAPTM4A	DDX39B	LARS		ICOSLG			SUPT16H			
					DHX32	SOD2	OGFRL1		METTL9			FOXN3			
					ACVR2B	CPOX	GSK3A		DNM1L			IRS4			
					EFHD2	CEACAM6	NM_006333 (hsa)		ASCC3			PPP6R3			
					SDPR	KIAA0040	ATP6V0A1		SYT7			QKI			
					OXA1L	PYGB	BTF3L4		LRFN4			HOXA5			
					MTR	CENPE	FUT11		CUX1			SMU1			
					MLLT4	OAZ2	NA		RIMKLB			PBRM1			
					DPY19L4	GNL2	HSS00368570 (hsa)		PTPRJ			UBE3C			
					SLC7A1	RC3H1	UCHL3		GPR137			RFC4			
					ICE1	ZNF609	PTPLAD1		GLCE			CENPF			
					FOXP4	SNX13	ZNF451		RAC1			IKZF5			
					EIF5	SMIM3	Contig36261_RC (hsa)		INSR			HCFC1			
					SERP1	FASTKD5	ARHGAP29		MAPRE1			SF3B2			
					PTK2	PLXNA4	PGRMC1		OTUD4			CD46			
					SLC2A3	PJA2	PSMC3IP		DPP7			AHCTF1			
					C9orf41	ZBTB33	Contig48494_RC (hsa)		XYLB			RAPGEF6			
					VPS4B	CPVL	SFR1		PRMT1			FBXO3			
					MRPS7	ZNF257	KHDRBS3		ANKFY1			AHDC1			
					CNN2	HOMER1	MKI67		CTPS2			TMEM14C			
					NLK	UBE2M	HDAC11		ABR			MKLN1			
					SOGA1	KLHL9	RRP1B		NOLC1			SEC14L2			
					MAVS	MARCH5	LMAN1		ATP6V1A			SMIM15			
					TIMM10	TFIP11	HSS00142930 (hsa)		PMPCB			TXNRD1			
					TMEM242	NGRN	UBE2V2		KMT2D			MRPS9			
					TES	RCBTB1	HSS00090480 (hsa)		KDM6A			PLK2			
					TICRR	GUF1	NA		UBE2K			DCTN3			
					ZBTB37	AURKA	PRKAR1A		KDM6B			PPM1A			
					LYPD3	IL6ST	RHBDL1		RARA			EARS2			
					SETDB1	ACTR1A	AK057835 (hsa)		DOLK			PARP16			
					ETNK1	NUP43	EP300		ANKRD46			MYO9A			

			PEG10	FUBP1	hCT2306120 (hsa)		UBE2D4			PIGP			
			PPP1R16B	VEZF1	HDLBP		SESN2			PRELID1			
			MOB3B	UPF2	AHSA2		SDE2			RPS24			
			PISD	MYO5C	ZNF35		EBAG9			RAD23A			
			LCORL	PHF16	PPARG		TRIO			USP21			
			AKT1S1	UBAP2L	KCNMB4		KCNC4			HEATR5B			
			CBLB	RTF1	ENST00000295192 (hsa)		NELFE			TP53INP1			
			TMEM254	AKAP10	ARL5B		TNFRSF10B			LRP8			
			GDA	ZNF682	Y16700 (hsa)		KDELR1			ELMSAN1			
			GPR125	GPI	ICK		GMPPB			MRPS15			
			EML5	LASP1	hCT2328745 (hsa)		RDH10			LZTFL1			
			ZNF217	MECOM	NCAPH		C16orf72			OCIAD1			
			LARP4B	FOXO3	NKX3-1		RPL19			SUB1			
			JOSD1	DHX9	PUM2		NBPF3			DIAPH1			
			FAF2	ABLIM1	NFIB		OBSL1			GLI3			
			NXF1	IFNG	AK024121 (hsa)		IGF2BP3			RNASEK			
			NUP214	FGD4	RSBN1L		TXLNG			MECP2			
			NAA20	STARD7	CHML		ELAVL1			ZFP36L1			
			PSD3	UBN2	OSBPL10		PHKA1			ST6GALNAC6			
			LIF	C22orf39	hCT32256.3 (hsa)		TMEM2			WDR19			
			LARP1	ANXA11	PITPNB		ZBTB21			SMAP1			
			CEP192	PXDN	CLN6		PSMC3			GOLGA4			
			KLHL42	TMEM218	FBXW2		EMP2			GPBP1L1			
			PRDX4	RAD54L2	ATF7IP		PSMD12			CDCA7L			
			TAF1	MATR3	C18orf54		EMC4			KLF3			
			HNRNPUL2	GAS2L1	CLVS1		AP2A1			UBA1			
			ABCF2	CRIPT	CDK12		HSPA8			EPS15			
			PIAS3	AKAP12	EHBP1L1		NF1			CLDND1			
			NCAM2	UFD1L	USP14		FIGN			STK35			
			KIAA1147	EIF4G3	hCT2250864 (hsa)		CCNY			KIFC1			
			KIF3B	RBM12B	WHSC1		TOPBP1			ZFAND1			
			ZNF697	ERH	Contig57226_RC (hsa)		FZD1			PCBP1			
			ARAP3	ZNF92	CREB3L2		HMG20A			YARS			

			TRIM59	UHMK1	EPM2AIP1		CDK13		TBC1D14			
			PGM2L1	POU3F3	CKAP4		KIAA0100		KXD1			
			WDHD1	NAPEPLD	POLQ		MED28		FBXO8			
			MAL2	CLSPN	CNTNAP3		UBAP2		MTPN			
			DDX24	DOCK9	SAMD3		GPX8		INTS7			
			FUBP3	TNRC6A	TNFSF13		RANBP2		RRAGA			
			EBF3	TMEM255B	BIVM		MGRN1		E2F6			
			C4orf29	MAP3K2	CLIP2		SUV420H1		TFRC			
			CRTC2	UTP18	CERS6		ULK3		ZMIZ1			
			DUS1L	USP9X	ENST00000291738 (hsa)		ZNF345		VIM			
			KANSL1	MBNL1	UBE2D1		YY1		DSN1			
			DDX18	ZFP69	PCAF (hsa)		RARS		SPAG9			
			PMAIP1	ZNF66	TARBP2		ARID5B		TMED10			
			CERS2	PDXDC1	PPP2R3B		DUSP7		WDR73			
			SLC9A1	NDUFA2	SMC2		CLDN1		SETD8			
			SP4	GMCL1	PCDH7		KLHDC8B		KPNA1			
			ELAVL4	INPP5A	NM_016125 (hsa)		COL1A2		COG8			
			MRPS18B	ZMAT2	PLCE1		CLP1		AAR2			
			PDHA1	AVL9	CCNE1		RBPJ		SLC25A24			
			FZD4	KIF21A	CENPE		MFSD5		CEP350			
			KIAA1432	CHAMP1	FKBP15		MDM4		POMGNT1			
			BRD7	REEP1	NFKBIL1		ALX1		ATRX			
			SFXN1	PPA1	SNX13		CSNK1G1		FKBP5			
			CUL5	EIF4EBP2	COCH		ADRM1		ZNF773			
			TRMT5	CSTF2T	SLC16A7		PLXND1		ZNF614			
			NFAT5	SLC39A10	SLC1A4		C19orf10		SET			
			MFSD9	SESN3	RACGAP1		GTPBP2		LYSMD2			
			PPP1R12A	PTX3	TMEM55B		MGAT1		ZMYM2			
			LIPT2	FAT1	CHDH		HIST1H2BL		HMOX1			
			SLC35F3	HRSP12	ZBTB33		MAT2A		GIPC1			
			MBTD1	DCAF12	AB040883 (hsa)		ZNF451		CNNM4			
			TET3	PDPK1	GPSM2		MAX		HOXB9			
			DROSHA	EIF1AD	FBXO31		TLE3		DENND4C			
			SLC5A6	NOP9	ANAPC15		NRARP		WASF2			
			NAV1	ZNF665	RAB3B		TNKS1BP1		BEND3			

			CCSER1	ZNF483	SCD		UBE2O			ARHGDIA			
			KLHDC10	DKK1	UBE2D2		ARRDC3			SLC38A2			
			ZNF655	MEMO1	FAXC		USP42			BAZ1A			
			DDX54	GORASP2	EAF2		PGRMC1			EIF4B			
			CSNK1D	HMGCS1	PLEKHG2		DOT1L			PARN			
			SLC5A3	BNIP3L	NUF2		SLC20A1			AHCYL1			
			ZBED6	NCBP1	NGRN		LAMC1			QPRT			
			LFNG	LIMA1	PAPD7		USP24			DDB1			
			ATP5G3	SHANK2	MIS12		RAB40C			EIF4G1			
			SF3B3	PNMA1	CEP162		NEDD4			RBBP5			
			RAB30	ZNF615	RLIM		ERC1			SVIP			
			KDELC1	CCNI	IL6ST		TSPAN14			ZC3H13			
			KCTD21	ZNF606	FUBP1		KLC2			SHISA5			
			AGO1	CUL3	ENPP4		TGFBR3			DENND6A			
			NECAP1	BTBD7	ZNF704		LAMP2			NSA2			
			S1PR3	OPA1	FAM222A		ASCC2			KDM3B			
			TBC1D4	TACO1	C9orf163		CYP20A1			JMJD1C			
			MICU3	FAM63B	LRRFIP1		FAM160 B1			GBAS			
			AP1S2	RUFY2	EPS8		PROSER1			C5orf51			
			DLX5	RPL23A	TDP1		TTLL11			NCAPG			
			OGDH	PCBP2	LRRC1		NR2C1			TRIM33			
			KLRG2	HIPK1	NSF		CSNK1E			TGOLN2			
			ATG101	AHNAK	UBAP2L		SLC4A5			THEG			
			YIF1B	TRPM7	MPP2		CYB561A 3			PTGS2			
			NAV2	LUZP1	ZNF573		MIDN			SLC39A9			
			SLIT2	NRIP1	ALDH9A1		PRDM2			GTF3A			
			KLKB1	ZBTB34	GAB2		DDAH1			CYB5R4			
			ROBO4	AP1G1	AL133090 (hsa)		SLC10A7			WDR45B			
			PCNX	PAFAH1B 2	TCEA2		CBLL1			SFPQ			
			TTC37	CIZ1	SLC7A2		ARL5B			DAG1			
			PKP4	VMP1	ACSL1		ZCCHC14			RMDN3			
			C1orf52	NDUFV1	37500		PIGT			NCOA2			
			STAMBPL1	ZIC5	BASP1		SAMD10			HIST1H4H			
			KHSRP	PDE7A	CTCF		FBXO7			GAD2			
			CLIP1	DUSP5	PXDN		TP53BP1			SAP130			
			EEF2	PPIF	DIMT1		PPP2CB			PABPC1			
			EXOSC1	ZDHHC17	TXN2		TTLL4			TXN			
			RAB3IP	OSBPL8	Contig60864_R C (hsa)		RPUSD3			EI24			
			BTG2	ZNF106	AKAP12		PLAG1			MEPCE			

					PDLIM1	KIAA1467	GNRH1		S100PBP			APBA2			
					ARFGAP2	NCKAP1	TMEM30A		PUM2			SEC62			
					FBN1	KHNYN	LHFP		AHR			APIP			
					TMTC4	ABCC5	ATP5L		REV3L			PIK3R1			
					TOMM5	NKRF	ZBTB14		RSBN1L			PCMTD2			
					GPD1L	ULK1	BC004287 (hsa)		CELSR3			RB1			
					PAQR7	ZCCHC11	CLSPN		ATF7IP			PNRC2			
					ZNF711	FOSL1	MRPL50		BRD3			SCYL2			
					MFN1	LEPROT	TMEM201		CDK12			BCAP29			
					PIK3R3	HEBP2	GABPB2		POMK			CHIC1			
					ZDHHC2	CLDN12	COBLL1		CNOT1			ANKRD27			
					CRTAP	PTAR1	ANKHD1-EIF4EBP3		ZNF281			CSNK2A1			
					PPP1R15B	LAMP1	CNEP1R1		EIF4E2			FAM118A			
					WDR33	EXOC8	MCAM		COX10			CDKN2AIP			
					ABHD17B	NDRG4	GSG2		TOPORS			SOS1			
					CCT3	MOB1A	MYPN		MPL			KDM5A			
					FBXO28	ACAT1	PRKD3		FAM120B			RBBP8			
					DCP2	KIRREL	ZBTB18		TSPYL1			NAMPT			
					FBN2	ITGA6	TOP1P2 (hsa)		CKAP4			MTSS1			
					XPO6	TADA1	REEP1		ATP8B3			NXN			
					RNF4	G2E3	POLL		AKAP5			USO1			
					TXLNA	SUCO	MBNL3		VPS41			CPEB2			
					CATSPERB	EHD1	LDLR		FOXP2			CDC42SE2			
					TMEM173	UQCRQ	FOCAD		ZFYVE20			C2orf49			
					MARS2	ZNF518A	ASRGL1		SEC24B			H2AFY2			
					HMGXB4	XRCC5	NM_032706 (hsa)		BRD1			GPS2			
					FAM208B	RYBP	HEATR5A		CAPN1			COX4I1			
					TOP2A	IRX2	OSBPL6		EML2			ZMPSTE24			
					CCND1	SYMPK	Contig10441_RC (hsa)		RGS16			YWHAZ			
					IMPAD1	TXNL4A	ZBTB4		TARBP2			NXT2			
					PBXIP1	DNAJA2	TRIM2		PAPD4			ARL5A			
					ILF3	NME1	HMGCS1		IDE			PSMB2			
					TSSC1	GTF2A1	CCRN4L		FOXC2			ADRA2C			
					LHFPL2	XPO1	BARD1		ZFR			SEMA4C			
					WDR82	C9orf156	AASDHPPT		VPS33A			RGS2			
					EPHB4	UNC13B	Contig49700_RC (hsa)		SUZ12			ZNF280D			
					ERGIC3	TMEM170A	BCOR		FAM216A			ATP5H			

					FNDC3B	GPBP1	GAREML		SMIM3			LDOC1L			
					JUNB	PHIP	NCBP1		BSG			TMPO			
					MEAF6	PTEN	MTAP		CDC27			TEX261			
					CTNNB1	PCF11	OPA1		KPNA6			UBALD1			
					CNOT7	ZNF652	FAM63B		ESYT2			RPAP1			
					ZNF678	NFATC3	HSS00060369 (hsa)		GNB2L1			PDSS2			
					POLDIP3	MTF2	AFAP1L1		ELOVL7			SLC30A9			
					PUM1	TUBB4B	RASSF8		VWF			TSPAN4			
					ELOVL1	DUSP16	MPPED2		BAHD1			GNL3			
					SPEN	MARCH6	TRPM7		TLCD1			PEX11B			
					MAPKAPK 5	HPS5	NRIP1		PRUNE			EEF1A1			
					THOC3	PALLD	ZBTB34		GPSM2			JAZF1			
					SNX27	SIX4	SH2B1		PHLDA3			GNPNAT1			
					PTBP3	SGK3	hCT2318966 (hsa)		ABI1			H3F3A			
					CBX1	SPTBN1	KCTD15		C5orf24			METTL9			
					PPT1	POLR1A	CEP55		DDA1			CERS5			
					CRY2	ATL3	RAB8B		RPL22L1			LARP4			
					NEBL	LIN54	UBE2W		SLC6A9			POP4			
					GPD2	TRIM25	NA		ATP6V1C 1			CUX1			
					CTNNA1	LATS1	MKNK2		SCD			SEC31A			
					IRF2BPL	ACADM	PSRC1		SEH1L			BACE2			
					PMM2	IRAK1	ZIC5		PDIK1L			PSMG1			
					GFPT1	ADAMTS 9	FAM132B		ZNF644			TNFRSF21			
					INPP4A	EIF3E	JADE1		ZNF512B			ZC3H14			
					NR3C1	ZNF98	ARPC4		THOC2			TIPRL			
					TEX22	ATP13A3	OSBPL8		ACTR1A			LDHA			
					RFTN1	SPATS2	DLAT		PLEKHA1			IRS1			
					E2F5	TRIB1	UPP1		SGOL1			GLCE			
					SHROOM4	ZNF550	NEDD1		ZNF704			RGCC			
					SSFA2	MAPK1	NASP		RPN2			RAC1			
					ST3GAL2	NUP188	NKRF		SCAF1			SNX2			
					MB21D2	PNRC1	GATAD1		AKR1A1			PIP5K1A			
					NRD1	SDCBP	DYRK1A		VEZF1			GABPA			
					MMP1	POMGNT 2	CMTM6		UPF2			FAM120C			
					POLR2H	NUPL1	RNF6		BAMBI			C6orf62			
					GIGYF2	GPR126	MOAP1		TDP1			OTUD4			
					KCTD2	SPTSSB	SGOL2		NSF			STC2			
					MYLIP	P2RY10	ATG10		TAF5L			XYLB			

			RPL17-C18orf32	C5orf15	UHRF1BP1		SCAF8		CHCHD4		
			TRAPPC9	ITPR2	SLC35B4		PXT1		PRKAA1		
			CTTNBP2	DYRK1B	ZNF449		AKAP10		IFRD1		
			BHLHE40	RRAGC	THBD		TMEM8A		ANKFY1		
			ARL6IP1	RNASEH2B	PRDM4		UBR7		FANCE		
			EPHA3	PRDX3	AK055620 (hsa)		CHD5		ATF1		
			ASPM	GREB1	JAGN1		MEX3D		PTP4A1		
			CDV3	FGF7	SGPP1		LASP1		RPL26		
			DOCK1	KDR	hCT2306937 (hsa)		FUT10		FGFR1OP2		
			NPEPPS	RAD21	ATP1A1		SEC22C		NOLC1		
			C2orf44	TET2	SCRN3		IGSF3		FGF2		
			FAR1	VPS35	Contig48185_RC (hsa)		XPO4		SHOC2		
			AQP1	AP3M1	XM_070846 (hsa)		VPS37C		CDC42		
			SAMD4A	GULP1	BX487921 (hsa)		DHX9		SEPN1		
			NUFIP2	KMT2C	CLSTN1		DERL1		BBS9		
			GCC2	GGA1	RYBP		FBXO22		DTX3		
			NEK4	GRB7	SPIB		NDC1		UBE2K		
			WBP1L	GTSE1	RSBN1		TULP4		NUP50		
			RAPGEF2	BTG1	LSM14B		STARD7		BCORL1		
			TSR1	TRMT10C	NAPA		UBN2		IPO5		
			SEC23B	VAT1	RFWD3		GATAD2B		AIF1L		
			BLOC1S6	TLDC1	TCF7		MYO1E		RARA		
			MSL1	UBE4B	XPO1		SLC35E3		SH3RF1		
			AOX1	NUCB1	PHIP		PXDN		PPP1R9B		
			TMEM243	REEP5	TMEM194 (hsa)		RAB22A		SLC25A11		
			MRPL30	PARD6B	ZNF652		NHSL2		RAP1GDS1		
			LRRC58	SERBP1	NFATC3		SLC31A2		C6orf89		
			SLC27A1	TRIM32	RSE_00000212796 (hsa)		TNC		TRIO		
			DYNC1LI2	DNMBP	DUSP16		AP3S2		BMP2		
			CISH	MTCL1	DNAJB4		TEAD4		AUNIP		
			PAICS	HMGB1	SPTBN1		ZNF619		SLC36A1		
			CCDC28B	DIS3	AGL		RBM12B		CNOT3		

					HIF1A	TAF3	Contig31122_RC (hsa)		CDKN1A			KIAA1107			
					PPP1R10	ZNF623	KIAA1946 (hsa)		LHFP			SRSF4			
					HELZ	RAB20	CD55		DCAF15			CITED2			
					POGZ	MYL12A	KIAA1841		BIRC3			KDELR1			
					PAIP2	ZBTB25	IRAK1		RALB			PIK3C2A			
					RGMB	PDP2	IRAK4		CCNJ			CIPC			
					SACS	NOC4L	DAAM1		KIAA0232			RDH10			
					C2CD2	ZNF675	PLK4		C3orf38			CHD2			
					CTNNBL1	KPNB1	EMC1		MFSD8			YEATS2			
					C8orf4	RHOB	DEAF1		CHTF8			MRPL3			
					SLIT1	CBX5	CDKN2D		SDC4			SKA1			
					ANP32B	TNKS	GRHL1		HIAT1			NLGN1			
					SEC14L1	MXRA7	IL6		TNRC6A			TSC22D2			
					MAPK8	SNRPC	RAB27A		MAP3K2			PYGO2			
					KPNA5	JMJD6	RNF141		ARF6			PDCL			
					ICE2	ZNF91	MFI2		KHDC1			SSX2IP			
					PPAP2B	RAB3C	NIPBL		MBNL1			SNRNP40			
					STAT2	DYNC1I2	PLOD1		NFIA			ZBTB21			
					PDE8A	SPOPL	GPR126		MON2			DHX15			
					DNAJB9	IREB2	ERCC6L		ZFHX4			KDELC2			
					RIT1	BZW2	UACA		MYBBP1A			BCL2L12			
					HERC2	ASXL2	EMB		INPP5A			ZC3H11A			
					CYP24A1	TGFBR2	WRNIP1		ZMAT2			ADNP			
					CSNK1A1	C3orf52	RAD21		ARSJ			HSPA8			
					CREBRF	GLYR1	SYNGR1		PLD3			CCNY			
					CGGBP1	RAB14	HMMR		ATG9A			TOPBP1			
					DNAJC13	MAML1	USP18		SCYL1			STRN			
					KIDINS220	PON2	AB033076 (hsa)		EIF4EBP2			MADD			
					LNPEP	SNRNP200	GULP1		MLXIP			SMARCD2			
					MED26	SGK1	PLAU		LDLR			ZNF398			
					AFF3	NAA38	FGFR1OP		FAT1			EIF2S1			
					STAU1	RAPH1	LEAP2		EIF1AD			DSG2			
					OGG1	GALNT16	PUS1		PRDX6			SLC50A1			
					FOXN3	RAB5B	HOXA10		RPRD1A			TFDP1			
					RMND5B	CAMK2N1	CCDC174		TPT1			HIVEP1			
					SRPK1	AKAP11	RNF26		TTI1			SP1			
					HOXA5	GLYCTK	UGT8		GORASP2			MRPL14			

					LIMK2	MDM2	hCT2309798 (hsa)		HMGCS1			RARS			
					MRPL16	ZNF669	AK075186 (hsa)		AASDHPPT			B4GALT3			
					HAX1	TNRC6C	RAD9B		LIMA1			INCENP			
					NR2C2	N4BP1	TMEM136		SLC25A27			MCRS1			
					KIF4A	ABL1	CBX3		EMD			CCDC15			
					CENPF	PLCB4	Contig753_RC (hsa)		PCBP2			NCSTN			
					SULT4A1	FAM98B	CCDC176		RASSF8			GOT2			
					IKZF5	BMP4	MSRB3		EDN1			ADRM1			
					SON	MEIS1	LYSMD3		PRKDC			SLC39A6			
					HMCN1	VKORC1L1	AB002442 (hsa)		DGKD			PROSC			
					PIM3	BAZ1B	TOP1		SMAD7			GTPBP2			
					HCFC1	SMG1	SGCB		SRGN			MRPL24			
					SF3B2	CCZ1	PSMD6		E2F4			MAT2A			
					CD46	MCM3	KIF5C		ENTHD2			ZNF451			
					AHCTF1	STK24	NEIL3		TNFSF9			GNPTAB			
					RGL2	BMPR2	hCT1776373.2 (hsa)		RAB8B			AP3D1			
					FBXO3	DMXL2	ENTPD7		UBE2W			CNOT2			
					PRKAG1	RTN3	AF125104 (hsa)		SLC15A3			AHSA1			
					AHDC1	CCNL1	Contig48208_RC (hsa)		SPHK1			SERF2			
					FAM83D	FLVCR1	COQ3		PLK1			YBX1			
					MKLN1	TOR1AIP1	Contig41618_RC (hsa)		ST14			ARRDC3			
					PLK2	ATF4	ZNF248		MTOR			USP42			
					GALNT1	NNT	TRAIP		VPS25			NFE2L2			
					TMEM91	ZNF141	PODXL		PIEZO1			BBS10			
					MYO9A	CLOCK	OTUB2		DLAT			MBNL2			
					STRN4	KIF5B	AK093202 (hsa)		PPM1G			STX6			
					EDNRB	NCOA1	GALNT12		AURKB			SLC20A1			
					AGAP2	HLA-A	40057		ZNF3			MELK			
					SFXN5	CCDC6	SPOP		RAB5C			LAMC1			
					VCAN	HCCS	Contig53126_RC (hsa)		GPR157			PRICKLE2			
					USP21	EPB41L3	ZNF165		ABCC5			USP24			
					MTCH1	MLLT10	Contig44271_RC (hsa)		ANAPC13			RAI14			

					USP46	METAP2	ZFP30		HYAL2			NEDD4			
					WSB1	CDC42BP B	NOB1		SRSF3			GMEB2			
					ACO1	RPS6KB1	SLC1A5		ZCCHC11			MKI67			
					TP53INP1	PLEKHA2	NA		ARHGEF1 7			TSPAN14			
					INSIG2	CD44	C1QTNF3-AMACR		RNF6			UBE2H			
					GPAM	TFAM	C15orf40		SNX25			SCOC			
					MLH3	PRKACB	SLC39A1		BCAR3			RRP1B			
					ITFG3	ATR	THOP1		LEPROT			TGFBR3			
					FXYD6	MAP1B	FERMT1		NR1D2			GSTA4			
					LIFR	R3HDM2	RAB14		CLDN12			PRKAR1A			
					LBH	DCHS1	ZNF641		MAN1A1			PEF1			
					POLR1B	KIAA1430	TMEM200A		UHRF1BP 1			TLK1			
					BCL7B	COL4A1	MED22		SRSF10			UBL4A			
					SUB1	SLC30A1	MMP20		PTAR1			PROSER1			
					DIAPH1	UBR4	Contig19213_R C (hsa)		KREMEN 1			CSNK1E			
					PTP4A2	ZKSCAN1	HCN4		CYB5R1			EP300			
					TMEM9B	NACC2	PRRC2A		MOB1A			AMOTL2			
					CHERP	CDC25A	BC040287 (hsa)		DOCK7			BRE			
					SPRYD3	PPP1CB	RP11-679B17.1 (hsa)		CDC42SE 1			DDAH1			
					PRKCSH	C2orf69	PTS		CHST1			SLC10A7			
					CLIP4	BNIP2	AL049452 (hsa)		GFOD2			TMEFF2			
					RELN	C2orf43	EXO1		IGDCC4			PGRMC2			
					ZFP36L1	TRIM28	RAB5B		ATP1A1			PPARG			
					COL5A1	ATP5O	CLK4		LUC7L			TMEM245			
					PLEKHA3	CAPN7	AKAP11		SUCO			ST13			
					GOLGA4	LRFN5	SLC46A3		SNF8			MAL			
					GPBP1L1	SNUPN	MDM2		UTRN			ZCCHC14			
					BLOC1S5-TXNDC5	IL6R	SP9		MAP2K4			ARFIP1			
					KLF3	KDM5B	CMPK1		SAMD9L			ICK			
					ATXN1	GSPT1	AK054764 (hsa)		PBX2			HARS			
					UBA1	AGPAT9	MTDH		SYMPK			SLC44A1			
					EPS15	NDEL1	GPR137C		ITPR3			DNAJC21			
					ESM1	ADCY6	TOMM6		DNAJA2			FBXO7			
					PRDX1	TUSC2	NIN		CEP290			TP53BP1			

					STK35	RMND5A	KIF24		SPHK2			FAM73A			
					PPP2R1A	TXNIP	FANCI		RFWD3			S100PBP			
					FAM126A	PIP4K2B	MCM10		GTF2A1			PUM2			
					IGF2R	ZNF230	hCT1645074.2 (hsa)		XPO1			ZNF780A			
					DYNLT1	NUDT3	NIPA1		TRIM71			NFIB			
					KANK1	TADA2A	CAB39L		UNC13B			REV3L			
					UBAP1	MARCKS	AUP1		TUBGCP6			EXOC2			
					SLC16A1	ARHGAP23	NM_144710 (hsa)		MTF2			COPG2			
					CLPX	ZNF792	FAM110C		FGFR1			BCL2L1			
					GALNT3	ZNF99	TMC3		TRAM1			TWISTNB			
					SOCS5	SYNRG	CCZ1		FUCA1			C12orf57			
					E2F6	ZNF729	BMPR2		SPTBN1			AKT3			
					SKIL	ASF1A	RPS29		POLR1A			RSBN1L			
					TFRC	NPHP3	hCT2261744 (hsa)		HNRNPUL1			KIF20B			
					BIRC6	MLLT6	FLVCR1		PHF21A			GLTSCR1			
					IKBKAP	AC004076.7	TOR1AIP1		THNSL1			STON2			
					SLC25A23	PECAM1	ATF4		SLC35F6			PIAS2			
					BLVRB	BRD4	GINM1		NFIC			ATF7IP			
					FLNB		SEC24D		RNF213			BRD3			
					DSN1		AP3M2		CTPS1			UBE3D			
					DAPK1		AC099522.1		INF2			NVL			
					BMP2K		NR2F2		DCAF16			WDR3			
					NEK6		RASL11B		ATP13A3			DYNLT3			
					JRKL		KIF5B		TRIB1			HPRT1			
					SPAG9		CXorf10 (hsa)		ZBED9			YIPF5			
					SUPT6H		hCT1812405.2 (hsa)		MAPK1			BMPR1A			
					MAPKAPK3		TCHP		MFI2			SDHA			
					TMED10		PLS3		PNRC1			CNOT1			
					ATAD2B		HOXA13		TTLL9			ZNF281			
					REL		RPS6KB1		FNDC3A			ATG3			
					IFI16		BBS7		DNTTIP2			USP14			
					SRPR		USP34		POMGNT2			HERC4			
					FRMD6		AF086561 (hsa)		NUPL1			WHSC1			
					HAUS6		XIAP		RAD54L			STARD13			
					DMTF1		hCT2256635 (hsa)		DDX26B			COX15			

					WT1		CA8		LRRC41			UBE2E1			
					GLO1		CTGF		NUP205			PHF12			
					VCP		ZKSCAN1		FGF1			TOPORS			
					CBX6		SENP5		AMOTL1			CREB3L2			
					CHD3		CDC25A		SRF			ATXN7			
					ZMYM1		PPP1CB		MAPK6			TSPYL1			
					ECT2		BNIP2		CREBBP			SCAF11			
					XPR1		PEX1		PPP2R1B			CYB561			
					MAFK		SYNCRIP		SESTD1			UBE2G2			
					ANKLE2		hCT2284819 (hsa)		RNASEH2 B			RPL37A			
					ZBED6CL		SCMH1		POLR2G			ARMC1			
					ZMYM2		TOR3A		GREB1			BIVM			
					SMURF1		DSP		TMPPE			SEC24B			
					GPR56		XM_208797 (hsa)		CSF1			BRD1			
					APLP2		Contig50272_R C (hsa)		GNG5			ADCK1			
					GIPC1		IL6R		DHCR7			NAP1L3			
					NDUFV3		hCT2259022 (hsa)		DIAPH2			STAMBP			
					TM2D3		BIRC4 (hsa)		SLC43A3			CERS6			
					GMPS		DGCR11 (hsa)		PCGF3			G3BP2			
					PRNP		TXNIP		AP3M1			SMC2			
					HSP90AB1		HIST2H4A		KMT2C			PCDH7			
					ZFYVE9		HIST1H4J		WDR77			CEP170			
					NFKBIA		HIST1H4B		CBS			SH3BGRL2			
					MYO5A		PIP4K2B		MGAT4A			SUZ12			
					BAZ1A		TBC1D22B		PUS1			DNAH14			
					HIC1		CPNE4		ADCY9			ARCN1			
					ZNF100		RP5-850E9.3		ULK2			HNRNPH1			
					SNAPC4		NUDT3		HELLS			CCNE1			
					RBBP5		ZNF160		HOXA10			GNL2			
					NIPA2		ZNHIT3		CCDC174			AKAP13			
					CEMIP		DUSP14		USP22			CDC27			
					PPP4R1		ZNF587B		KATNAL1			KPNA6			
					SHISA5		MLLT6		SERTAD2			GNAI2			
					RAB18		BGLAP		RNF26			IGFBP3			
					TMEM87B				DLGAP4			CD300LG			
					C17orf58				UBE4B			S100Z			
					DENND6A				BAP1			RALGDS			
					HIST1H2B G				RAB10			RACGAP1			
					XYLT2				PATL1			ZNF211			

					KDM3B				PARD6B			TPST2			
					JMJD1C				ZNF567			C16orf70			
					COPZ1				TRIM32			ZBTB33			
					RNMT				MTCL1			SHMT1			
					GLUL				ALG3			C5orf24			
					RBM26				SETD1A			ABAT			
					GBAS				ARF4			ZNF257			
					GALK2				ZNRF3			SCD			
					KIAA2013				GNB4			SEH1L			
					TLE4				ZNF623			SAMD4B			
					MID2				CERS4			PDIK1L			
					NCAPD3				PDP2			LAPTM4B			
					CLK1				RNF139			HOMER1			
					ZNF138				GPR161			ACSL4			
					H2AFZ				STK40			KLHL9			
					C5orf51				RHOB			MED23			
					TRIM33				SIM2			BAK1			
					ELOVL5				BRWD1			UBE2G1			
					MMD				CBX5			CYB5B			
					CDC42BPA				STARD4			CEP162			
					ARIH2				SOX13			PCMTD1			
					PATZ1				RAD9A			AURKA			
					DAG1				JMY			RLIM			
					RMDN3				DMKN			ESPN			
					CACNG4				TBL1XR1			TATDN3			
					RNGTT				WDR37			ATG5			
					RP11-385D13.1				DNASE1			ATXN2			
					NCOA2				SNRPC			NAPB			
					POP5				LETM1			ACTR1A			
					SAP130				ZNF410			PLEKHA1			
					PABPC1				GATA6			SGOL1			
					CDYL				TIMM17B			KIAA1244			
					EI24				SOX12			ROGDI			
					MEPCE				ARHGAP28			COQ10A			
					RBM47				BZW2			KCNG3			
					THAP11				TIE1			VEZF1			
					CDCA7				TGFBR2			UBAP2L			
					HOOK1				EIF1			PRMT5			
					KLHDC3				C3orf52			TRIM5			
					SEC62				BICC1			CTTNBP2NL			

					NOL11				RAB14			BEX4			
					SEC24A				MAML1			GPI			
					PIK3R1				ZFAND3			LASP1			
					NEO1				USP10			ZHX2			
					PCMTD2				PPARD			MECOM			
					ARMCX3				SNRNP200			FOXO3			
					TEAD1				COPS6			XPO4			
					PNRC2				PUF60			IRF2			
					SCYL2				MARCH3			VPS37C			
					GEM				RAPH1			ZNF585A			
					ERP29				FADS3			PPP1R8			
					BAZ2A				JAK1			ABLIM1			
					SATB1				EXO1			SLC7A2			
					SMARCC1				RAB5B			TULP4			
					TGIF1				MAP7D1			C12orf4			
					ASH2L				ZNF142			STARD7			
					FKBP8				MDM2			ATF3			
					MARCKSL1				ELOVL4			GATAD2B			
					TMEM185B				SH3TC1			BASP1			
					WDTC1				WNT9A			CDIPT			
					KDM5A				MTDH			GFRA1			
					SEPP1				NIN			RRP12			
					CPEB2				PRR11			SPTSSA			
					TMCC1				SCML1			RAB22A			
					RORA				RNF216			PKN2			
					SLC22A23				ING1			COBL			
					CDK18				AP5M1			COG3			
					MRPL43				NIPA1			AKAP12			
					PTPN11				AUP1			MCCC2			
					BUD31				STK36			ELOF1			
					HSP90B1				VPS54			CDKN1A			
					MAD2L2				MRPS31			MED20			
					ARPC1A				EHMT1			UHMK1			
					CDC42SE2				COIL			DCAF15			
					C2orf49				SMG1			PKD2			
					FARP2				MCM3			KIAA0232			
					TMEM39B				RPS29			SYNE2			
					PHC2				TMBIM6			SDC4			
					DYNC2LI1				PGAM5			GPT2			
					GTF2H2				MVP			TNRC6A			
					ERAL1				H1FX			ULBP2			
					NRP1				NNT			MAP3K2			

					CRELD1				EIF4G2			RSPRY1			
					MPP6				RHOG			C1orf109			
					PRKAR2A				GTF2E1			MBNL1			
					PRKX				SYT16			CHEK2			
					ZNF592				NR2F2			ISL1			
					TUT1				ZMYND8			WDR4			
					BHLHE22				SLC27A4			TMEM201			
					ZC3HAV1				ZNF740			NFIA			
					TAF4B				ZDHHC7			GABPB2			
					SRRM2				RNH1			EIF6			
					SS18L1				ANKMY2			TBP			
					DLGAP5				WBP11			HIST2H2A C			
					AGPAT6				PAX3			ABHD12			
					KLHL15				ADAM15			COBLL1			
					FXR1				ARHGEF2 8			OGFOD1			
					SLC30A9				XIAP			KIAA0930			
					PPIL4				MAP1B			ELL			
					PEX11B				SH2D3A			TNFAIP1			
					SENP1				SOX5			PTPMT1			
					JAZF1				VAMP2			RFC5			
					GNPNAT1				QRICH1			ELOVL6			
					GALNT10				SMPD4			CHAMP1			
					METTL9				CDC25A			SNX6			
					EPHA2				FNIP2			ALG13			
					MID1IP1				PEX1			LRR1			
					ASCC3				BIRC2			MRPL42			
					EIF5B				RIMKLA			EIF4EBP2			
					CERS5				MEX3C			RRM2			
					LARP4				ADPRH			SMG7			
					BACE2				INO80D			VGLL4			
					PTPRJ				RIN3			CKAP2			
					DIS3L				KDELR2			TGIF2			
					KITLG				FHOD1			SESN3			
					TNFRSF21				KDM5B			COPA			
					KBTBD11				C9orf142			KLHL12			
					SYNE1				GSPT1			FAT1			
					SRPX2				COA7			OTULIN			
					IRS1				SMCR8			C17orf51			
					WDFY3				TUSC2			FOCAD			
					LZTS2				RMND5A			HRSP12			
					GLCE				ZCCHC3			DCAF12			
					RAC1				KMT2B			PDPK1			

					NRP2				SOCS7			COQ6			
					SNX2				HIST1H2BB			PRDX6			
					GABPA				LHX1			TMEM230			
					CD99L2				NR4A1			ZCCHC8			
					FAM120C				FPGT			LRP3			
					TNFRSF25				MARCKS			ADARB1			
					SPRY4				PCGF2			SETD9			
					CHRAC1				ZNF792			PPM1B			
					ANKRD13C				RASL10B			ZBTB4			
					RANGAP1				VAMP3			PEX12			
					DUSP8				BIVM-ERCC5			HMGCS1			
					STC2				HIST1H4D			PTPN12			
					MAPK3				URGCP-MRPS24			MMACHC			
					PRKAA1				DDX52			BARD1			
					CAP2				ZNF280B			BCOR			
					UBXN2A				BRD4			ZNF615			
					CTNNAL1							TUB			
					ANKFY1							CCNI			
					FAM208A							BTBD7			
					SDAD1							OPA1			
					NOLC1							GID4			
					MESDC2							DDIAS			
					ATP6V1A							GPM6B			
					CDC42							CACNA2D1			
					KMT2D							PHF13			
					SEPN1							PRPF18			
					ADAMTS1							MAP1A			
					DNAJB14							CCNK			
					FSCN1							PCBP2			
					PAK6							RASSF8			
					KDM6A							AHNAK			
					UBE2K							38961			
					NUP50							TRPM7			
					KDM6B							LUZP1			
					BCORL1							ZSWIM5			
					PEA15							ZBTB34			
					IPO5							AP1G1			
					IPO4							SMAD7			
					LDHB							PAFAH1B2			

SESN1	CIZ1
AIF1L	CEP55
TSEN34	MKNK2
SESN2	RBM4B
VPS52	ZIC5
SDE2	DCLRE1B
HSPA4L	HERPUD2
YRDC	PLK1
POLR3C	ZNF367
KIF11	DUSP5
CS	GDI2
FRYL	RPL35
C2CD4A	TRRAP
COMMD2	ARPC4
DSC3	ZFP36L2
TIA1	OSBPL8
PIK3C2A	DLAT
CIPC	NEDD1
C16orf72	HIST1H1B
KCNJ6	PHF5A
MRGBP	RAB5C
RPL19	KIAA1467
TSC22D2	NCKAP1
TXLNG	NFYC
KIAA0247	MAFG
TMEM2	DYRK1A
LYPLA2	HYAL2
ADCK3	FUT8
DNAJC12	SKA3
MIOS	MOAP1
SNRNP40	TMEM70
ZBTB21	WIPI2
DHX15	ALDOA
BAG2	LEPROT
DYNLL2	AGTPBP1
AP2A1	HEBP2
CALD1	NR1D2
ADNP	EXOC8
RPS27L	SMARCD3
C6orf120	MOB1A
NF1	PRDM4
ZNF430	ZNF574
STRN	LPHN1
HMG20A	CLUAP1

					CDK13							VCPIP1			
					KIAA0100							HEY1			
					UBAP2							CAMK1D			
					NIFK							SPIN1			
					CACYBP							EXOSC9			
					RANBP2							SUCO			
					SUV420H1							UBE2B			
					DTNA							MAP2K4			
					ZFP1							FLOT2			
					SPSB1							RYBP			
					HIVEP1							SSBP3			
					SP1							PPARGC1 B			
					ARID5B							OAT			
					B4GALT3							RELA			
					DUSP7							RSBN1			
					ABCB7							ZNF22			
					COL1A2							RND3			
					JAG1							CBR4			
					FAM98A							VAV3			
					MDM4							ITPR3			
					OGFRL1							TMCO3			
					CSNK1G1							NME1			
					ACVR1C							FAM188A			
					KRT80							B3GNT2			
					SLC39A6							CHMP1A			
					FUT11							GTF2A1			
					SMIM13							DTNBP1			
					PLXND1							TRIM71			
					C19orf10							TSPYL5			
					ONECUT2							FBXO33			
					KLC1							NFATC3			
					TWSG1							SUPT7L			
					AGAP5							DUSP16			
					AKIRIN1							PALLD			
					NPTN							SGK3			
					TNRC18							DDIT3			
					ZNF451							ERRFI1			
					GNPTAB							PHF21A			
					MTERFD3							LATS1			
					TLE3							FOXO1			
					AP3D1							CENPU			
					CNOT2							PLEKHM1			
					RASSF3							AP1B1			

					TAF7							IRAK1			
					USP42							DAAM1			
					NFE2L2							SOX2			
					EPC2							PLK4			
					ANAPC5							RNF213			
					GFM1							TMEM214			
					MBNL2							BBX			
					STX6							DCAF16			
					DOT1L							LMAN2			
					SLC20A1							RAD17			
					PPP6R1							ATP13A3			
					NDUFA8							PPFIA1			
					LAMC1							NSFL1C			
					PRICKLE2							LAMTOR1			
					USP24							NUP188			
					RAI14							BRI3			
					RPS6KA3							SLAIN2			
					NFE2L1							NUPL1			
					NEDD4							BRF2			
					FOXJ3							TLN1			
					ERC1							CCDC59			
					CDK19							SPTSSB			
					ATP5E							MICAL3			
					MKI67							PHKB			
					SCOC							RAD51AP1			
					IGF1							HIST1H2A C			
					ELP6							SRF			
					TGFBR3							TRIT1			
					CNTLN							CREBBP			
					UBE2F							FBXW11			
					LAMP2							GATAD2A			
					PRKAR1A							CHD9			
					SUMO3							POLR2G			
					FAM160B1							RFC3			
					UBL4A							TMPPE			
					KIF26A							H1F0			
					MIDN							ZHX1			
					AMOTL2							DHCR7			
					BRE							DIAPH2			
					PRDM2							ZNF33A			
					ZNF714							NKIRAS2			
					DDAH1							SCARB2			
					SLC10A7							AP3M1			

					FCGRT							CIRBP			
					PTPRA							KMT2C			
					AHSA2							LPCAT3			
					BTG3							GOLGA3			
					PPARG							HOXA10			
					MYT1L							KATNAL1			
					POU2F2							RNF26			
					PAAF1							VAT1			
					MGEA5							SEPHS2			
					DENND2D							CASC3			
					ARL5B							UBE4B			
					CASP8							PATL1			
					NOTCH3							REEP5			
					TTLL4							ZCCHC17			
					PLAG1							PARD6B			
					PUM2							SDHD			
					NFIB							TMUB1			
					THUMPD1							VPS72			
					C2orf47							FNBP4			
					TWISTNB							ARF4			
					RASEF							WDR83OS			
					C12orf57							ZNRF3			
					CHML							TMUB2			
					OSBPL10							KIAA1279			
					STON2							ENAH			
					ATF7IP							ZCCHC24			
					IGFBP2							GMEB1			
					MEDAG							TNPO3			
					PLEKHB1							GPR107			
					COL11A1							NUDT11			
					GJC1							MAPKAPK 2			
					LY6G5B							PCBP4			
					POMK							ENTPD7			
					CNOT1							RNF139			
					ZNF281							MRPL19			
					EFNB1							STK40			
					ATG3							KPNB1			
					CEP135							FDX1L			
					UBE2D3							RHOB			
					ZZEF1							BRWD1			
					STARD13							TTPAL			
					PHF12							RPE			
					TSHZ1							NOP58			

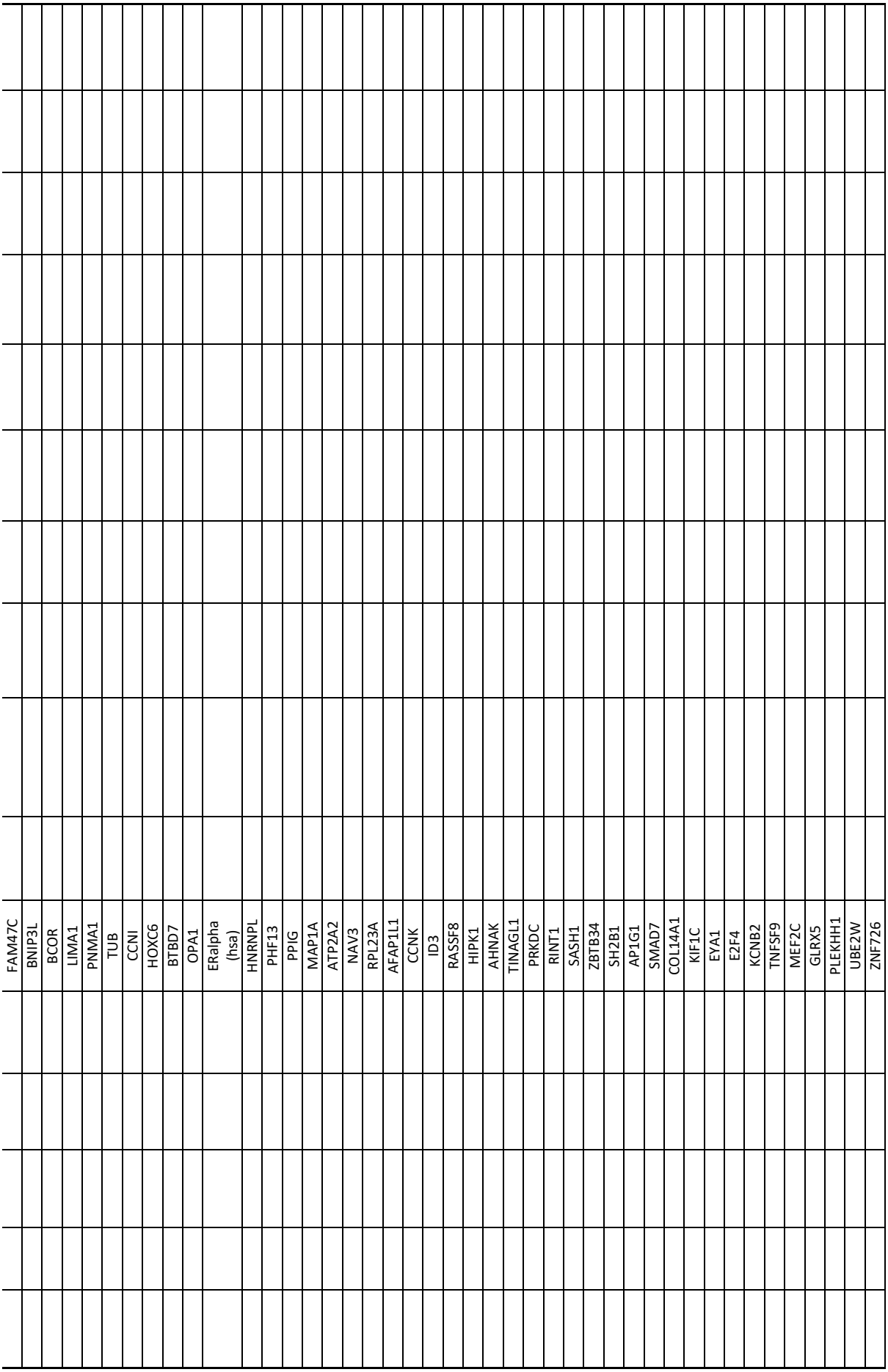
					PPARGC1A							PIAS1			
					TSPYL1							TBL1XR1			
					SCAF11							WDR37			
					CBY1							SETX			
					CYB561							NAA60			
					CDH5							ZNF91			
					CKAP4							RAB3C			
					PARP1							SPOP			
					UBE2G2							SPOPL			
					H3F3B							TRMT10A			
					GPR160							FOXF2			
					SEC24B							PIGA			
					MPHOSPH 9							IREB2			
					CAPN1							CCP110			
					SLC4A8							ARHGAP2 8			
					CLIP2							GOLPH3			
					CNN3							SNRPB2			
					TMEM150 C							ASXL2			
					CD99							KLHL18			
					GAPVD1							MEF2A			
					G3BP2							ISY1			
					RBBP7							SLC39A1			
					B4GALNT1							DDHD2			
					TP53I11							CENPN			
					ZFR							RAB14			
					TTC39A							MAML1			
					SUZ12							MTRNR2L 8			
					TRAPPC10							ZFAND3			
					HNRNPH1							USP10			
					OAZ2							PVRL1			
					ABCA5							PPARD			
					CCDC88C							RPL15			
					FAM210A							ETFA			
					RC3H1							ARHGAP5			
					SPX							RAPH1			
					CDC27							UBE2C			
					KPNA6							KHDRBS1			
					IGFBP3							RBM15			
					FASTKD5							RAB5B			
					CAPN15							STX2			

					MYCT1						AKAP11				
					ZNF81						ZNF142				
					GALNT2						ELK1				
					LONRF2						PSMB5				
					PJA2						BAG5				
					ZBTB33						PRR11				
					FAM193B						PPAP2A				
					ABI1						TNRC6C				
					VPS39						N4BP1				
					DDA1						RNF216				
					ABAT						ING1				
					NCDN						AP5M1				
					ATP6V1C1						KIAA1715				
					NGLY1						PCMT1				
					CYP1B1						VKORC1L1				
					LAPTM4B						EIF2B2				
					KLHL9						BAZ1B				
					MARCH5						MCM3				
					UBE2G1						PPFIA3				
					SNRPB						BMPR2				
					TFIP11						RITA1				
					RCBTB1						KIAA1551				
					SREK1IP1						PNISR				
					RLIM						TMBIM6				
					IL6ST						ARHGAP1 2				
					SEMA7A						GNPAT				
					ZNF462						ATF4				
					THOC2						RPL12				
					NAPB						SEC61G				
					ENPP1						NNT				
					FAM102A						EIF4G2				
					LEF1						BCL2L11				
					OCRL						CLOCK				
					MRPL28						RHOG				
					KMT2A						ZMYM3				
					RPN2						KIF5B				
					EPS8						MTF1				
					VEZF1						HOXC8				
					UPF2						LMO4				
					BAMBI						CCDC6				
					CLK2						AJUBA				
					PRMT5						ZNF740				
					PACSIN2						EPB41L3				

ZER1	MYO1B
ALDH9A1	ABCF1
ZNF682	SIN3A
GPI	PLS3
TPP1	ZNF689
LIMK1	WBP11
LASP1	USP34
RLF	AP2M1
MECOM	ATP5C1
FOXO3	PRKACB
DHX9	KCTD20
ANKRD28	MOSPD2
DERL1	TBXA2R
IFNG	MIER3
FBXO22	VAMP2
LMNA	DHX33
SLC7A2	FAM134B
CANT1	SIAH2
TULP4	PFKFB3
FN1	LHX2
FAM136A	SLC30A1
STARD7	POLH
IFFO1	TRIM28
37500	NCALD
IPO8	MSX2
RNASET2	MTFR1
ERI3	MEX3C
ANXA11	ATP5O
CTCF	R3HDM1
GFRA1	DSP
RRP12	KDELR2
PXDN	KIAA1586
RAB22A	ABRACL
TMEM218	ADCY6
TXN2	TUSC2
RAD54L2	RMND5A
SEMA4D	ZCCHC3
LRIF1	RBM15B
TNC	TXNIP
ROBO1	HIST1H4I
AKAP12	PIGW
CALR	SEC22B
ELOF1	CSNK2A3
EIF4G3	SRGAP2

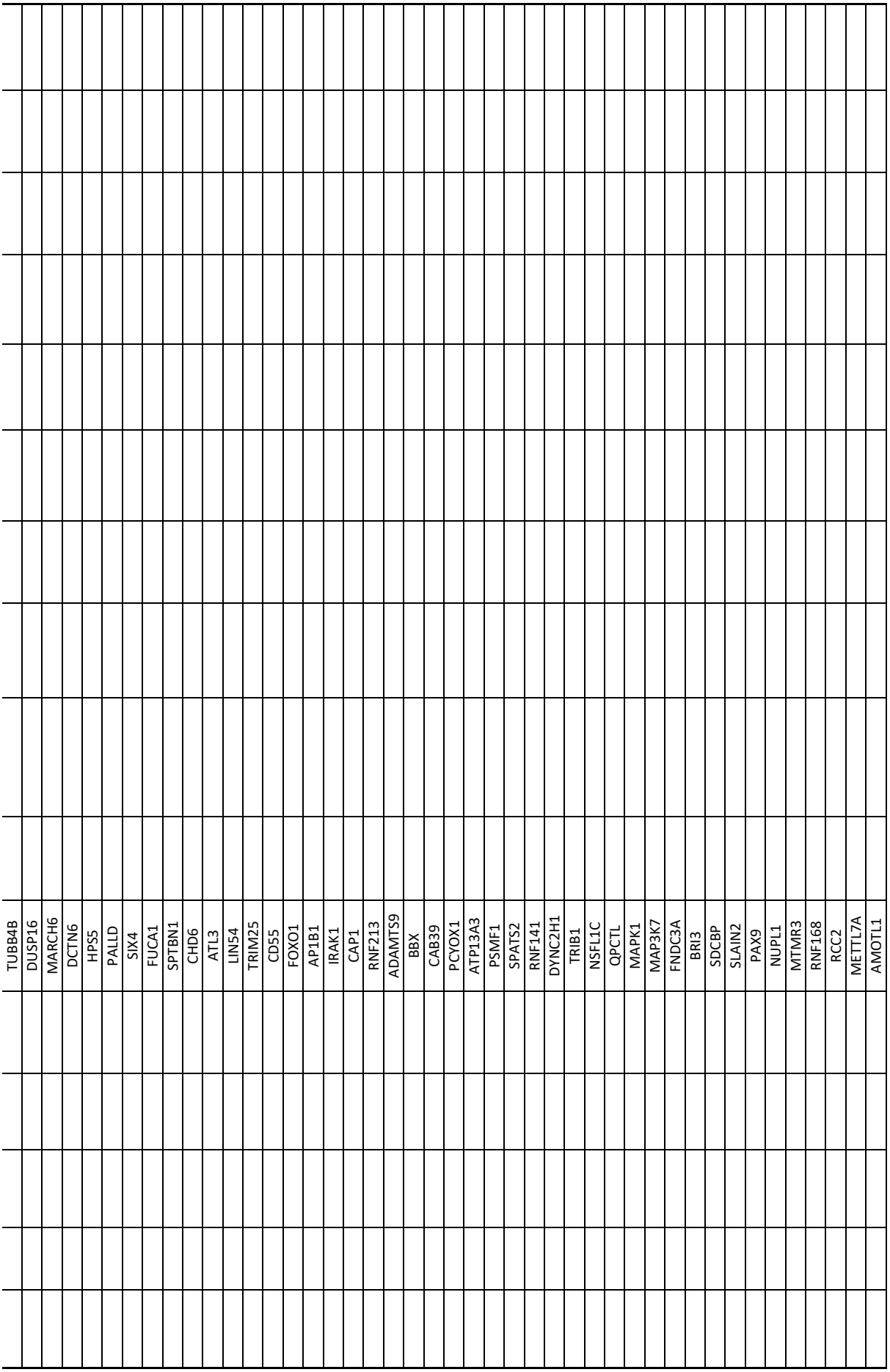
					TMEM30A							LHX1			
					RBM12B							RSC1A1			
					CDKN1A							RP5-850E9.3			
					ERH							NUDT3			
					PCIF1							C5orf22			
					STMN1							HIST1H3E			
					UHMK1							NCOA4			
					RALB							NR4A1			
					POFUT1							PALM2-AKAP2			
					NAPEPLD							MARCKS			
					SYNE2							ZNF613			
					CPT1B							ZNF792			
					C3orf38							VAMP3			
					GTPBP10							ACACA			
					CDC42EP3							TFDP2			
					GPT2							GPR179			
					DOCK9							ZFP91			
					TNRC6A							MLLT6			
					MDN1							CTDSP2			
					ZNF662							CWC25			
					AKR7A2							FAM58A			
					USP9X							TMEM185A			
					IRX3							ORAI1			
					MBNL1							ZNF280B			
					ISL1							PSMB3			
					WDR4							RNF115			
					TMEM201							BRD4			
					EIF6										
					TBP										
					MYBBP1A										
					COBLL1										
					ATP2B4										
					LTBP1										
					NOL10										
					PAN3										
					ANKHD1-EIF4EBP3										
					MKNK1										
					ZMAT2										
					MCAM										
					SEMA6D										

					KIF21A										
					PRKD3										
					ABCA12										
					ZNF579										
					CHAMP1										
					HSD17B12										
					KIAA0586										
					ZBTB18										
					REEP1										
					PRKCH										
					PPA1										
					HHLA3										
					MON1B										
					CARM1										
					TUBA4A										
					MLXIP										
					KIFC3										
					PDZD11										
					RRM2										
					SMG7										
					VGLL4										
					CKAP2										
					SESN3										
					LDLR										
					FAT1										
					TSEN54										
					C17orf51										
					HNRNPA2B1										
					TEX15										
					FOCAD										
					HRSP12										
					PDPK1										
					NOP9										
					PPM1H										
					NUP155										
					TPT1										
					TRIM44										
					DKK1										
					LRP3										
					ZBTB1										
					POLA1										
					FIP1L1										
					HMGCS1										


FAM47C
BNIP3L
BCOR
LIMA1
PNMA1
TUB
CCNI
HOXC6
BTBD7
OPA1
ERalpha (hsa)
HNRNPL
PHF13
PPIG
MAP1A
ATP2A2
NAV3
RPL23A
AFAP1L1
CCNK
ID3
RASSF8
HIPK1
AHNAK
TINAGL1
PRKDC
RINT1
SASH1
ZBTB34
SH2B1
AP1G1
SMAD7
COL14A1
KIF1C
EYA1
E2F4
KCNB2
TNFSF9
MEF2C
GLRX5
PLEKHH1
UBE2W
ZNF726

ABCA1
TMEM127
RAN
NDUFV1
MKNK2
PSRC1
FANCA
ZIC5
DCLRE1B
STAT6
STAB1
PDE7A
PALM2
PPIF
ADD1
GDI2
JADE1
MTOR
TRRAP
ARPC4
PRR13
VPS25
ZFP36L2
PPM1G
PLEKHA6
PHF5A
ZNF106
NCKAP1
MAGT1
GATAD1
DYRK1A
CTDP1
CMTM6
ABHD17C
FOSL1
MOAP1
WIPI2
C1orf115
ALDOA
BCAR3
NR1D2
SLC35B4
PTAR1
DENND2C

					EXOC8										
					SLC30A2										
					METTL3										
					MOB1A										
					ZNF574										
					NFASC										
					VCPIP1										
					ITGA6										
					ZFY										
					SGPP1										
					PDZD8										
					ATP1A1										
					PFKL										
					SUCO										
					ST20										
					UBE2B										
					UTRN										
					MAP2K4										
					XPC										
					EHD1										
					RUSC2										
					AKAP6										
					SHH										
					ZNF518A										
					RYBP										
					RND3										
					PTRH2										
					TMCO3										
					COL5A3										
					NME1										
					ZNF618										
					GTF2A1										
					TANC2										
					XPO1										
					PHIP										
					FAM171A1										
					FBXO33										
					SCNN1A										
					ZNF408										
					CPEB3										
					PCF11										
					ZC3H4										
					MTF2										

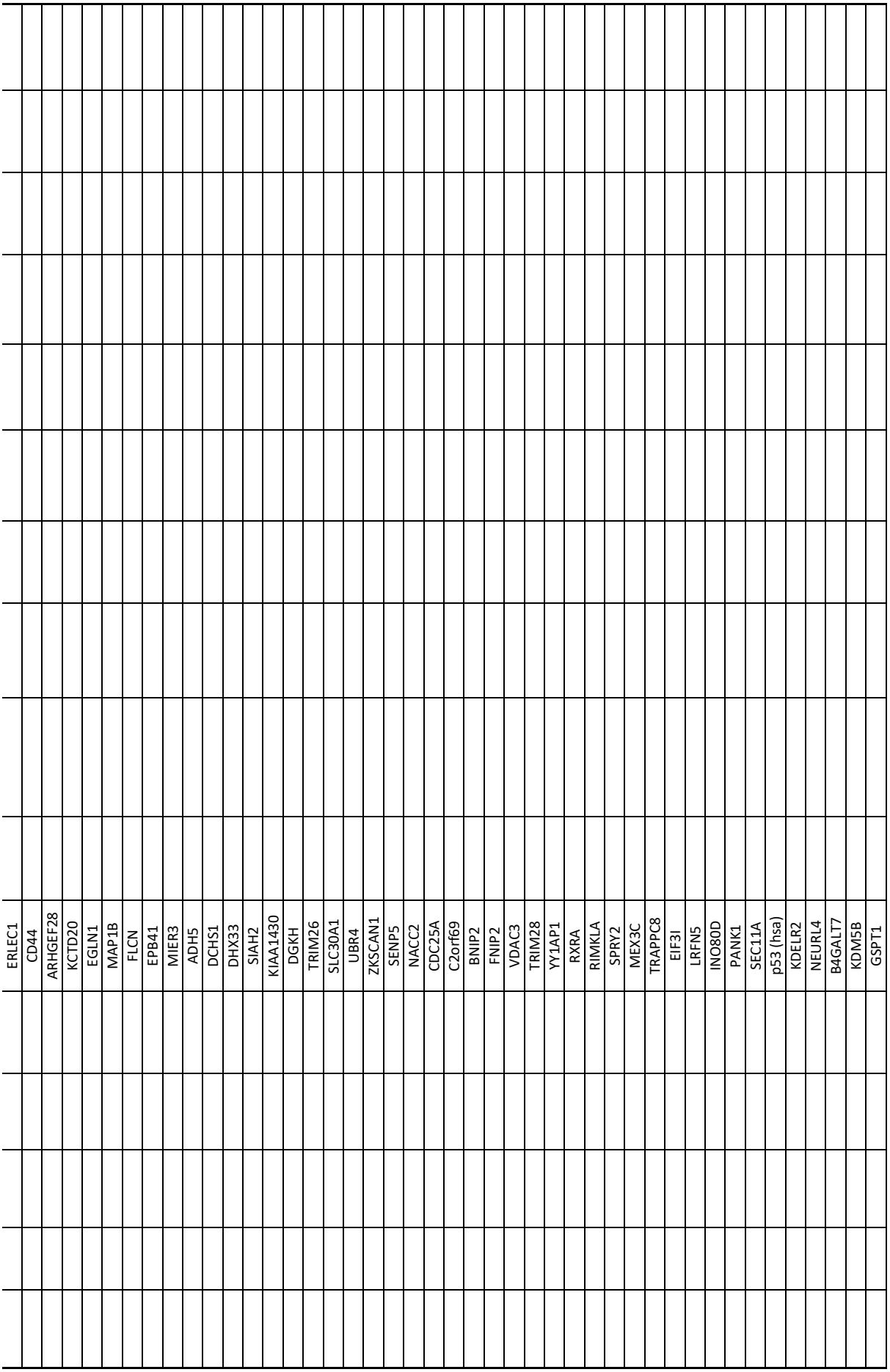

TUBB4B
DUSP16
MARCH6
DCTN6
HPS5
PALLD
SIX4
FUCA1
SPTBN1
CHD6
ATL3
LIN54
TRIM25
CD55
FOXO1
AP1B1
IRAK1
CAP1
RNF213
ADAMTS9
BBX
CAB39
PCYOX1
ATP13A3
PSMF1
SPATS2
RNF141
DYNC2H1
TRIB1
NSFL1C
QPCTL
MAPK1
MAP3K7
FNDC3A
BRI3
SDCBP
SLAIN2
PAX9
NUPL1
MTMR3
RNF168
RCC2
METTL7A
AMOTL1

					MAPK6										
					MED6										
					PTGFRN										
					CREBBP										
					FBXW11										
					GATAD2A										
					TMEM231										
					PPP2R1B										
					SESTD1										
					C9orf114										
					PRDX3										
					ABL2										
					CSF1										
					ZHX1										
					IL10										
					RAD21										
					HECTD4										
					HMMR										
					AP3M1										
					CIRBP										
					DNAJC27										
					KMT2C										
					HEATR2										
					APLN										
					TBPL1										
					GRB2										
					MGAT4A										
					FAM84B										
					CALCRL										
					BTG1										
					HOXA10										
					ITSN1										
					KIAA0895										
					USP22										
					EIF2B4										
					RNF26										
					C17orf70										
					DLL4										
					SLFN5										
					SMAD1										
					UBE4B										
					VBP1										
					RAB10										
					PCM1										

PARD6B
SERBP1
TMUB1
CPNE8
TRIM32
DNMBP
MTCL1
C1GALT1C1
LYSMD3
TAF3
KLF6
SRP14
NGDN
KIAA1279
MYBL2
ENAH
ARHGEF26
ITPK1
RAB20
ZBTB25
GPR107
DNAJC2
MAPKAPK2
GPR176
ZNF675
RNF139
KNTC1
SLC12A7
KPNB1
RHOB
C19orf24
BRWD1
CBX5
TTPAL
CDK10
JMY
MXRA7
TBL1XR1
40057
WDR37
SETX
GALNT5

					C14orf132										
					LETM1										
					GATA6										
					RNASE1										
					ATRN										
					IREB2										
					SPIDR										
					CCP110										
					GPS1										
					SECISBP2L										
					TGFBR2										
					SLC1A5										
					MEF2A										
					EIF1										
					C3orf52										
					BICC1										
					RAB14										
					ORC5										
					PON2										
					MTSS1L										
					TMEM60										
					ATP2C2										
					PLA2G15										
					SGK1										
					PRRC2A										
					ZBTB47										
					ARHGAP5										
					ATP5SL										
					RAPH1										
					KHDRBS1										
					PLEKHG1										
					JAK1										
					RPA2										
					RAB5B										
					CABIN1										
					MAF1										
					CLDN3										
					AKAP11										
					GLYCTK										
					SLC46A3										
					MDM2										
					TTC5										
					SLC18B1										
					PGK1										

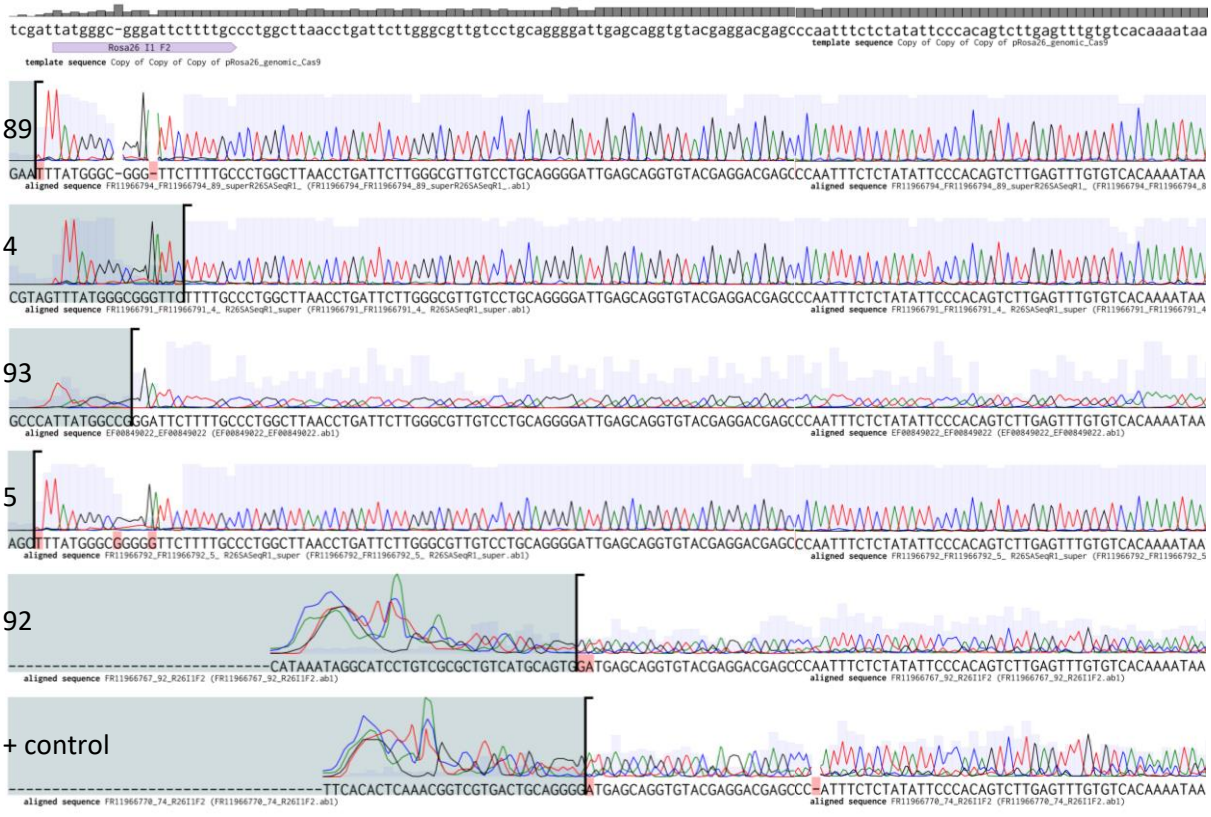
					ARMCX2										
					WNT9A										
					MTDH										
					TOMM6										
					SYTL5										
					NIN										
					ZNF669										
					TNRC6C										
					SLC35E2B										
					N4BP1										
					SCML1										
					ABL1										
					KIAA1429										
					EHMT1										
					SVIL										
					STK24										
					BMPR2										
					KIAA1551										
					PNISR										
					TMBIM6										
					ARHGAP12										
					TOR1AIP1										
					ZW10										
					EIF4G2										
					BCL2L11										
					ECE1										
					NR2F2										
					ZMYND8										
					MAP3K5										
					ZMYM3										
					KIF5B										
					MTF1										
					ILK										
					ZFP91-CNTF										
					LIN7C										
					SP6										
					SIN3A										
					CDC42BPB										
					PDGFRB										
					HOXA13										
					RPS6KB1										
					ARRB1										

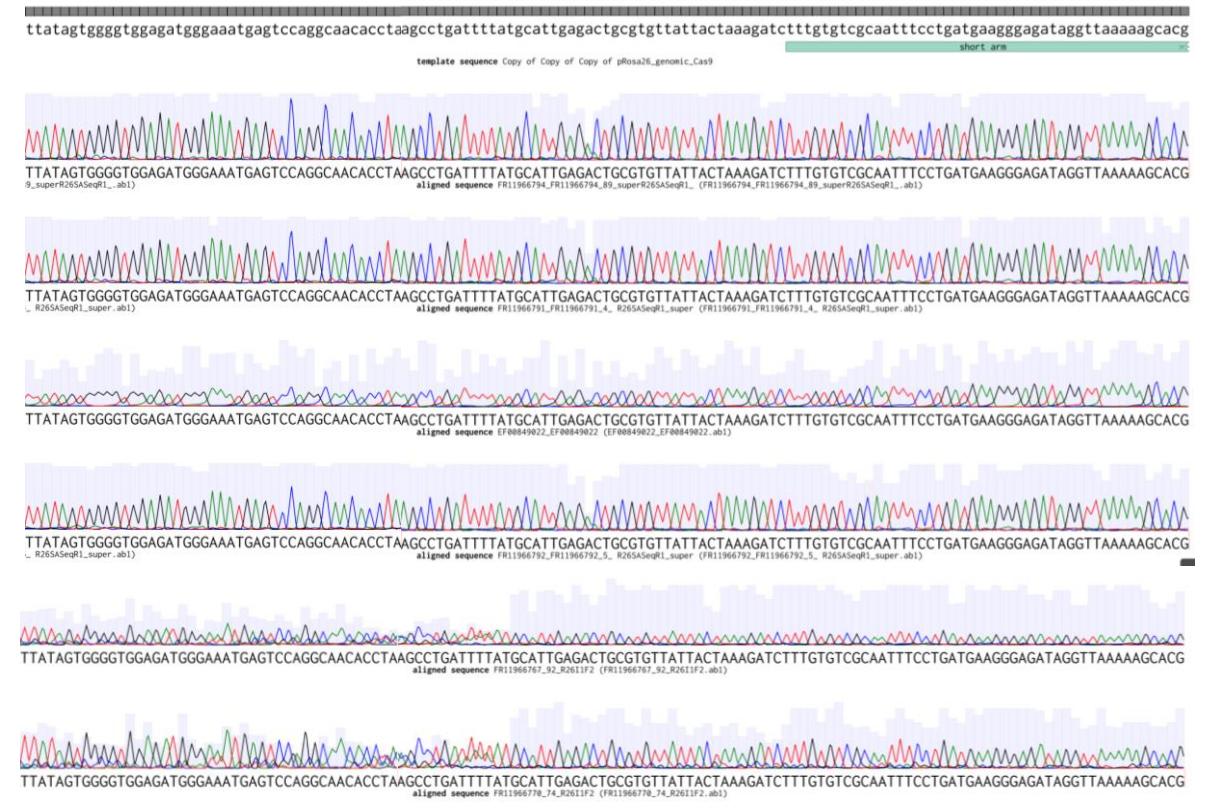

ERLEC1
CD44
ARHGEF28
KCTD20
EGLN1
MAP1B
FLCN
EPB41
MIER3
ADH5
DCHS1
DHX33
SIAH2
KIAA1430
DGKH
TRIM26
SLC30A1
UBR4
ZKSCAN1
SENP5
NACC2
CDC25A
C2orf69
BNIP2
FNIP2
VDAC3
TRIM28
YY1AP1
RXRA
RIMKLA
SPRY2
MEX3C
TRAPPC8
EIF3I
LRFN5
INO80D
PANK1
SEC11A
p53 (hsa)
KDELR2
NEURL4
B4GALT7
KDM5B
GSPT1

					NDEL1										
					COA7										
					ADCY6										
					SMCR8										
					UTP11L										
					TUSC2										
					ELF1										
					RMND5A										
					OR7C2										
					ZCCHC3										
					TXNIP										
					KMT2B										
					HIST1H2AH										
					HIST1H4E										
					PIGW										
					GGNBP2										
					PILRB										
					PIP4K2B										
					GTF2H5										
					ZNF8										
					NUDT3										
					C5orf22										
					HIST1H2BJ										
					SNX12										
					MARCKS										
					GPR75-ASB3										
					MYO19										
					VAMP3										
					RPS6KA2										
					ACACA										
					ZFP91										
					CTDSP2										
					CPNE1										
					PECAM1										
					BRD4										
					ZNF28										
					F11R										

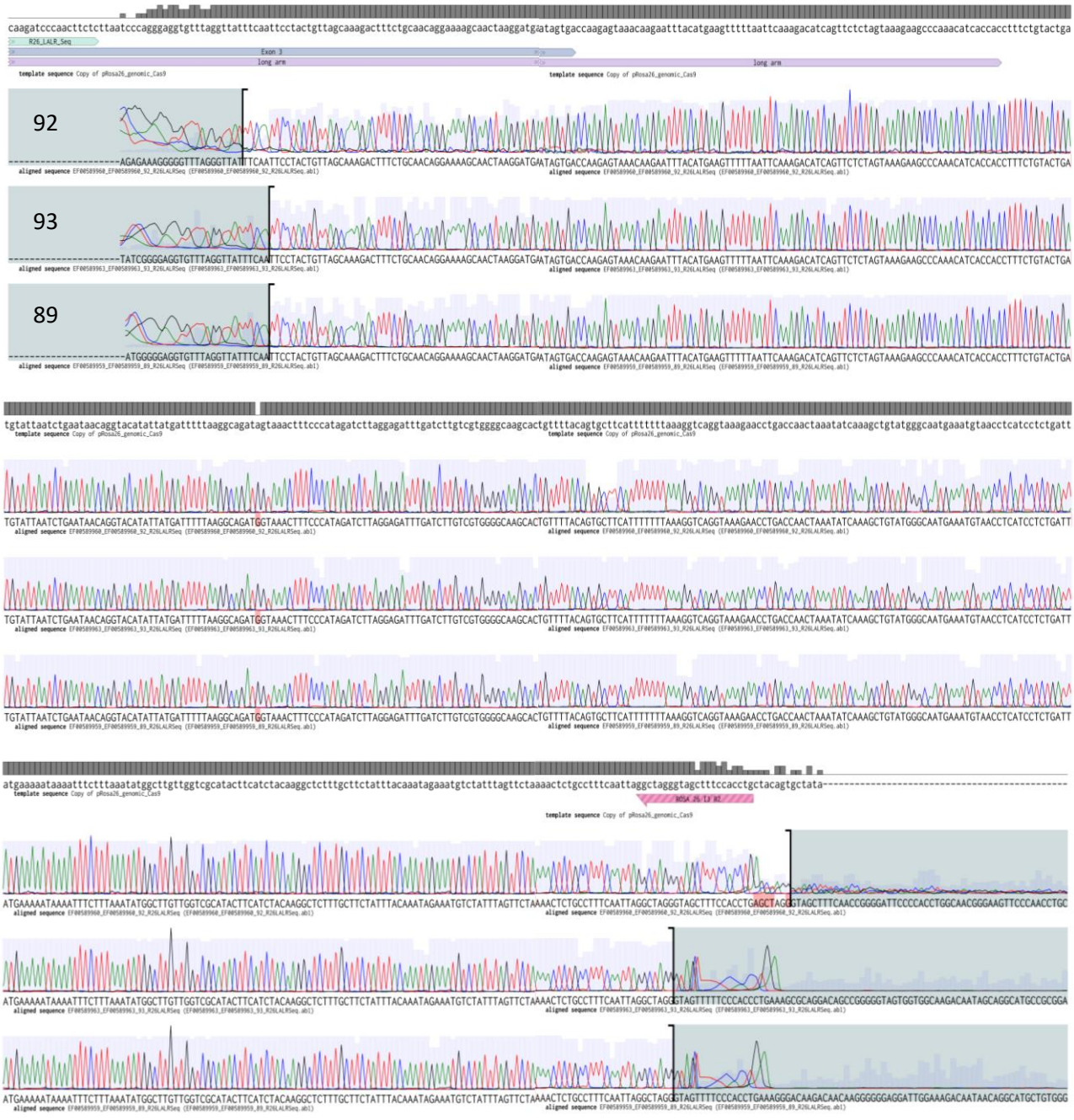
10.3 Sequencing analysis of the Cas9 targeted clones

Sequencing analysis of PCR amplification across the 5’ junction of the vector and the target site using the primer Rosa26 I1 F2 and Rosa26 Loc2R of the clones 89, 4, 93, 5, 92 and a positive control (from top to bottom) was performed using R26_SA_Seq_R1.

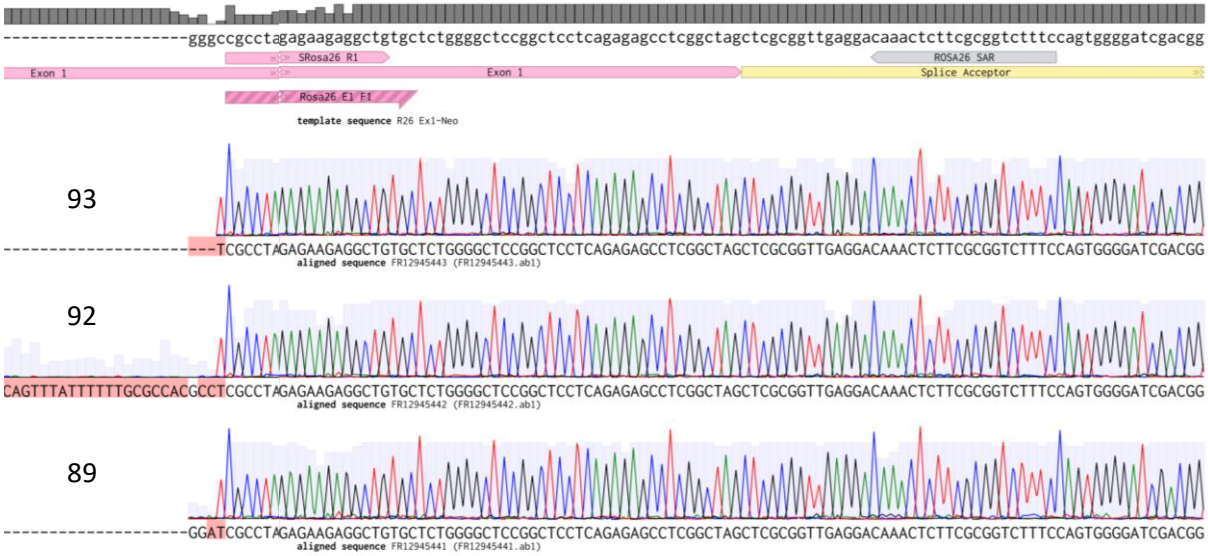


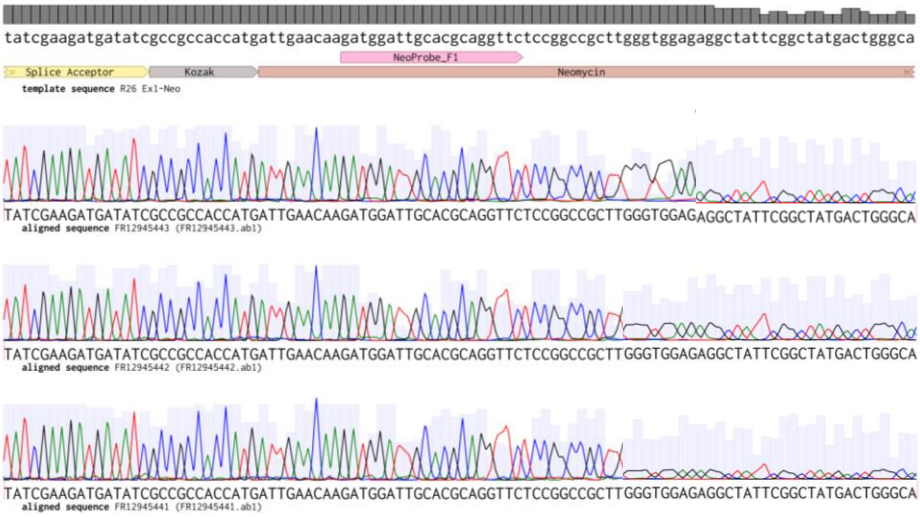


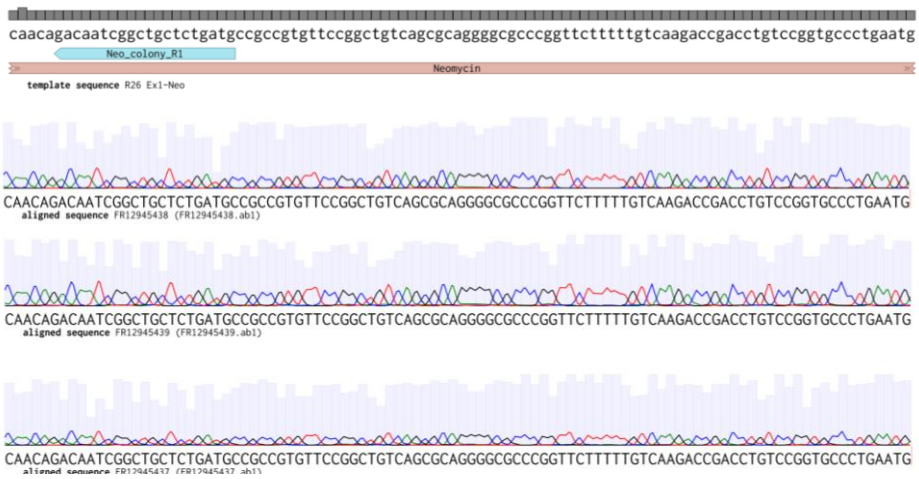
Sequencing analysis of PCR amplification across the 3'junction of the vector and target site 3' LR PCR using the primer Cas9_3'LR_for1 and Rosa26 I3 R2 of the clones 92, 93 and 89 was performed using Rosa26 I3 R2.



Sequencing analysis of the RT-PCR amplification from exon 1 of ROSA26 locus to the neomycin resistance gene using the primer Rosa26 E1 F1 and Rosa26 Loc2R and Rosa26 Loc3R of the clones 93, 92 and 89 was performed using the primer Neo_colony_R1 and Rosa26 Loc3R.



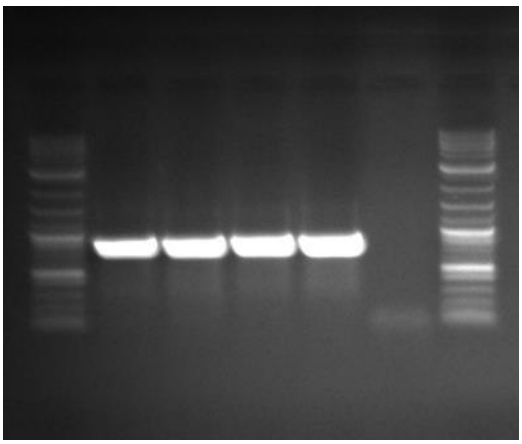


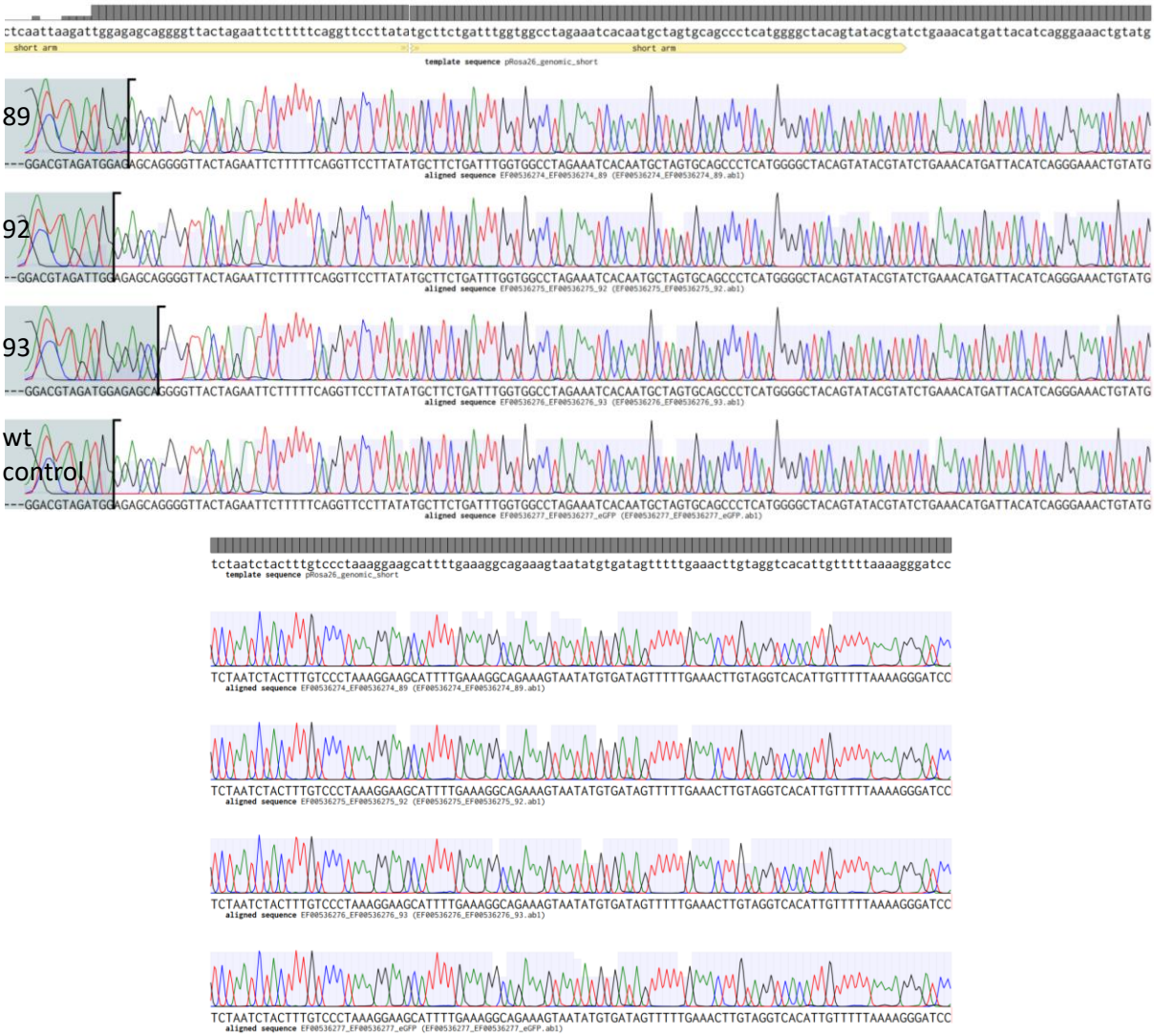


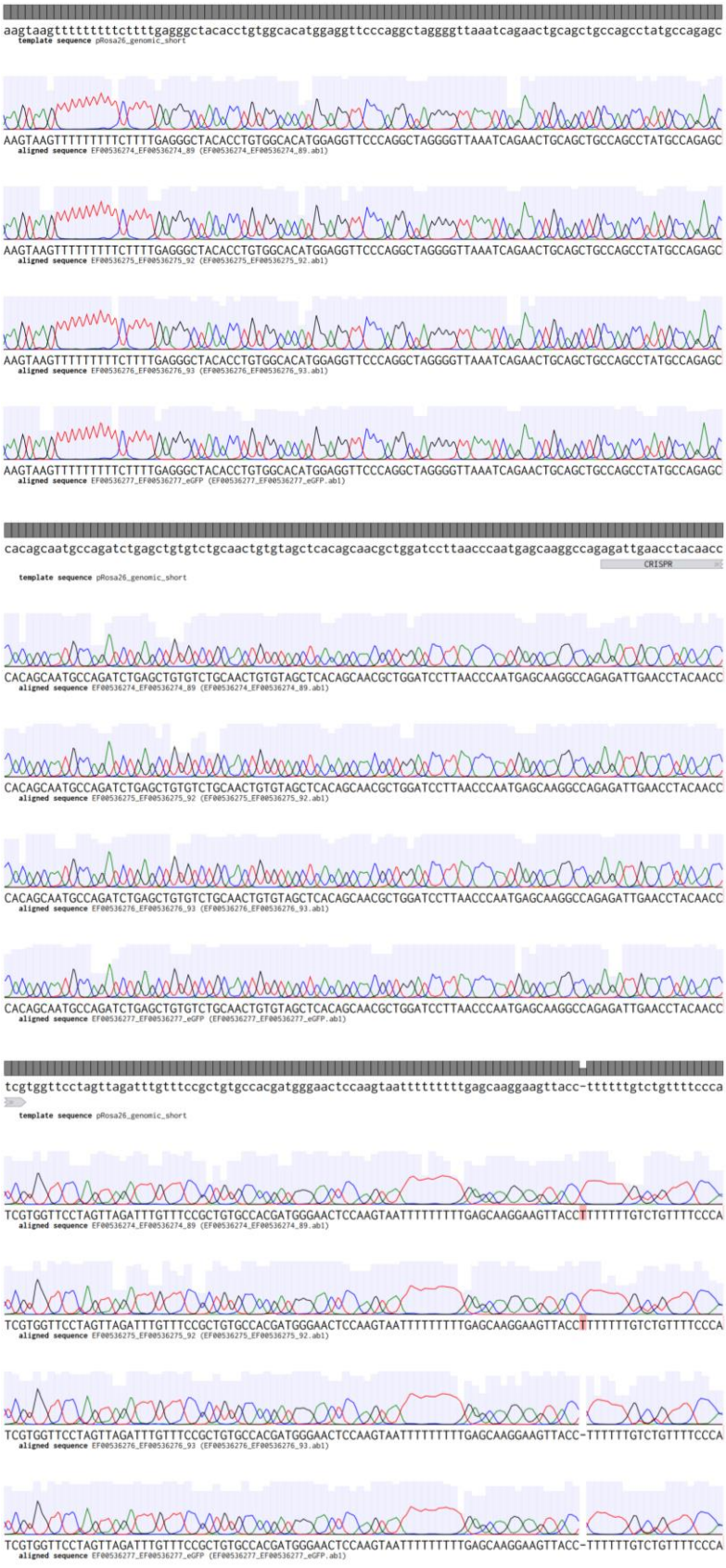

aactgcaggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcgggaagggactggct
Neomycin
template sequence R26 Ex1-Neo
AACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCT
aligned sequence FR12945438 (FR12945438.ab1)
AACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCT
aligned sequence FR12945439 (FR12945439.ab1)
AACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCT
aligned sequence FR12945437 (FR12945437.ab1)
gctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacg
Neomycin
template sequence R26 Ex1-Neo
GCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACG
aligned sequence FR12945438 (FR12945438.ab1)
GCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACG
aligned sequence FR12945439 (FR12945439.ab1)
GCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACG
aligned sequence FR12945437 (FR12945437.ab1)
cttgatccggctacctgcccattcgaccaccaagcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctgg
Neomycin
template sequence R26 Ex1-Neo
CTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGG
aligned sequence FR12945438 (FR12945438.ab1)
CTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGG
aligned sequence FR12945439 (FR12945439.ab1)
CTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGG
aligned sequence FR12945437 (FR12945437.ab1)

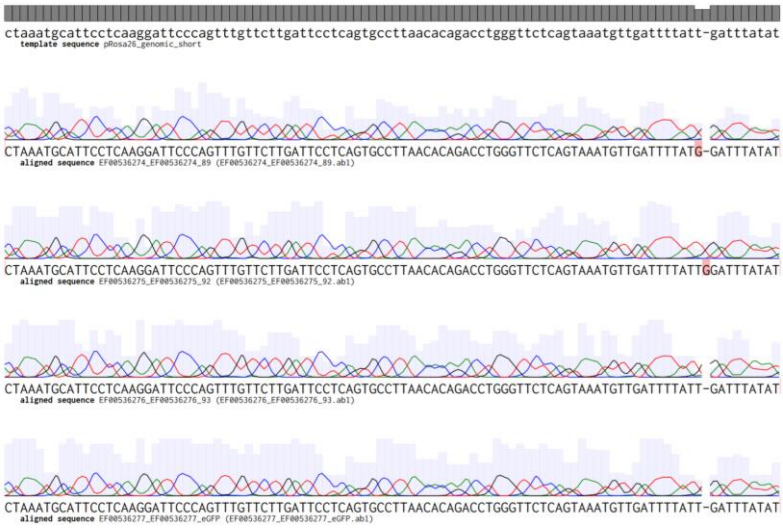

acgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacccatggcgatgcctg
Neomycin
template sequence R26 Ex1-Neo
ACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTG
aligned sequence FR12945438 (FR12945438.ab1)
ACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTG
aligned sequence FR12945439 (FR12945439.ab1)
ACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTG
aligned sequence FR12945437 (FR12945437.ab1)
cttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctacc
Neomycin
template sequence R26 Ex1-Neo
CTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACC
aligned sequence FR12945438 (FR12945438.ab1)
CTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACC
aligned sequence FR12945439 (FR12945439.ab1)
CTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACC
aligned sequence FR12945437 (FR12945437.ab1)

Gel electrophoresis and sequencing analysis of the endogenous ROSA26 locus of the clones 89, 92 and 93 using the primer Rosa26 I1 F2and Rosa26 I1 R3.






aagtaagttttttttttctttgagggctacacctgtggcacatggaggttcccaggctaggggttaaatcagaactgcagctgccagcctatgccagagc
template sequence pRosa26_genomic_short
AAGTAAGTTTTTTTTTTCTTTTGAGGGCTACACCTGTGGCACATGGAGGTTCCCAGGCTAGGGGTTAAATCAGAACTGCAGCTGCCAGCCTATGCCAGAGC
aligned sequence EF00536274_EF00536274_89 (EF00536274_EF00536274_89.ab1)
AAGTAAGTTTTTTTTTTCTTTTGAGGGCTACACCTGTGGCACATGGAGGTTCCCAGGCTAGGGGTTAAATCAGAACTGCAGCTGCCAGCCTATGCCAGAGC
aligned sequence EF00536275_EF00536275_92 (EF00536275_EF00536275_92.ab1)
AAGTAAGTTTTTTTTTTCTTTTGAGGGCTACACCTGTGGCACATGGAGGTTCCCAGGCTAGGGGTTAAATCAGAACTGCAGCTGCCAGCCTATGCCAGAGC
aligned sequence EF00536276_EF00536276_93 (EF00536276_EF00536276_93.ab1)
AAGTAAGTTTTTTTTTTCTTTTGAGGGCTACACCTGTGGCACATGGAGGTTCCCAGGCTAGGGGTTAAATCAGAACTGCAGCTGCCAGCCTATGCCAGAGC
aligned sequence EF00536277_EF00536277_eGFP (EF00536277_EF00536277_eGFP.ab1)
cacagcaatgccagatctgagctgtgtctgcaactgtgtagctcacagcaacgctggatccttaacccaatgagcaaggccagagattgaacctacaacc
CRISPR
template sequence pRosa26_genomic_short
CACAGCAATGCCAGATCTGAGCTGTGTCTGCAACTGTGTAGCTCACAGCAACGCTGGATCCTTAACCCAATGAGCAAGGCCAGAGATTGAACCTACAACC
aligned sequence EF00536274_EF00536274_89 (EF00536274_EF00536274_89.ab1)
CACAGCAATGCCAGATCTGAGCTGTGTCTGCAACTGTGTAGCTCACAGCAACGCTGGATCCTTAACCCAATGAGCAAGGCCAGAGATTGAACCTACAACC
aligned sequence EF00536275_EF00536275_92 (EF00536275_EF00536275_92.ab1)
CACAGCAATGCCAGATCTGAGCTGTGTCTGCAACTGTGTAGCTCACAGCAACGCTGGATCCTTAACCCAATGAGCAAGGCCAGAGATTGAACCTACAACC
aligned sequence EF00536276_EF00536276_93 (EF00536276_EF00536276_93.ab1)
CACAGCAATGCCAGATCTGAGCTGTGTCTGCAACTGTGTAGCTCACAGCAACGCTGGATCCTTAACCCAATGAGCAAGGCCAGAGATTGAACCTACAACC
aligned sequence EF00536277_EF00536277_eGFP (EF00536277_EF00536277_eGFP.ab1)
tcgtggttcctagttagatttgtttccgctgtgccacgatgggaactccaagtaatttttttttgagcaaggaagttacc-ttttttgtctgttttccca
template sequence pRosa26_genomic_short
TCGTGGTTCCTAGTTAGATTTGTTTCCGCTGTGCCACGATGGGAACTCCAAGTAATTTTTTTTTGAGCAAGGAAGTTACCTTTTTTTGTCTGTTTTCCCA
aligned sequence EF00536274_EF00536274_89 (EF00536274_EF00536274_89.ab1)
TCGTGGTTCCTAGTTAGATTTGTTTCCGCTGTGCCACGATGGGAACTCCAAGTAATTTTTTTTTGAGCAAGGAAGTTACCTTTTTTTGTCTGTTTTCCCA
aligned sequence EF00536275_EF00536275_92 (EF00536275_EF00536275_92.ab1)
TCGTGGTTCCTAGTTAGATTTGTTTCCGCTGTGCCACGATGGGAACTCCAAGTAATTTTTTTTTGAGCAAGGAAGTTACC-TTTTTTGTCTGTTTTCCCA
aligned sequence EF00536276_EF00536276_93 (EF00536276_EF00536276_93.ab1)
TCGTGGTTCCTAGTTAGATTTGTTTCCGCTGTGCCACGATGGGAACTCCAAGTAATTTTTTTTTGAGCAAGGAAGTTACC-TTTTTTGTCTGTTTTCCCA
aligned sequence EF00536277_EF00536277_eGFP (EF00536277_EF00536277_eGFP.ab1)


ctaaatgcattcctcaaggattcccagtttgttcttgattcctcagtgccttaacacagacctgggttctcagtaaatgttgattttatt-gatttatat
template sequence pRosa26_genomic_short
CTAAATGCATTCCTCAAGGATTCCCAGTTTGTTCTTGATTCCTCAGTGCCTTAACACAGACCTGGGTTCTCAGTAAATGTTGATTTTATG-GATTTATAT
aligned sequence EF00536274_EF00536274_89 (EF00536274_EF00536274_89.ab1)
CTAAATGCATTCCTCAAGGATTCCCAGTTTGTTCTTGATTCCTCAGTGCCTTAACACAGACCTGGGTTCTCAGTAAATGTTGATTTTATTGGATTTATAT
aligned sequence EF00536275_EF00536275_92 (EF00536275_EF00536275_92.ab1)
CTAAATGCATTCCTCAAGGATTCCCAGTTTGTTCTTGATTCCTCAGTGCCTTAACACAGACCTGGGTTCTCAGTAAATGTTGATTTTATT-GATTTATAT
aligned sequence EF00536276_EF00536276_93 (EF00536276_EF00536276_93.ab1)
CTAAATGCATTCCTCAAGGATTCCCAGTTTGTTCTTGATTCCTCAGTGCCTTAACACAGACCTGGGTTCTCAGTAAATGTTGATTTTATT-GATTTATAT
aligned sequence EF00536277_EF00536277_eGFP (EF00536277_EF00536277_eGFP.ab1)

11. Acknowledgments

An dieser Stelle möchte ich gerne die Möglichkeit ergreifen und mich herzlichst bei den Menschen bedanken, ohne deren Mithilfe, Unterstützung und Motivation der Beginn sowie die Fertigstellung dieser Promotionsarbeit unmöglich gewesen wäre.

Zu Beginn möchte ich meiner Doktormutter Frau Prof. Dr. Angelika Schnieke danken, die mir die Möglichkeit eröffnet hat, am Lehrstuhl für Biotechnologie der Nutztiere im spannenden Bereich der Onkologie und neuster Forschungserkenntnisse zu arbeiten und zu promovieren. Ich möchte mich auch herzlich für die stete Unterstützung durch wertvolle Hilfestellungen bei Problemen sowie allen meinen Vorhaben bedanken, welche ich zu keiner Zeit als selbstverständlich empfunden habe. Zudem gilt ein sehr großer Dank der Studienstiftung des deutschen Volkes, welche mich während drei meiner vier Jahre im Labor nicht nur finanziell, sondern auch ideell unterstützt hat. Mein Dank gebührt auch meinem Zweitprüfer Herrn Prof. Dr. Dieter Saur, welcher während der gesamten Zeit meiner Doktorarbeit ein sehr wichtiger und geschätzter Kollaborationspartner war. Außerdem möchte ich Herrn Prof. Dr. Harald Luksch meinen herzlichsten Dank aussprechen, welcher mich seit meinem ersten Semester begleitet, bereits als Zweitprüfer meiner Masterarbeit fungierte und sich ohne zu zögern bereit erklärt hat, den Vorsitz meiner Prüfungskommission zu übernehmen.

Ein ausgesprochenes Dankeschön gilt ebenso Herrn Dr. Alex Kind, der mir inhaltlich und sprachlich stets unter die Arme gegriffen hat und sowohl bei meinem Review-Artikel als auch bei der sprachlichen Gestaltung dieser Arbeit stets eine große Hilfe war.

Ganz besonders möchte ich meinen Betreuern Frau PD. Dr. Tatiana Flisikowska und Herrn PD Dr. Krzysztof Flisikowski danken. Sie haben mich zu jeder Zeit mit Ihrer wissenschaftlichen Expertise und Tatenkraft unterstützt. Ob an Wochenenden, Feiertagen, im Feierabend oder im anspruchsvollen Alltag haben sie sich immer Zeit genommen und waren stets mit einem Rat, einem offenen Ohr oder tröstenden Worten für mich da. Diese wohlwollende Unterstützung und gute Freundschaft ist keine Selbstverständlichkeit, weshalb sehr dankbar bin, solche Betreuer an meiner Seite gehabt zu haben.

Die Laborarbeit kann nicht immer alleine gestemmt werden, so gebührt meinen Kollaborationspartnern Frau Dr. Monika Stachowiak, Herrn Prof. Dr. Hubert Pausch, Herrn Prof. Dr. Hans-Rudolf Fries, Herrn Dr. Hongen Xu, Herrn Dr. Stefan Bauersachs und Frau Dr. Christine Wurmser mein aufrichtiger Dank, welche durch das Teilen Ihrer Expertise oder gemeinsamen Durchführung wichtiger Experimente und Analysen, einen signifikanten Beitrag an dieser Arbeit geleistet haben.

Auch im Laboralltag haben liebe Menschen Abläufe durch eine freundliche Atmosphäre und produktive Zusammenarbeit erleichtert. Daher auch ein großes Dankeschön an unsere TAs und Sekretärin Kristina Mosandl, Peggy Müller-Fliedner, Nina Simm, Sulith Christan, Marlene Stumbaum,

Toni Kuhnt, Robert Grötschel, Sandra Wantschner, Johanna Tebbing, Alex Carrapeiro und Barbara “Bobbylein” Bauer.

Außerdem möchte ich mich bei Joanna Madej, Agnieszka Bak und Guanglin Niu bedanken, die mich ganz uneigennützig bei Experimenten unterstützt haben. Auch die Studenten Mona Baumgartner, Sandra Romero und Johanna Steinhard haben im Rahmen ihrer Forschungspraktika einen Beitrag zu dieser Arbeit geleistet, wofür ich gerne Danke sagen möchte.

Weiterhin vielen Dank an die PostDocs Anja Saalfrank, Simone Kraner Schreiber, Konrad Fischer und Hicham Sid für ihren Rat und ihre Unterstützung. Meinen Mitdoktoranden Daniela Huber, Denise Nestle-Nguyen, Shun Li, Rahul Dutta, Benedikt Baumer, Erica Schulze, Andrea Schäffler, Romina Hellmich, Bernhard Klinger, Guanglin Niu, Liang Wei, Melanie Manyet, Alessandro Grodziecki, Daniela Kalla und Beate Rieblinger möchte ich für spannende Diskussionen, Abende am See und Kollegialität danken.

Das Arbeiten mit Schweinen benötigt geschulte Tierpfleger. So haben Viola und Steffen Loebnitz stets eine erstklassige Arbeit mit den Tieren geleistet, was meine eigene Arbeit nicht nur erleichtert, sondern auch bereichert hat. Auch die lieben Gespräche während der Plasmaprozessierung und Ihr Entgegenkommen im Rahmen meiner Experimente werden mir immer im Kopf bleiben. Dafür möchte ich mich ausdrücklich bedanken.

Die Freundschaften mit Tatiana, Erica, Denise, Rahul und Tinie, welche im Rahmen dieser Arbeit zustande gekommen sind und weit über diese hinausreichen, stellen ein großes Glück für mich dar.

Vielen Dank an meine Eltern Dagmar und Uwe Wander, ohne deren Unterstützung ich nie die Möglichkeiten gehabt hätte, all das zu erreichen, was ich heute geschafft habe.

Ganz besonders möchte ich mich für die starke emotionale Unterstützung meiner lieben Schwiegereltern Ilona und Mike Perleberg, meines lieben Opas Horst Kerzig, meines Bruders Sebastian und seiner kleinen Familie Steffanie und Benjamin Wander, sowie meiner wundervollen Freunde Denise Nestle-Nguyen, Melanie Gieseke, Daniela Dichtler, Maja Huber, Franziska und Fabian Drasdo, Maria Frystacki, Julia Hofmann, Franka Hirsch, Alina Huntgeburth, Tinie Wurmser, Simone Jung, Linda Tuchen, Sarah Bynevelt, Colleen Blake und Heidi Bisping-Arnold bedanken. Ihr ward stets für mich da und immer an meiner Seite, hierfür möchte ich mich von ganzem Herzen bei Euch bedanken.

Abschließend gibt es noch eine Person die mich über alle Höhen und Tiefen zusammengehalten hat, mir über jede Hürde geholfen, und jeden Moment zu etwas Besonderem gemacht hat, Kevin Perleberg. Mein Schatz ich liebe dich über alle Maßen und ich bin soooooo unsagbar Dankbar für deine Unterstützung. Ohne dich hätte ich das hier nie, nie, niemals geschafft. Your love is lifting me up!