

7-14-2015

# Recurrent Fusion Genes in Gastric Cancer: CLDN18-ARHGAP26 Induces Loss of Epithelial Integrity.

Fei Yao

Jaya P Kausalya


Yee Yen Sia

Audrey S M Teo

Wah Heng Lee

*See next page for additional authors*

Follow this and additional works at: <https://mouseion.jax.org/stfb2015>

 Part of the [Life Sciences Commons](#), and the [Medicine and Health Sciences Commons](#)

---

## Recommended Citation

Yao, Fei; Kausalya, Jaya P; Sia, Yee Yen; Teo, Audrey S M; Lee, Wah Heng; Ong, Alicia G M; Zhang, Zhenshui; Tan, Joanna H J; Li, Guoliang; Bertrand, Denis; Liu, Xingliang; Poh, Huay Mei; Guan, Peiyong; Zhu, Feng; Pathiraja, Thushangi Nadeera; Ariyaratne, Pramila N; Rao, Jaideepraj; Woo, Xing Yi; Cai, Shaojiang; Mulawadi, Fabianus H; Poh, Wan Ting; Veeravalli, Lavanya; Chan, Chee Seng; Lim, Seong Soo; Leong, See Ting; Neo, Say Chuan; Choi, Poh Sum D; Chew, Elaine G Y; Nagarajan, Niranjan; Jacques, Pierre-Étienne; So, Jimmy B Y; Ruan, Xiaoan; Yeoh, Khay Guan; Tan, Patrick; Sung, Wing-Kin; Hunziker, Walter; Ruan, Yijun; and Hillmer, Axel M, "Recurrent Fusion Genes in Gastric Cancer: CLDN18-ARHGAP26 Induces Loss of Epithelial Integrity." (2015). *Faculty Research 2015*. 121.

<https://mouseion.jax.org/stfb2015/121>

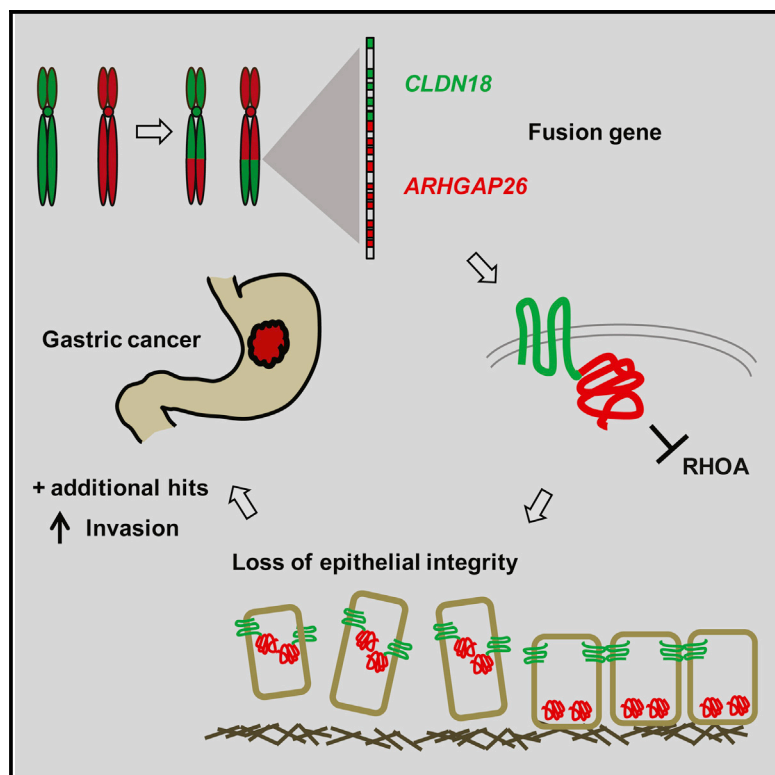
---

**Authors**

Fei Yao, Jaya P Kausalya, Yee Yen Sia, Audrey S M Teo, Wah Heng Lee, Alicia G M Ong, Zhenshui Zhang, Joanna H J Tan, Guoliang Li, Denis Bertrand, Xingliang Liu, Huay Mei Poh, Peiyong Guan, Feng Zhu, Thushangi Nadeera Pathiraja, Pramila N Ariyaratne, Jaideepraj Rao, Xing Yi Woo, Shaojiang Cai, Fabianus H Mulawadi, Wan Ting Poh, Lavanya Veeravalli, Chee Seng Chan, Seong Soo Lim, See Ting Leong, Say Chuan Neo, Poh Sum D Choi, Elaine G Y Chew, Niranjana Nagarajan, Pierre-Étienne Jacques, Jimmy B Y So, Xiaoan Ruan, Khay Guan Yeoh, Patrick Tan, Wing-Kin Sung, Walter Hunziker, Yijun Ruan, and Axel M Hillmer

## Recurrent Fusion Genes in Gastric Cancer: *CLDN18-ARHGAP26* Induces Loss of Epithelial Integrity

### Graphical Abstract



### Authors

Fei Yao, Jaya P. Kausalya, Yee Yen Sia, ..., Walter Hunziker, Yijun Ruan, Axel M. Hillmer

### Correspondence

hunziker@imcb.a-star.edu.sg (W.H.), yijun.ruan@jax.org (Y.R.), hillmer@gis.a-star.edu.sg (A.M.H.)

### In Brief

Yao et al. identify somatic structural variations (SVs) in gastric cancer (GC) that create five recurrent fusion genes. Comparisons between chromatin interaction and SV data suggest chromatin structure contributes to the emergence of SVs. Comprehensive functional analyses reveal that *CLDN18-ARHGAP26* impairs epithelial integrity and wound healing.

### Highlights

- Chromatin structure contributes to the shape of genome rearrangements in GC
- *CLDN18-ARHGAP26* is a recurrent fusion gene in 3% of Asian GCs
- *CLDN18-ARHGAP26* impairs epithelial barrier properties and wound healing
- *CLDN18-ARHGAP26* contributes to GC by loss of *CLDN18* and gain of *ARHGAP26* functions

### Accession Numbers

GSE26954  
GSE30833  
SRP035443



# Recurrent Fusion Genes in Gastric Cancer: *CLDN18-ARHGAP26* Induces Loss of Epithelial Integrity

Fei Yao,<sup>1,2,20</sup> Jaya P. Kausalya,<sup>3,20</sup> Yee Yen Sia,<sup>1,2</sup> Audrey S.M. Teo,<sup>1,2</sup> Wah Heng Lee,<sup>4</sup> Alicia G.M. Ong,<sup>3</sup> Zhenshui Zhang,<sup>5</sup> Joanna H.J. Tan,<sup>1</sup> Guoliang Li,<sup>6</sup> Denis Bertrand,<sup>4</sup> Xingliang Liu,<sup>1</sup> Huay Mei Poh,<sup>1</sup> Peiyong Guan,<sup>4,7</sup> Feng Zhu,<sup>2,8</sup> Thushangi Nadeera Pathiraja,<sup>1,2</sup> Pramila N. Ariyaratne,<sup>4</sup> Jaideep Raj Rao,<sup>9</sup> Xing Yi Woo,<sup>10</sup> Shaojiang Cai,<sup>4</sup> Fabianus H. Mulawadi,<sup>4</sup> Wan Ting Poh,<sup>10</sup> Lavanya Veeravalli,<sup>11</sup> Chee Seng Chan,<sup>11</sup> Seong Soo Lim,<sup>5</sup> See Ting Leong,<sup>12</sup> Say Chuan Neo,<sup>12</sup> Poh Sum D. Choi,<sup>12</sup> Elaine G.Y. Chew,<sup>1</sup> Niranjan Nagarajan,<sup>4</sup> Pierre-Étienne Jacques,<sup>13</sup> Jimmy B.Y. So,<sup>2,14,15</sup> Xiaoan Ruan,<sup>10,12</sup> Khay Guan Yeoh,<sup>2,8,15</sup> Patrick Tan,<sup>1,2,16,17</sup> Wing-Kin Sung,<sup>4,7</sup> Walter Hunziker,<sup>3,18,21,\*</sup> Yijun Ruan,<sup>19,21,\*</sup> and Axel M. Hillmer<sup>1,2,21,\*</sup>

<sup>1</sup>Cancer Therapeutics and Stratified Oncology, Genome Institute of Singapore, Singapore 138672, Singapore

<sup>2</sup>The Singapore Gastric Cancer Consortium, National University of Singapore, Singapore 119228, Singapore

<sup>3</sup>Epithelial Cell Biology Laboratory, Institute of Molecular and Cell Biology, Singapore 138673, Singapore

<sup>4</sup>Computational and Systems Biology, Genome Institute of Singapore, Singapore 138672, Singapore

<sup>5</sup>Human Genetics, Genome Institute of Singapore, Singapore 138672, Singapore

<sup>6</sup>National Key Laboratory of Crop Genetic Improvement, Center for Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, China

<sup>7</sup>School of Computing, National University of Singapore, Singapore 117417, Singapore

<sup>8</sup>Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore

<sup>9</sup>Department of General Surgery, Tan Tock Seng Hospital, Singapore 308433, Singapore

<sup>10</sup>Personal Genomic Solutions, Genome Institute of Singapore, Singapore 138672, Singapore

<sup>11</sup>Research Computing, Genome Institute of Singapore, Singapore 138672, Singapore

<sup>12</sup>Genome Technology and Biology, Genome Institute of Singapore, Singapore 138672, Singapore

<sup>13</sup>Département de Biologie, Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada

<sup>14</sup>Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore

<sup>15</sup>National University Health System, Singapore 119228, Singapore

<sup>16</sup>Duke-NUS Graduate Medical School, Singapore 169857, Singapore

<sup>17</sup>Cancer Science Institute of Singapore, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117599, Singapore

<sup>18</sup>Department of Physiology, National University of Singapore, Singapore 117597, Singapore

<sup>19</sup>The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032, USA

<sup>20</sup>Co-first author

<sup>21</sup>Co-senior author

\*Correspondence: [hunziker@imcb.a-star.edu.sg](mailto:hunziker@imcb.a-star.edu.sg) (W.H.), [yijun.ruan@jax.org](mailto:yijun.ruan@jax.org) (Y.R.), [hillmer@gis.a-star.edu.sg](mailto:hillmer@gis.a-star.edu.sg) (A.M.H.)  
<http://dx.doi.org/10.1016/j.celrep.2015.06.020>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

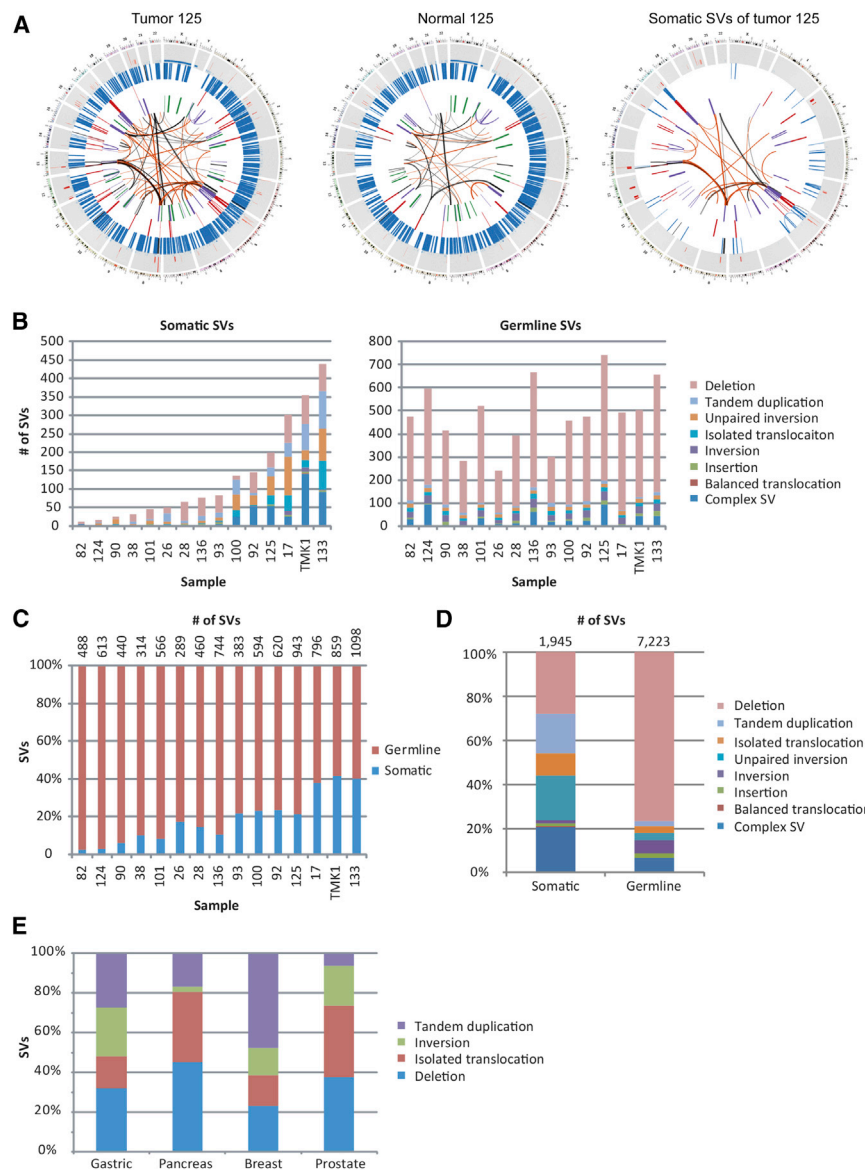
## SUMMARY

Genome rearrangements, a hallmark of cancer, can result in gene fusions with oncogenic properties. Using DNA paired-end-tag (DNA-PET) whole-genome sequencing, we analyzed 15 gastric cancers (GCs) from Southeast Asians. Rearrangements were enriched in open chromatin and shaped by chromatin structure. We identified seven rearrangement hot spots and 136 gene fusions. In three out of 100 GC cases, we found recurrent fusions between *CLDN18*, a tight junction gene, and *ARHGAP26*, a gene encoding a RHOA inhibitor. Epithelial cell lines expressing *CLDN18-ARHGAP26* displayed a dramatic loss of epithelial phenotype and long protrusions indicative of epithelial-mesenchymal transition (EMT). Fusion-positive cell lines showed impaired barrier properties, reduced cell-cell and cell-extracellular matrix adhe-

sion, retarded wound healing, and inhibition of RHOA. Gain of invasion was seen in cancer cell lines expressing the fusion. Thus, *CLDN18-ARHGAP26* mediates epithelial disintegration, possibly leading to stomach H<sup>+</sup> leakage, and the fusion might contribute to invasiveness once a cell is transformed.

## INTRODUCTION

Worldwide, nearly one million new gastric cancer (GC) cases were diagnosed and more than 700,000 deaths occurred in 2008 (Jemal et al., 2011). More than 70% new GC cases and deaths came from developing countries, with the highest incidence in Eastern Asia (Jemal et al., 2011). Most GCs are diagnosed at an advanced stage, which limits the current treatment strategies with the overall 5-year survival rate for distant or metastatic disease of ~3% (Janjigian and Kelsen, 2013). On the molecular level, GC is heterogeneous and only the amplified



**Figure 1. Characteristics of Somatic SVs Identified by DNA-PET in GC**

(A) SV filtering procedure for GC patient 125 is shown. SVs are plotted by Circos (Krzywinski et al., 2009) across the human genome arranged as a circle with the copy number in the outer ring (copy number > 2.8 is shown in red; copy number < 1.5 is shown in blue), followed by deletion (blue), tandem duplications (red), inversions (green), and unpaired inversions (purple), and in the inner ring, inter-chromosomal isolated translocations (orange). SVs identified in the blood of patient 125 (middle) were subtracted from SVs identified in gastric tumor of patient 125 (left), resulting in the somatically acquired SVs specific for the tumor (right). Circos plots of all 15 GCs can be found in Figure S1.

(B) Distribution of somatic and germline SVs of 15 GCs.

(C) Proportion of somatic SVs and germline SVs in 15 GCs. SV counts are shown on top.

(D) Composition of somatic SVs in GC compared with germline SVs. SV counts are shown on top.

(E) Comparison of somatic SV compositions of GC with reported somatic SVs for pancreatic cancer (Campbell et al., 2010), breast cancer (Stephens et al., 2009), and prostate cancer (Berger et al., 2011). SVs were reduced to four categories to allow comparison.

fragments (Hillmer et al., 2011; Nagarajan et al., 2012) to characterize the genomic structural rearrangements of 15 GCs and their impact on genes and gene fusions. We identified *CLDN18-ARHGAP26*, *CLEC16A-EMP2*, *SNX2-PRDM6*, *KMT2C (MLL3)-PRKAG2*, and *DUS2 (DUS2L)-PSKH1* as recurrent fusion genes with frequencies between 2% and 5% by an extended screen. Detailed functional evaluation suggests that *CLDN18-ARHGAP26* negatively affects cell-cell and cell-matrix interactions and epithelial barrier function,

*ERBB2* is a therapeutic target (Janjigian and Kelsen, 2013). Therefore, identifying additional molecular biomarkers in GC genomes will potentially result in early diagnosis and new treatment strategies. Whereas recent whole-genome- and exome-sequencing studies have identified recurrently mutated genes (Kakiuchi et al., 2014; Nagarajan et al., 2012; Wang et al., 2011, 2014; Zang et al., 2012), genome rearrangements in GC have not been studied in great detail. Genomic rearrangements can have dramatic impact on gene function by amplification, deletion, or gene disruption and can create fusion gene proteins with new functions or locations. There is still a large gap in our understanding of the functional role of the recurrent fusion genes and the underlying mechanisms of genomic rearrangements that create them.

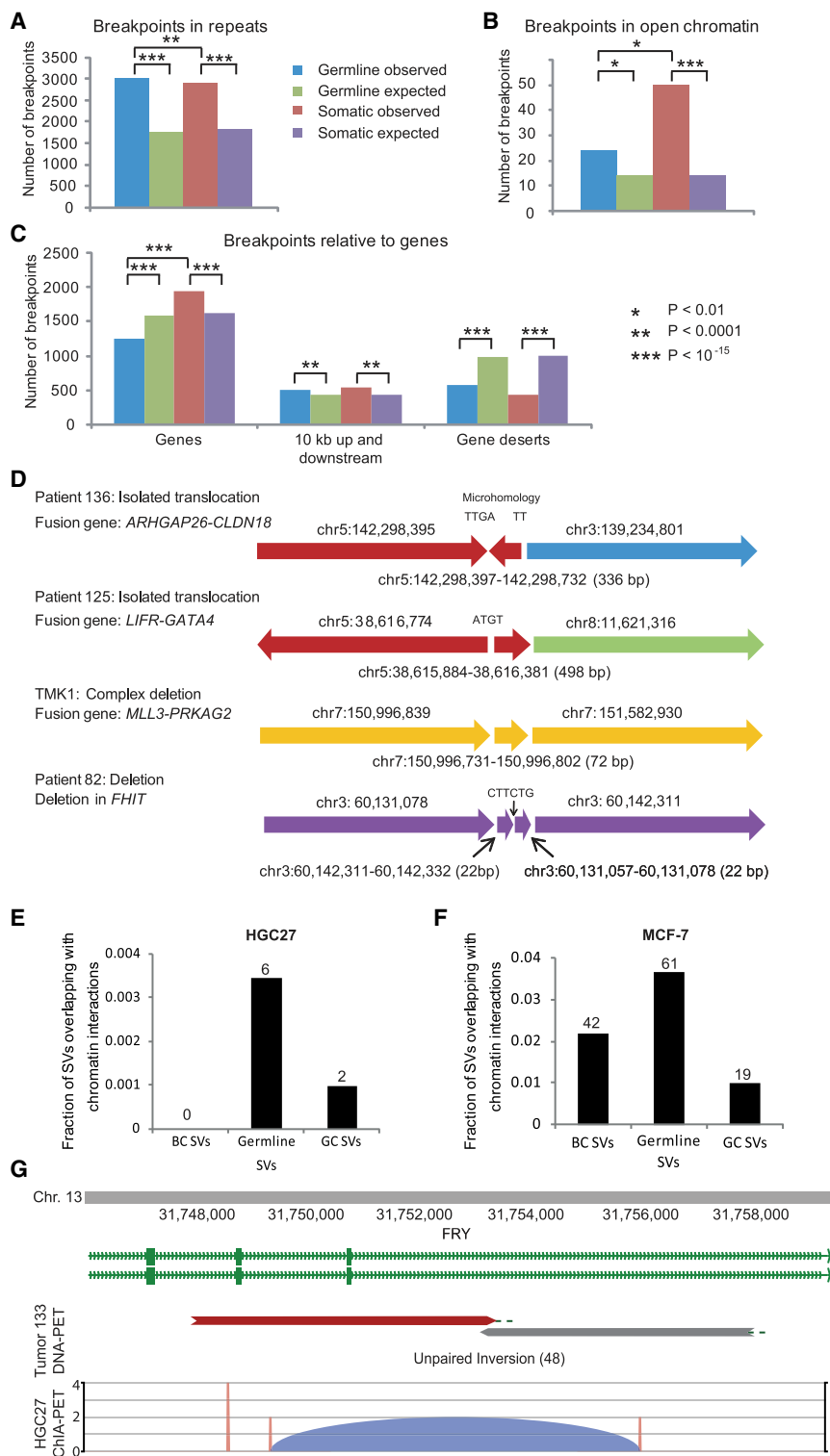
In the present study, we extended our earlier analyses by DNA paired-end-tag (DNA-PET) sequencing of 10-kb DNA

thereby potentially contributing to gastritis, a known risk factor for GC.

## RESULTS

### SVs in GC Identified by Whole-Genome DNA-PET Sequencing

We sequenced genomic DNA from 14 primary gastric tumors including ten paired normal samples (clinical data in Table S1) and GC cell line TMK1 by DNA-PET (Hillmer et al., 2011; Yao et al., 2012). With approximately 2-fold bp coverage and 200-fold physical coverage of the genome (Table S2), we identified 1,945 somatic structural variations (SVs) (Figures 1A–1C and S1; Table S3; Supplemental Experimental Procedures and Supplemental Results) with significant differences in SV distributions between germline and somatic SVs ( $p = 2.2 \times 10^{-16}$ ;  $\chi^2$



## Figure 2. Breakpoint Features of Somatic SVs Provide Mechanistic Insights

(A–C) Characterization of breakpoint locations of somatic SVs in GC. Coordinates of repeats and genes were downloaded from UCSC genome browser (Rhead et al., 2010), and open chromatin regions were compiled from Encyclopedia of DNA Elements (ENCODE) (Djebali et al., 2012).

(D) Gene-involving rearrangements can have insertions of small DNA fragments originating from one of the SV breakpoints. Arrows represent genomic fragments. Breakpoint coordinates are indicated, and micro-homologies are shown above breakpoint pairs.

(E and F) Overlap of somatic SVs identified by DNA-PET in breast cancer (BC) (n = 1,935; Hillmer et al., 2011) and GC (n = 1,945) and germline SVs in GC patients (n = 1,667) with long-range chromatin interactions bound to RNA polymerase II in gastric cancer cell line HGC27 (E; n = 7,623) and in breast cancer cell line MCF-7 (F; n = 87,253; Li et al., 2012). Absolute numbers are shown above bars. Fraction of SVs overlapping with ChIA-PET interactions is calculated relative to the total number of SVs of each data set (e.g., GC SVs). All SV/chromatin interaction overlaps are significantly higher than expected by chance (p < 0.001; permutation based).

(G) Example of an overlap of a somatic unpaired inversion in GC and a chromatin interaction. Coordinates of chromosome 13 are shown on top. UCSC gene track is displayed in green. The PET mapping coordinates of a somatic 5.3-kb unpaired inversion of GC tumor 133 are shown with the upstream mapping region in red and the downstream mapping region in gray. Number in brackets indicates number of non-redundant PET reads connecting the two regions (cluster size). (Bottom) Chromatin interaction is identified by ChIA-PET in cell line HGC27. Each pink peak represents one end of a mapped chromatin interaction, and the blue arch shows an interaction between two breakpoint regions.

tandem duplications than prostate cancer and more inversions than pancreatic cancer (Figure 1E), indicating that each cancer type bears its own rearrangement pattern.

## Characteristics of Somatic SVs in GC Provide Insight into Rearrangement Mechanisms

Both germline and somatic breakpoints were enriched in repeat regions (p < 10<sup>-5</sup>; Figure 2A) and open chromatin domains (p < 10<sup>-21</sup>;  $\chi^2$  test; Figure 2B) whereas only somatic breakpoints were enriched in genes (p < 10<sup>-15</sup>;  $\chi^2$  test) and germline breakpoints were depleted in genes (p < 10<sup>-15</sup>;  $\chi^2$  test; Figure 2C). This might reflect the negative selection for gene-disruptive rearrangements in germline and, in contrast, the pro-cancer potential for somatic rearrangements altering

tests; Figure 1D), suggesting different mutational or selective mechanisms. Compared to other cancer types that have been analyzed for SVs in detail (Berger et al., 2011; Campbell et al., 2010; Stephens et al., 2009), GC showed a higher proportion of

gene structures. These observations suggest that transcriptionally active parts of the genome are more prone for somatic rearrangements in GC (more details in the [Supplemental Results](#)).

We noticed in 2% of validated fusion points a characteristic pattern where the inserted sequence originated from a locus near the fusion point ([Figure 2D](#); [Supplemental Results](#)). Three of these cases created fusion genes (*ARHGAP26-CLDN18*, *LIFR-GATA4*, and *MLL3-PRKAG2*). Intriguingly, the same rearrangement characteristics have been described for a translocation with the same gene, *ARHGAP26*, in a patient with juvenile myelomonocytic leukemia creating a fusion with *KMT2A* ([Borkhardt et al., 2000](#)). The repeated observation of neighbor locus sequence insertions at rearrangement points suggests a specific mechanism, which might be transcription coupled.

We tested whether the rearrangement partner sites of somatic SVs tend to be in spatial proximity within the nucleus by searching for overlap between rearrangement points of SVs and chromatin interaction analysis by paired-end-tag (ChIA-PET) sequencing data ([Li et al., 2012](#)). We performed ChIA-PET sequencing for the GC cell line HGC27 and compared the derived chromatin interactions ( $n = 7,623$ ) with the SVs of the 15 GCs (1,667 germline and 1,945 somatic SVs). We found six germline and two somatic SV overlaps, more than expected by chance ( $p < 0.001$ ; permutation based; [Figures 2E](#) and [2G](#); [Supplemental Experimental Procedures](#)), indicating that chromatin interactions might contribute to the shape of germline and somatic GC SVs. We performed the same analysis with breast cancer cell line MCF-7 (87,198 chromatin interactions; [Li et al., 2012](#)), representing another epithelial cancer, and observed overlaps with 61 germline and 19 somatic GC SVs, respectively ( $p < 0.001$ ; permutation based). Chromatin interactions in HGC27 overlapped better with somatic SVs in gastric than in breast cancer, whereas chromatin interactions in MCF-7 overlapped better with somatic SVs in breast than GC ([Figures 2E](#) and [2F](#)), suggesting that tissue-specific chromatin interactions can contribute to the formation of somatic SVs.

### Recurrent Fusion Genes in GC

Using the somatic SVs of the 15 GCs, we were able to predict 136 fusion genes ([Tables S3](#) and [S6](#)), validated 97 of them by genomic PCR and Sanger sequencing, and confirmed the expression of 44 by RT-PCR in the respective tumors ([Table S6](#)). Fifteen expressed fusion genes were in frame. Because constitutively active oncogenic fusion genes are usually in-frame fusions, we focused on this category to screen an additional set of 85 GC tumor/normal pairs by RT-PCRs and found *SNX2-PRDM6* in one additional tumor, *CLDN18-ARHGAP26* and *DUS2L-PSKH1* in two additional tumors, *MLL3-PRKAG2* in three additional tumors, and *CLEC16A-EMP2* in four additional tumors, giving overall frequencies of 2%–5% ([Figures 3A](#) and [S2](#); [Table S6](#)). We performed statistical simulations to assess the significance of such rates of recurrence ([Experimental Procedures](#)) and found that they were not expected by chance ( $p = 0.00472$ ), with higher levels of significance for two rediscoveries ( $p = 9.98 \times 10^{-5}$ ) and three rediscoveries ( $p = 1.11 \times 10^{-5}$ ). This suggests that these fusion genes are not randomly created but most likely by targeted rearrangement mechanisms and/or that the resulting fusion genes provide selective advantages.

To explore whether the fusion genes provided selective advantages, we used a network fusion centrality analysis ([Wu et al., 2013](#)) to predict driver fusion genes. Among the 136 fusion genes of our study, 38 were classified as potential driver fusion genes, including *CLDN18-ARHGAP26*, *SNX2-PRDM6*, and *MLL3-PRKAG2* ([Table S7](#)).

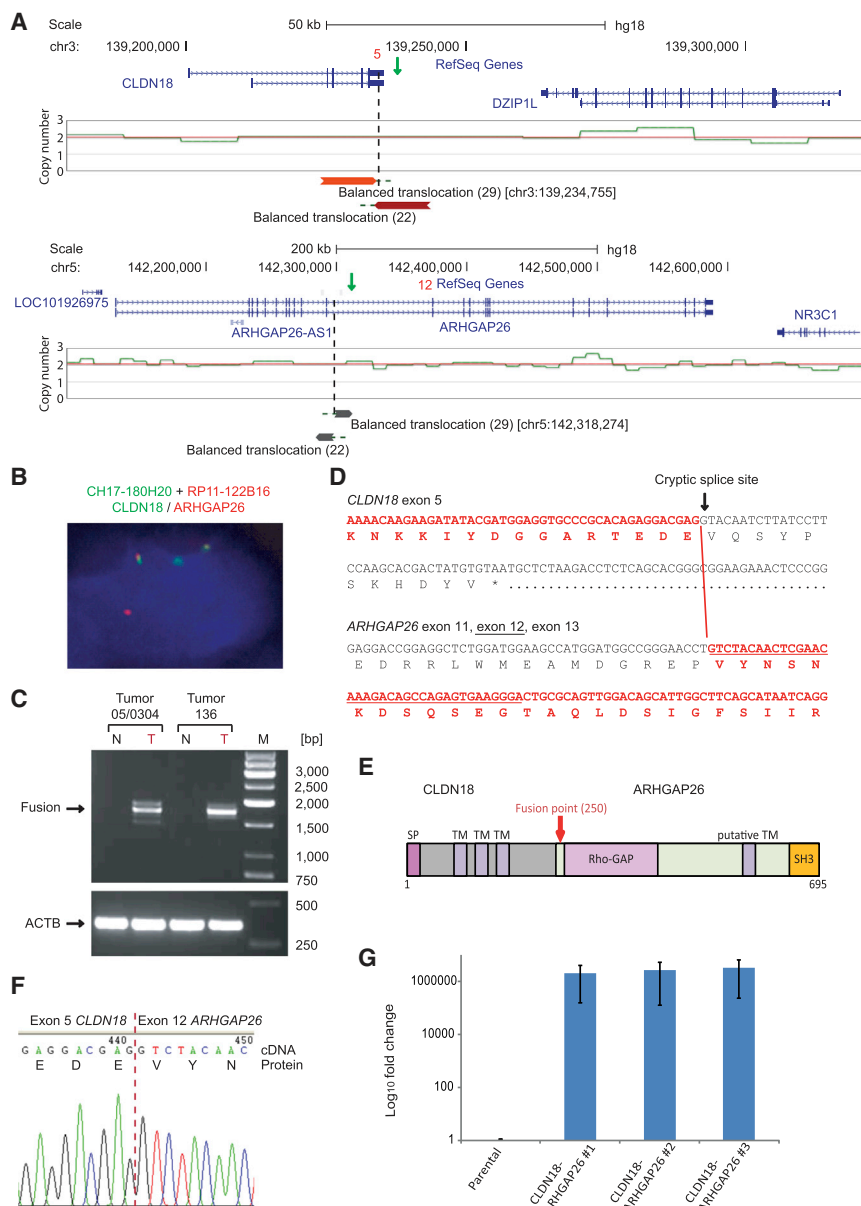
Further, we investigated the ratio of the discordant PETs (dPETs) that connect the fusion gene creating rearrangement points versus concordant paired-end tags (cPETs) that cover the fusion gene regions in the wild-type configuration. If the fusion gene rearrangements are early rearrangements, it is expected that such rearrangements can be observed from more cells and a relatively high dPET/cPET ratio from these rearrangements should be found. We compared the dPET/cPET ratio of the fusion gene rearrangements with all other somatic rearrangements. The fusion gene rearrangement point of *CLDN18-ARHGAP26*, *MLL3-PRKAG2*, and *SNX2-PRDM6* had higher dPET/cPET ratios than the median of the ratios for the somatic rearrangements of the samples in which they were discovered (tumor 136, TMK1, and tumor 125, respectively; data not shown), suggesting that the fusion genes are created by relatively early rearrangements.

We tested for the impact of the recurrent fusion genes on proliferation by knocking down *MLL3-PRKAG2* and *DUS2L-PSKH1* in TMK1 and establishing stable cell lines expressing *CLDN18-ARHGAP26*, *CLEC16A-EMP2*, and *SNX2-PRDM6*. We found evidence for proliferative supporting function for *MLL3-PRKAG2* and *CLEC16A-EMP2* ([Figure S3](#); [Supplemental Results](#); see below). For rearrangement hot spots and global impact of somatic SVs on genes, see [Tables S4](#) and [S5](#) and the [Supplemental Results](#).

### Phenotypic Changes by *CLDN18-ARHGAP26* Are Not a Result of Classical Epithelial-Mesenchymal Transition

Because *CLDN18* is an essential tight junction (TJ) component in the stomach with its deficiency causing paracellular  $H^+$  leakage ([Hayashi et al., 2012](#)) and *ARHGAP26* likely affects adhesion of cells to the extracellular matrix (ECM) through its regulation of *RHOA*, a gene recently shown to be mutated in GC ([Kakiuchi et al., 2014](#); [Wang et al., 2014](#)), we performed a deeper analysis of *CLDN18-ARHGAP26*. On the genomic level, we validated the *CLDN18-ARHGAP26* rearrangement in tumors 136 and 07K611 by fluorescence in situ hybridization (FISH) ([Figures 3B](#) and [S2K](#)) and PCR/Sanger sequencing ([Figure 3C](#)). Using custom capture sequencing, we verified the genomic fusion point on chromosome 3 in tumor 07K611 to be 2,342 bp downstream of *CLDN18* ([Figures 3A](#) and [S2A](#)). Array expression analysis of tumor 136 suggested that it belongs to the microsatellite stable/TP53<sup>-</sup> expression profile class defined recently ([Cristescu et al., 2015](#); data not shown).

In all three tumors with *CLDN18-ARHGAP26* fusions, the transcripts were joined by a cryptic splice site within the coding region of exon 5 of *CLDN18* and the regular splice site of exon 12 of *ARHGAP26* ([Figures 3D](#) and [3F](#)). *CLDN18-ARHGAP26* encodes a 75.6-kDa fusion protein containing all four transmembrane domains of *CLDN18* and the RhoGAP domain of *ARHGAP26* ([Figure 3E](#)) but lacking the C-terminal PDZ-binding motif of *CLDN18* that mediates interactions with TJ proteins (TJP1,



**Figure 3. Recurrent *CLDN18-ARHGAP26* In-Frame Fusions in GC**

(A) RefSeq gene track (top), copy number of tumor 136 by DNA-PET sequencing (middle), and PET mapping of a somatic balanced translocation with breakpoints in *CLDN18* and *ARHGAP26* in tumor 136 (bottom). Numbers of fused exons are shown in red. Mapping regions of DNA-PET clusters are shown by red and gray arrow heads with cluster size in brackets and dashed lines at Sanger sequencing validated breakpoint coordinates in squared brackets. Location of genomic breakpoints of tumor 07K611 (chr3:139,237,526 and chr5:142,309,897) are indicated by green arrows.

(B) Validation of genomic rearrangement by FISH of tumor 136.

(C) RT-PCRs of tumor/normal pairs of two gastric cancers with *CLDN18-ARHGAP26* fusions. RT-PCRs for  $\beta$ -actin (*ACTB*) serve as positive control. M, marker; N, normal gastric tissue; T, gastric tumor.

(D) Cryptic splice site in the coding region of exon 5 of *CLDN18* results in the extension of the open reading frame into *ARHGAP26*. Sequences of the fusion transcript are highlighted in red.

(E) Protein domain ideogram of *CLDN18-ARHGAP26*. TM, transmembrane.

(F) Sanger sequencing chromatogram of RT-PCR of *CLDN18-ARHGAP26* of tumor 136. Fusion point between *CLDN18* and *ARHGAP26* is indicated by red dashed line.

(G) qRT-PCR for the *CLDN18-ARHGAP26* fusion transcript in MCF10A non-transfected cells and stable cell lines with *CLDN18-ARHGAP26*-expressing vector.

Data are presented as mean  $\pm$  SD (G).

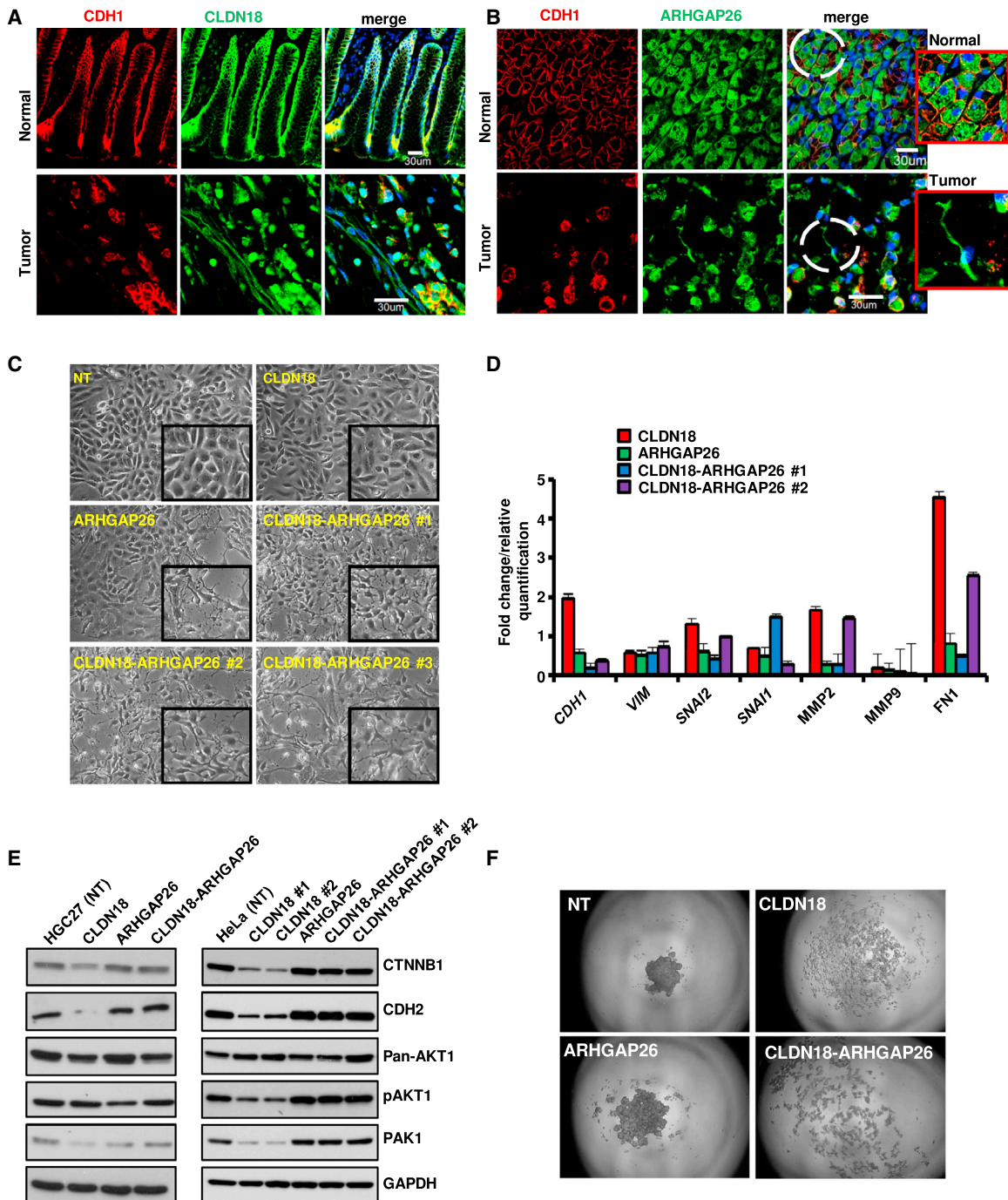
TJP2, and TJP3). *ARHGAP26* (*GRAF1*) binds to focal adhesion kinase (FAK or PTK2), which modulates cell growth, proliferation, survival, adhesion, and migration (Doherty and Lundmark, 2009). *ARHGAP26* can also negatively regulate the small GTP-binding protein RHOA, which is well known for its growth-promoting effect in RAS-mediated malignant transformation (Hildebrand et al., 1996; Qian et al., 2010).

*CLDN18* protein was observed in the plasma membrane of epithelial cells lining the gastric pit region and at the base of the gastric glands as previously reported in normal human stomach specimens (Sahin et al., 2008; Figure 4A). *ARHGAP26* was previously detected on pleiomorphic tubular and punctate membrane structures in HeLa cells (Lundmark et al., 2008). We observed *ARHGAP26* in normal stomach on vesicular structures

restricted to parietal cells that specifically express the  $H^+/K^+$  ATPase or “proton pump” (Figures 4B and S4A). Stomach tumor specimens of patient 136, with a cancer stage of 4 and one metastasis, expressing *CLDN18-ARHGAP26*, showed a diffused structure, characteristic for late-stage tumors, with some regions lacking E-cadherin (CDH1) staining (Figures 4A and 4B). *CLDN18-ARHGAP26* was present in both CDH1-positive and negative cells; with the CDH1-negative cells showing mesenchymal features such as elongated and spindled morphology (Figures 4A and 4B).

To understand the effect of the fusion protein to changes on epithelial integrity, we stably expressed *CLDN18*, *ARHGAP26*, or *CLDN18-ARHGAP26* in the non-transformed epithelial cell lines MCF10A (human breast; Figures 3G and S4B) and MDCK (canine kidney) and in the transformed cancer cell lines HGC27 (gastric) and HeLa (cervical cancer). Viewed by phase contrast, control and *CLDN18* in MCF10A and MDCK cell cultures, respectively, showed the characteristic epithelial morphology (Figures 4C and S4C). Whereas *ARHGAP26*-expressing MCF10A and MDCK cells were more spindle shaped and had short protrusions as previously reported (Taylor et al.,





**Figure 4. CLDN18-ARHGAP26-Fusion-Expressing Patient Specimen and MCF10A Cells Exhibit Loss of Epithelial Phenotype**

(A and B) *CLDN18* (A) and *ARHGAP26* (B) expression in normal and gastric tumor patient specimens. Human normal (top) and tumor (bottom) stomach sections were stained with DAPI and antibodies to CDH1 as well as *CLDN18* and *ARHGAP26*, respectively, for immunofluorescence analysis.

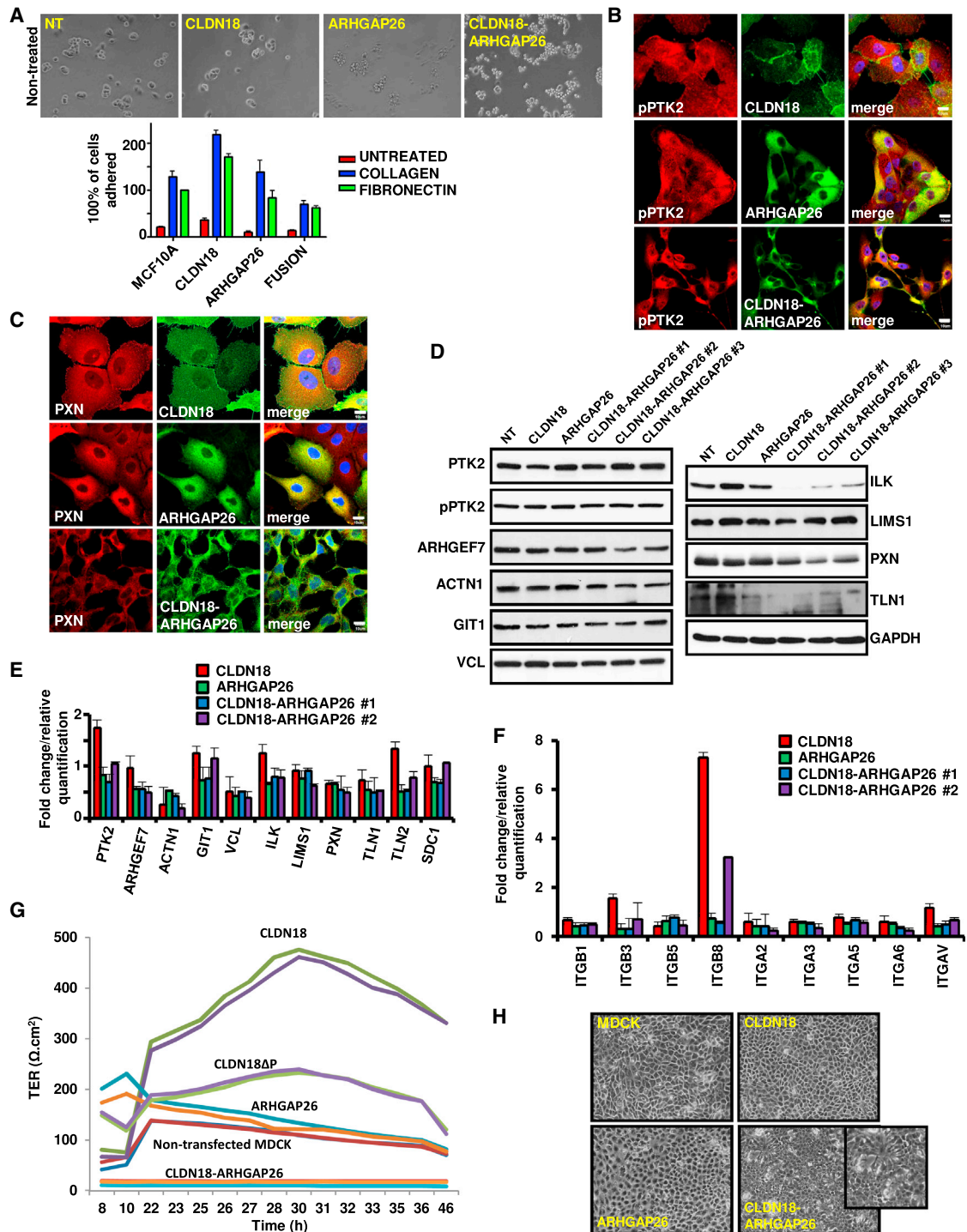
(C) *CLDN18-ARHGAP26*-fusion-expressing MCF10A cells display fusiform and protrusive morphology. Phase contrast images of non-transfected (NT) and stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* in MCF10A cells obtained at sub-confluent levels.

(D) qPCR of EMT markers in MCF10A cells stably expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26*, respectively.

(E) Western blot analysis of NT HGC27 and HeLa cells and stables expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* gene by immunoblotting for indicated proteins. GAPDH is used as loading control.

(F) Cell aggregation assay. MCF10A NT and stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* were plated as hanging drops, and phase contrast images were obtained the next day.

Data are presented as mean  $\pm$  SD (D). The scale bars in (A) and (B) represent 30  $\mu$ m.



**Figure 5. *CLDN18-ARHGAP26* Expression Results in Reduced Cell-ECM Adhesion**

(A) (Top) Cell-ECM adhesion assay. MCF10A stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* were seeded on untreated plates, and phase contrast images were obtained 2 hr after seeding. MCF10A NT cells were used as control. (Bottom) Quantification of cells that adhered to untreated collagen type I and fibronectin-treated surfaces is shown. The proportion of cells that adhered to NT MCF10A cells (100%) is shown.

(B) MCF10A stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* were fixed and immunostained with antibodies to activated PTK2 and HA or GFP.

(C) Absence of PXN in free edges in *CLDN18-ARHGAP26*-expressing MCF10A cells. MCF10A stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* were fixed and immunostained with antibodies to PXN and HA or GFP.

(legend continued on next page)

1999), *CLDN18-ARHGAP26* expression in both cell lines displayed a dramatic loss of epithelial phenotype and long protrusions, suggestive of epithelial-mesenchymal transition (EMT) (Figures 4C and S4C). Cell protrusion phenotype was also observed in transformed HGC27 and HeLa cells (Figures S4D and S4E).

To evaluate whether the phenotypic changes induced by *CLDN18-ARHGAP26* reflected a classical EMT, we investigated the expression of various EMT markers using qPCR. Whereas *CDH1* mRNA levels were reduced for *CLDN18-ARHGAP26*-expressing MCF10A cells, other EMT markers were unchanged or reduced, indicating that the morphological changes were not based on classical EMT (Figure 4D). In MDCK cells, *CDH1* levels were unchanged and mRNA of the master EMT regulators *SNAI1* and *SNAI2* (*SLUG*) were decreased (Figure S4F). MDCK-*CLDN18-ARHGAP26* showed a 5.2-fold increase in matrix metalloproteinase 2 (*MMP2*) mRNA levels relative to control MDCK cells (Figure S4F), suggesting changes in ECM adhesion induced by the fusion gene (Sahin et al., 2008).

Interestingly, expression of *CLDN18*, but not the fusion protein, downregulated N-cadherin (*CDH2*) and  $\beta$ -catenin (*CTNNB1*) expression in transformed HGC27 and HeLa cells (Figure 4E), suggesting that *CLDN18* can reverse the switch from an epithelial to a mesenchymal cadherin observed during EMT and suppress Wnt signaling, respectively. Wnt signaling is hyperactivated in many cancers, and *CDH2* expression activates AKT signaling (Tran et al., 2002), which is hyperactivated in many tumors. Although *CTNNB1* and *CDH2* were reduced in both cancer cell lines, pAKT1 protein levels, as well as those of the downstream effector p21-activated kinase (PAK1), were predominantly reduced in HeLa cells overexpressing *CLDN18* as compared to controls (Figure 4E). This might suggest a role for *CLDN18* as a tumor suppressor by dampening AKT1 and Wnt signaling.

### **CLDN18-ARHGAP26 Reduces Cell-ECM Adhesion**

ARHGAP26 likely affects adhesion of cells to the ECM through its interaction with PTK2 (FAK) and its regulation of RHOA, which in turn regulates focal adhesions (Taylor et al., 1999). Cell aggregation assays indicated poor aggregation for MCF10A-*CLDN18-ARHGAP26* cells (Figure 4F), suggesting that indeed the fusion gene causes epithelial changes that affect cell-cell interaction. Similar results were obtained with MDCK, HGC27, and HeLa cells (Figures S4H–S4J). Adhesion assays showed that control and MCF10A-*CLDN18* cells attached and spread on either untreated or ECM-coated surfaces. Not only did *ARHGAP26*-and, even more so, *CLDN18-ARHGAP26*-expressing cells attach

less efficiently to the surfaces (Figure 5A), but the cells that did attach were still rounded up 2 hr after seeding (Figure 5A), showing that the fusion gene potentiates the effect of *ARHGAP26* and strongly affects cell-ECM adhesive properties. Similar results were obtained in MDCK, HGC27, and HeLa cells (Figures S5A–S5F).

The SH3 domain of *ARHGAP26*, present in the fusion protein, binds to the focal adhesion molecules, PTK2 and PXN (paxillin) (Doherty et al., 2011). We therefore examined the effect of *CLDN18-ARHGAP26* expression on focal adhesion proteins. pPTK2 and PXN were detected at the free edge of MCF10A-*CLDN18*. These molecules were reduced in MCF10A-*ARHGAP26* but were absent from this location in MCF10A-*CLDN18-ARHGAP26* cells (Figures 5B and 5C). Western blot analysis for adhesion molecules associated with *ARHGAP26* or focal adhesion complex proteins showed reduced levels for integrin-linked kinase (ILK), Talin 1 (TLN1), and PXN in MCF10A-*ARHGAP26* and more pronounced so in MCF10A-*CLDN18-ARHGAP26* cells (Figure 5D). Significant decrease in levels of *ILK* and Talin 2 (*TLN2*) transcripts was observed in MCF10A-*ARHGAP26* and MCF10A-*CLDN18-ARHGAP26* cells by qPCR (Figure 5E). Changes in localization of focal adhesion molecules at free edges (pPTK2 and PXN) and protein expression patterns of focal adhesion components were also observed in MDCK, HGC27, and HeLa cells, indicating poor ECM adhesion of *CLDN18-ARHGAP26*-expressing cells (Figures S5G–S5I).

In addition to the cytoplasmic components of focal adhesions, we analyzed mRNA levels of integrin family members, which directly interact with the ECM components (Calderwood, 2004). Consistent with the poor attachment of MCF10A-*CLDN18-ARHGAP26* cells on collagen-coated surfaces (Figure 5A), these cells expressed reduced levels of integrin  $\beta$ 3 (*ITGB3*), integrin  $\beta$ 8 (*ITGB8*), and integrin  $\alpha$ V (*ITGAV*) (Figure 5F). A decrease in transcript levels of integrin subunits, in particular integrin  $\beta$ 1 (*ITGB1*), integrin  $\alpha$ 3 (*ITGA3*), and integrin  $\alpha$ 5 (*ITGA5*), was observed in MDCK-*CLDN18-ARHGAP26* cells (Figure S5J). In summary, overexpression of *ARHGAP26* and even more pronounced of the fusion gene disrupt ECM adhesion.

### **The Epithelial Barrier Promoted by CLDN18 Is Compromised by CLDN18-ARHGAP26**

Claudins are critical components of the paracellular epithelial barrier, including the protection of the gastric tissue from the acidic milieu in the lumen (Davenport, 1972a, 1972b, 1975). Alterations of this barrier function might cause chronic inflammation (Jovov et al., 2007; Sanders et al., 1985), a risk factor for the development

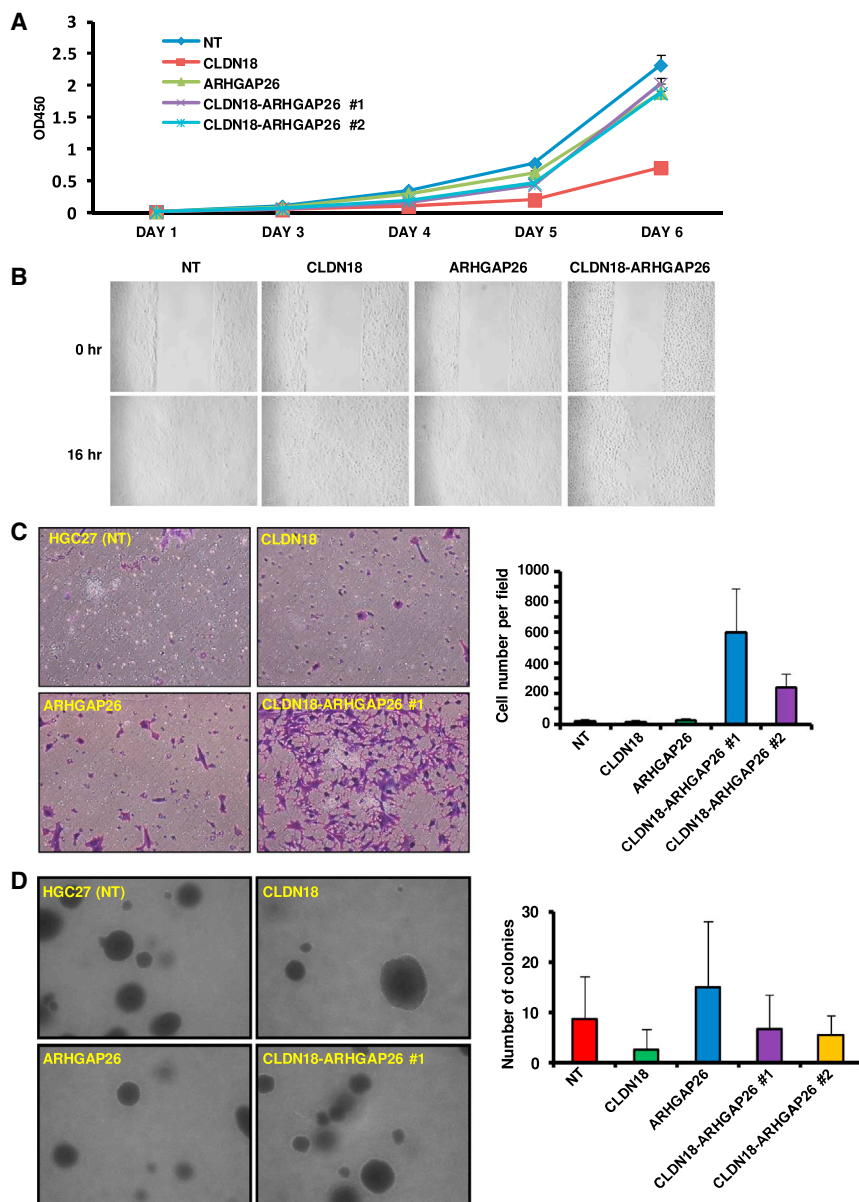
(D) Western blot analysis of focal adhesion molecule levels in MCF10A NT and stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26*. GAPDH was used as loading control.

(E) Reduced levels of focal adhesion molecules in *CLDN18-ARHGAP26*-expressing MCF10A. qPCR analysis of MCF10A stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* was performed for focal adhesion molecules. Fold changes were calculated relative to MCF10A NT cells.

(F) Reduction in integrin subunit levels in *CLDN18-ARHGAP26*-expressing MCF10A. qPCR analysis of MCF10A-*CLDN18*, -*ARHGAP26*, and -*CLDN18-ARHGAP26* stables was performed for integrin subunits. Fold changes were calculated relative to MCF10A NT cells.

(G) MDCK stable lines expressing *CLDN18*, *CLDN18* with inactivated C-terminal PDZ-binding motif (*CLDN18 $\Delta$ IP*), *ARHGAP26*, *CLDN18-ARHGAP26*, and NT MDCK cells were seeded on Transwell inserts, and TER values were measured over a period of 48 hr. Empty Transwell inserts were used as negative control. Experiments were performed in duplicate.

(H) Phase contrast images of NT MDCK and stables expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* at confluent levels. Data are presented as mean  $\pm$  SD (A, E, and F). Scale bars in (B) and (C) represent 10  $\mu$ m.



**Figure 6. CLDN18-ARHGAP26 Has a Cell-Context-Specific Impact on Proliferation, Wound Closure, and Invasion**

(A) Delayed cell proliferation rates in *CLDN18-ARHGAP26*-fusion-expressing MCF10A cells. MCF10A stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* were seeded at 800 cells in quadruplicate in 24-well plates. MCF10A NT cells were used as control.

(B) Wound healing assay of MCF10A stable lines expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26*. Phase contrast images were obtained at the start of the experiments and at intervals.

(C) HGC27 cells stably expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* fusion gene were seeded on Matrigel invasion chamber. NT HGC27 cells were used as control. Cells were fixed, washed, and stained with crystal violet to obtain phase contrast images (left) and to quantitate (right) the number of cells that invaded the Matrigel.

(D) HGC27 cells stably expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* were seeded on soft agar, incubated for 1 month, imaged (left), and counted (right). Parental lines stably transfected with vector were used as control. Data are presented as mean  $\pm$  SD (A, C, and D).

cytoskeleton, were not altered in MDCK cells expressing the fusion protein (Figure S5K), the expression of several other TJ components was upregulated in MDCK-*CLDN18-ARHGAP26*, possibly reflecting a compensatory mechanism (Figure S5L).

### CLDN18-ARHGAP26 Exerts Cell-Context-Specific Effects on Cell Proliferation, Invasion, and Migration

*CLDN18-ARHGAP26* reduced cell proliferation in transfected MDCK, HGC27, and HeLa cells compared to controls (Figure S6A). In MCF10A, *CLDN18-ARHGAP26* cells had lower cell proliferation compared to non-transfected and ARH-

GAP26 but a higher proliferation rate than MCF10A-*CLDN18* cells (Figure 6A), suggesting cell-context-specific differences. Interestingly, *CLDN18-ARHGAP26* expression in the four analyzed cell lines delayed wound closure (Figures 6B and S6B). Expression of *CLDN18-ARHGAP26* in MCF10A and MDCK cells had no effect on invasion and anchorage-independent growth (Figures S6C and S6D), which are features of cancer progression and metastasis. The fusion therefore might not be considered a classical oncogenic driver and, as expected, we found independent driver mutations in the *CLDN18-ARHGAP26*-positive tumors (Supplemental Results). We tested whether invasion and anchorage-independent growth were altered in cancer cell lines HGC27 and HeLa. Two independent HGC27 and HeLa cell lines stably expressing *CLDN18-ARHGAP26* showed 300- to 600-fold and 3- to 4-fold increase in cell invasion (Figures 6C and S6F). However,

of GC (Vannella et al., 2012), prompting us to explore the role of *CLDN18* and the fusion protein in barrier formation. Overexpression of *CLDN18*, which is not endogenously expressed in MDCK cells, resulted in a dramatic increase in the transepithelial electrical resistance (TER) of MDCK-*CLDN18* monolayers. Whereas *ARHGAP26* had no significant effect on the TER, *CLDN18-ARHGAP26* completely abolished the TER (Figure 5G). This effect did not simply reflect the lack of the C-terminal PDZ-binding motif, because a *CLDN18* construct where this C-terminal PDZ-binding motif was inactivated (*CLDN18 $\Delta$ P*) still increased the baseline TER of MDCK cells. Phase contrast images of confluent *CLDN18-ARHGAP26*-fusion-expressing MDCK cells showed that these cells failed to form tight monolayers, explaining the loss of TER (Figure 5H). Whereas expression levels and subcellular localization of TJP1, a scaffold protein that directly links claudins to the actin

CLDN18-ARHGAP26 in these transformed cells appeared to be less efficient in soft agar growth assays compared to controls (Figures 6D and S6G). These findings highlight different effects of the fusion protein on proliferation, invasion, and anchorage-independent growth in non-transformed and transformed cells and might suggest a role of the fusion protein driving late cancer events such as invasion.

### Both ARHGAP26 and CLDN18-ARHGAP26 Inhibit RHOA and Stress Fiber Formation

RHOA regulates many actin events like actin polymerization, contraction, and stress fiber formation upon growth factor receptor or integrin binding to their respective ligands (Nobes and Hall, 1995). ARHGAP26 stimulates, via its GAP domain, the GTPase activities of CDC42 and RHOA, resulting in their inactivation (Taylor et al., 1999). Given that the CLDN18-ARHGAP26 fusion protein retains the GAP domain of ARHGAP26, it may still be able to inactivate RHOA. To test this, we analyzed the effect of *CLDN18-ARHGAP26* expression on stress fiber formation and the presence and subcellular localization of active RHOA. In MCF10A, MDCK, HGC27, and HeLa cells, stable overexpression of *ARHGAP26* or *CLDN18-ARHGAP26* induced cytoskeletal changes, notably a reduction in stress fibers indicative of RHOA inactivation (Figures 7A and S7A–S7C). Labeling of stable MCF10A cell lines with an antibody that specifically recognizes activated RHOA showed reduced labeling in ARHGAP26 and even more so in CLDN18-ARHGAP26-fusion-protein-expressing cells (Figure 7B), whereas total RHOA levels remained unchanged in ARHGAP26-expressing cells and were reduced in fusion-expressing cells (Figures 7C and S7D). G-LISA assay measuring levels of active RHOA further confirmed these results (Figure 7D). These findings were also observed in MDCK cells (Figures S7E–S7G) and suggest that the GAP domain in the CLDN18-ARHGAP26 fusion protein retains its inhibitory activity on RHOA and lower total RHOA levels contribute to reduced RHOA function.

### CLDN18-ARHGAP26 Fusion Protein Suppresses Clathrin-Independent Endocytosis

ARHGAP26 has been shown to regulate, through its BAR and PH domains, clathrin-independent endocytosis. Changes in endocytosis can affect cell surface residence time and/or degradation of cell-ECM and cell-cell adhesion proteins as well as receptor tyrosine kinases, thereby altering cell adhesion, migration, and receptor tyrosine kinase signaling, which can drive carcinogenesis (Mellman and Yarden, 2013). MCF10A, MDCK, HGC27, and HeLa cells expressing the CLDN18-ARHGAP26 fusion protein showed a significant reduction of endocytosis (Figures 7E and S7H–S7J; Supplemental Results), consistent with the absence of the BAR and PH domains, which are essential for the role of ARHGAP26 in this process (Doherty et al., 2011), from the fusion protein. A summary of the effects of CLDN18-ARHGAP26 on the four analyzed cell lines is provided in Figure 7F.

## DISCUSSION

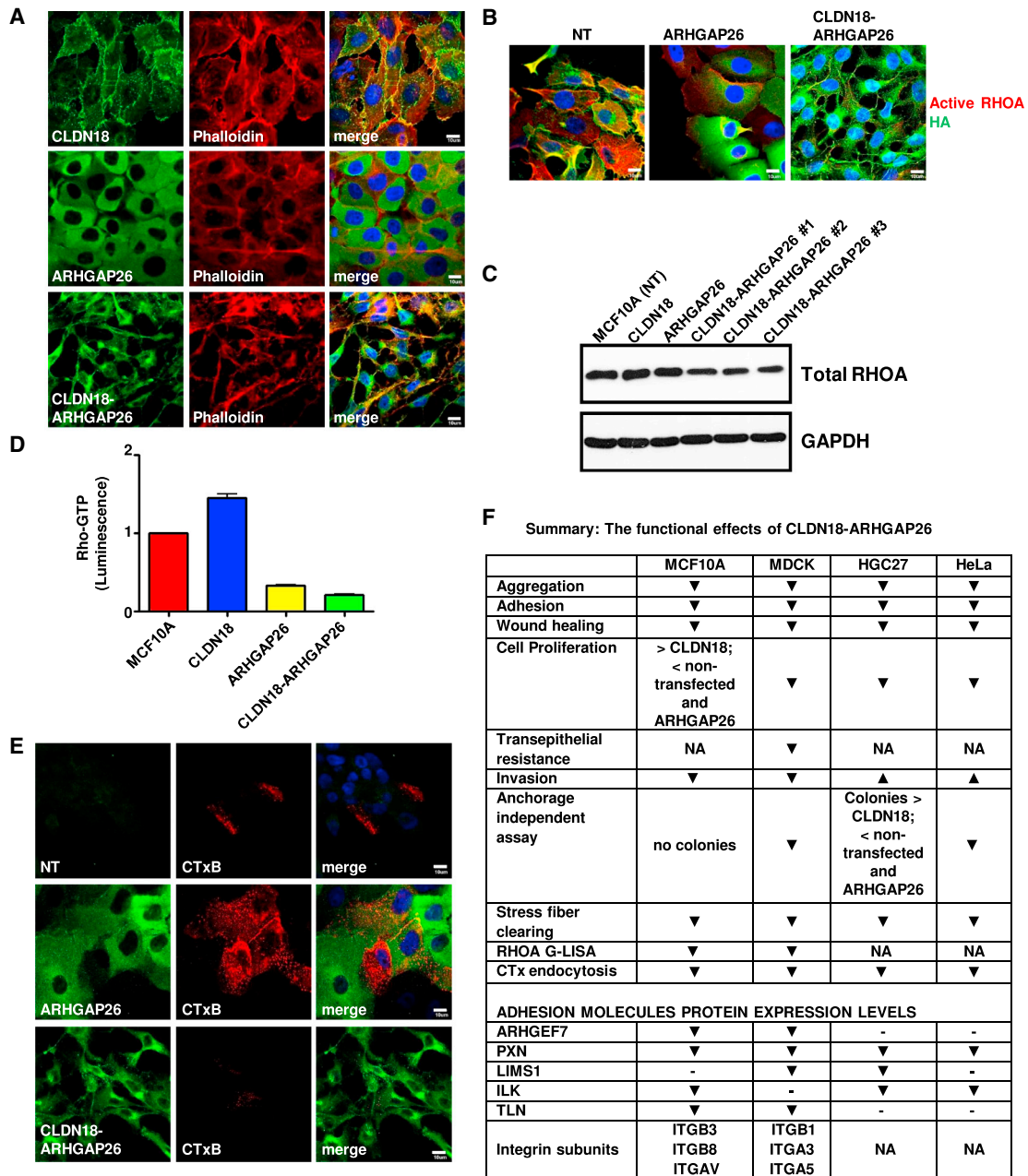
In our study, we observed recurrent somatic SVs and recurrent fusion genes in GC. Our simulations show that the rate of recur-

rent fusion genes could not be explained by chance, indicating that specific rearrangements are more likely to occur than others and/or that selective processes enrich for such rearrangements. By comparing the somatic SVs with a genome-wide view of chromatin interactions, we found significantly more overlaps of rearrangement sites with chromatin interactions than expected by chance, suggesting that the chromatin structure contributes to recurrent fusions of distant loci in GC.

We validated 136 fusion genes, evaluated their expression and reading frame characteristics, and identified five of these fusion genes as recurrent by an extended screen. Our detailed analysis of CLDN18-ARHGAP26 showed its functional properties by promoting both early cancer development and late disease progression. *CLDN18* and *ARHGAP26* are expressed in the gastric mucosa epithelium, where CLDN18 localizes to TJs and ARHGAP26 to punctate tubular vesicular structures of parietal epithelial cells. In normal gastric epithelia, we observed *ARHGAP26* expression to be specific to parietal cells. The *CLDN18-ARHGAP26* fusion gene under the influence of *CLDN18* promoter is probably expressed in all gastric epithelial cells and not just parietal cells, introducing the functions of ARHGAP26, in particular its RHOA GAP activity, to cell types normally not expressing this protein. Moreover, the *CLDN18-ARHGAP26* fusion gene links functional protein domains of a regulator of RHOA to a TJ protein, resulting in altered properties. These, as well as the aberrant localization of the GAP activity to the plasma membrane via the four transmembrane domains of CLDN18 possibly result in changes to cellular functions that are associated with GC.

Integrins and associated cytoplasmic focal adhesion complexes play central roles in ECM adhesion, epithelial differentiation, cell spreading, and cell migration (Lee and Gotlieb, 1999). ARHGAP26 interacts via its SH3 domain with the focal adhesion components PTK2 and PXN (Doherty et al., 2011). CLDN18-ARHGAP26 cells showed reduced levels of integrin-focal adhesion complex constituents, which is generally associated with decreased cell-cell and cell-ECM adhesion, and changes to ECM integrin cytoskeletal dynamics (Priddle et al., 1998; Zhang et al., 2008; Zheng et al., 2012) in MCF10A and MDCK cells.

ARHGAP26 negatively regulates RHOA activity via the GAP domain (Taylor et al., 1999), which is retained in the fusion protein. RHOA regulates many cellular processes, including cell-ECM and cell-cell adhesion, cytoskeletal dynamics, and vesicular trafficking, among others. Recently, recurrent mutations of *RHOA* have been identified in GC, resulting in proliferation gain and reduced anoikis (Kakiuchi et al., 2014; Wang et al., 2014). In our study, we identified another mechanism of altering RHOA function in GC. We observed lower levels of activated RHOA (e.g., GTP-RHOA) in both *ARHGAP26*- and *CLDN18-ARHGAP26*-expressing cells as compared to *CLDN18*-expressing MCF10A and MDCK cells. Stress fibers were no longer observed in four cell lines expressing either ARHGAP26 or the fusion protein, suggesting that the GAP domain of ARHGAP26, despite being tethered to the plasma membrane domain, can still inactivate RHOA involved in this process, although it remains possible that this is due to the observed reduction of total RHOA. Whereas ARHGAP26 and the fusion protein showed a similar effect on stress fibers, it is conceivable that the two proteins show different effects in epithelial cells, where the fusion



**Figure 7. CLDN18 and ARHGAP26 Modulate Epithelial Phenotypes**

(A) Actin cytoskeletal staining of MCF10A cells expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26*. Cells were immunostained with HA for CLDN18- and CLDN18-ARHGAP26-expressing cells and phalloidin conjugated with Alexa Fluor 594 fluorescence.

(B) Active RHOA immunofluorescence analysis in MCF10A NT cells and stables expressing *ARHGAP26* and *CLDN18-ARHGAP26*. Cells were stained with an antibody to active RHOA and DAPI.

(C) Western blot analysis of total RHOA in NT MCF10A and cells expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26*. Cells were immunostained with RHOA antibody and GAPDH.

(D) Reduced GAP activity in MCF10A stables expressing *ARHGAP26* and *CLDN18-ARHGAP26*. The GAP activity was analyzed in a pull-down assay (G-LISA; Cytoskeleton). The amount of endogenous active GTP-bound RHOA was determined in a 96-well plate coated with RDB domain of Rho-family effector proteins. The GTP form of RHO from cell lysates of the different stable lines bound to the plate was determined with RHOA primary antibody and secondary antibody conjugated to HRP. Luminescence values were calculated relative to NT MCF10A cells.

(E) Live MCF10A cells expressing *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* were incubated with Alexa-Fluor-594-conjugated CTxB for 15 min at 37°C followed by washing and fixation. Cells were immunostained with HA or GFP antibody and DAPI.

(F) Summary of the functional effects of *CLDN18-ARHGAP26* in MCF10A, MDCK, HGC27, and HeLa.

Data are presented as mean ± SD (D). Scale bars in (A), (B), and (E) represent 10 μm.

protein may mediate a localized inactivation of RHOA at junctional adhesion complexes and thereby affect the cortical actin cytoskeleton and cell-cell adhesion. Indeed, the epithelial morphology of ARHGAP26-expressing cells is less compromised than that of cells expressing the fusion. Another difference between ARHGAP26 and the fusion may result from the observation that GAP activity of oligophrenin-1 (OPHN1), like ARHGAP26 a member of the GRAF family, may be regulated by intramolecular interactions involving the N-terminal BAR and PH domains (Elvers et al., 2012). Because these domains are missing from the fusion protein, the GAP activity may no longer be regulated in the fusion protein. The BAR and PH domains are also required for the role of ARHGAP26 in clathrin-independent endocytosis (Doherty et al., 2011). Receptor tyrosine kinases can be endocytosed by both clathrin-dependent and independent pathways. Hence, disturbances in endocytic trafficking, as observed in cells overexpressing the fusion, could influence receptor tyrosine kinase signaling.

Isoform 2 of CLDN18 is stomach specific and a component of the paracellular barrier to hydrochloric acid (Hayashi et al., 2012). *Cldn18*<sup>-/-</sup> mice present with paracellular leakage of luminal H<sup>+</sup>, accompanied by inflammation and atrophic gastritis (Hayashi et al., 2012). Although MDCK-ARHGAP26 showed a similar TER as control cells, barrier function was abolished in cells expressing the fusion. In addition, the GAP domain linked to CLDN18 could lead to a localized junctional inactivation of RHOA, affecting the cortical actin at epithelial junctions.

Whether CLDN18-ARHGAP26 directly affects the intactness of the paracellular barrier by interfering with the function of CLDN18 or the localized inactivation of RHOA at junctions or indirectly by weakening cell-ECM, cell-cell adhesion (Arnold et al., 2013; Ma et al., 2013), or cytoskeletal alterations, the scenario results in an increased chronic paracellular H<sup>+</sup> permeability, which is associated with acute and atrophic gastritis, a risk factor for GC. Coupled with poor wound healing, damaged epithelial cells may not be actively replaced, resulting in gaps that enhance tissue damage, eventually leading to GC. Once the *CLDN18-ARHGAP26*-expressing cells become cancerous, they develop an invasive tendency, which is crucial to cancer progression.

## EXPERIMENTAL PROCEDURES

### Clinical Tumor Samples

Patient samples and clinical information were obtained from patients who had undergone surgery for GC at the National University Hospital, Singapore, and Tan Tock Seng Hospital, Singapore. Informed consent was obtained from all subjects, and the study was approved by the Institutional Review Board of the National University of Singapore (reference code 05-145) as well as the National Healthcare Group Domain Specific Review Board (reference code 2005/00440).

### DNA/RNA Extraction from Samples

Genomic DNA and total RNA extraction from tissue samples was performed using Allprep DNA/RNA Mini Kit (QIAGEN). Genomic DNA was extracted from blood samples with Blood & Cell Culture DNA kit (QIAGEN).

### Antibodies and Reagents

Primary and secondary commercial antibodies and reagents are described in the Supplemental Experimental Procedures.

### Cell Culture Conditions and Transfections

MCF10A, MDCK, HeLa, HGC27, and TMK1 cell lines were cultured according to standard conditions. Transient and stable transfection experiments of MDCK, HGC27, and HeLa were carried out using JetPrimePolyPlus transfection kit according to manufacturer's instructions. Stable transfectants were generated with G418 selection as outlined in previous protocol (Kausalya et al., 2001).

### DNA-PET Libraries Construction, Sequencing, Mapping, and Data Analysis

DNA-PET library construction of 10-kb fragments of genomic DNA, sequencing, mapping, and data analysis were performed as described in Hillmer et al. (2011) with refined bioinformatics filtering as described in Nagarajan et al. (2012). The short reads were aligned to the NCBI human reference genome build 36.3 (hg18) using Bioscope (Life Technologies). DNA-PET data of TMK1 and tumors 17, 26, 28, and 38 have been described previously (Hillmer et al., 2011; GEO:GSE26954) and of tumors 82 and 92 (Nagarajan et al., 2012; NCBI GEO: GSE30833). The SOLiD sequencing data of the eight additional tumor/normal pairs, the ChIA-PET-sequencing data of HGC27, and the exome sequencing data of tumors 136, 07K611, and 05/0304 can be accessed at NCBI's Sequence Read Archive (SRA): SRP035443. Procedures for the identification of recurrent genomic breakpoints of *CLDN18-ARHGAP26*, filtering of germline SVs in cancer genomes, and breakpoint distribution analyses are described in the Supplemental Experimental Procedures.

### HGC27/MCF-7 RNA Polymerase II ChIA-PET and GC DNA-PET Comparison

To investigate whether the two partner sites of germline and somatic SVs of our study were enriched for loci that are in proximity of each other in the nucleus, we tested for overlap of SVs with genome-wide chromatin interaction data sets derived from ChIA-PET sequencing of HGC27 and further of the breast cancer cell line MCF-7 with the rationale that some chromatin interactions might be conserved across different cell types. For HGC27 ChIA-PET description and further details, see the Supplemental Experimental Procedures.

### Driver Fusion Gene Prediction

The potential for driver fusion genes was predicted by using the tool as described (Wu et al., 2013), and the threshold value 0.37 was set for identifying the potential fusion drivers.

### In-Frame Fusion Gene Confirmation and Screening by RT-PCR

One microgram of total RNA was reverse transcribed to cDNA using SuperScript III First-Strand Synthesis System for RT-PCR (Life Technologies) according to the manufacturer's instruction. PCR was done with JumpStart REDAccuTaq LA DNA Polymerase (Sigma-Aldrich).

### Generation of Cell Lines

The pMXs-Puro retroviral vectors containing the particular fusion genes were co-transfected with pVSVG (pseudotyping construct) into GP2-293 cells using Lipofectamine 2000 to produce virus. HGC27, HeLa, and MCF10A cells were then infected with the viral supernatant containing expression constructs. Stable transfectants were obtained and maintained under selection pressure by puromycin dihydrochloride (Sigma; P9620). Viral transfection was used to establish all cell lines except *CLDN18*, *ARHGAP26*, and *CLDN18-ARHGAP26* in MDCK, HGC27, and HeLa, which were transfected by Lipofectamine 2000 (Life Technologies). Under selection conditions, clones were picked and maintained.

### Cell Aggregation, Cell Adhesion, and Wound Healing Assays

For cell aggregation assay, 20  $\mu$ l of  $1.2 \times 10^6$ /ml cells were plated on tissue culture dishes as hanging drops and phase contrast images were obtained the next day using Nikon Eclipse TE2000-S. For cell adhesion assay, detailed information is provided in the Supplemental Experimental Procedures. For wound healing assay, 70  $\mu$ l of  $7 \times 10^5$  cells/ml were plated on culture insert in  $\mu$ -Dish 35 mm (Ibidi). Prior to seeding, the  $\mu$ -Dish plates were treated with collagen type 1. The following day, the insert was peeled off to create a wound

and migration was imaged with Nikon Eclipse TE2000 until closure of the wound.

#### Cell Proliferation Assay

Eight hundred cells were seeded in quadruplicates for each condition in 24-well plates, and readings were taken according to manufacturer's instructions (Cell Proliferation Reagent WST-1; Roche) for 7 days. Absorbance was measured using Infinite M200 Quad4 Monochromator (Tecan) at 450 nm using a reference wavelength of 650 nm.

#### Cell Invasion Migration Assay

$5 \times 10^4$  stably transfected cells in RPMI serum-free media were plated into the Biocoat Matrigel invasion chamber according to manufacturer's instructions (Corning) with 5% FBS in media added as chemoattractant to the wells of the Matrigel invasion chamber for 24 hr. Further detail is provided in the [Supplemental Experimental Procedures](#).

#### TER Analysis

$2 \times 10^5$  stably transfected MDCK cells were seeded on 12-mm Transwell inserts (Corning) to obtain a polarized monolayer. The next day, the inserts were placed in CellZscope (nanoAnalytics) for TER measurements.

#### Soft Agar Colony Formation Assay

Five thousand cells of stable cell lines were added to 2 ml soft agar (0.35% Noble agar and 2× FBS media) and plated onto solidified base layers (0.7% Noble agar with 2× FBS) with triplicates set up for each experiment. Colonies were counted 2 to 4 weeks later.

#### ACCESSION NUMBERS

DNA-PET data for TMK1 and tumors 17, 26, 28, and 38 reported in this paper have been deposited to the NCBI GEO and are available under accession number GEO:GSE26954. Tumors 82 and 92 have been deposited to the GEO under accession number GEO:GSE30833. SOLiD sequencing data of the eight additional tumor/normal pairs; the ChIA-PET-sequencing data of HGC27; and the exome sequencing data of tumors 136, 07K611, and 05/0304 have been deposited to the NCBI Sequence Reads Archive and are available under accession number SRA:SRP035443.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Results, Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2015.06.020>.

#### AUTHOR CONTRIBUTIONS

Y.R., K.G.Y., and P.T. initiated the study. F.Y. and A.M.H. coordinated the study. F.Z., J.B.Y.S., J.R., and K.G.Y. recruited clinical samples and information and contributed to interpretation. F.Y., A.S.M.T., and Z.Z. performed DNA/RNA extractions and constructed DNA-PET libraries; X.R. managed SOLiD sequencing; S.T.L., S.C.N., and P.S.D.C. performed SOLiD sequencing; W.-K.S. oversaw DNA-PET bioinformatics analysis; W.H.L., P.N.A., X.Y.W., S.C., F.H.M., W.T.P., L.V., and C.S.C. performed bioinformatics analysis; J.H.J.T. performed breakpoint distribution and driver fusion gene analysis; F.Y. and A.S.M.T. performed fusion gene validation PCRs and RT-PCRs; P.-E.J. contributed active chromatin data and discussions; D.B. performed SV reoccurrence simulations under guidance of N.N.; H.M.P. constructed LMP ChIA-PET library of HGC27; F.Y. constructed sequence capture libraries; G.L. performed breakpoint/ChIA-PET correlation analysis; P.G. performed the dPET/cPET ratio of SVs breakpoints; X.L. performed mutation detection of whole-exome sequencing; S.S.L. performed FISH analysis; J.P.K., Y.Y.S., and E.G.Y.C. constructed expression plasmids; J.P.K. performed cell biological work of *CLDN18-ARHGAP26* with help of A.G.M.O. and with advice from W.H.; Y.Y.S. contributed to *CLDN18-ARHGAP26* functional studies and performed functional assays of four

additional fusion genes under guidance of T.N.P.; F.Y., J.P.K., W.H., and A.M.H. wrote the manuscript; and Y.R., A.S.M.T., and W.-K.S. contributed to it.

#### ACKNOWLEDGMENTS

We thank all patients who consented to donate tissue samples. This work was supported by the Agency for Science Technology and Research, Singapore, the Translational Clinical Research Flagship Programme, The Singapore Gastric Cancer Consortium - Improving Outcomes for Our Patients. Additional support was provided by the Genome Institute of Singapore and the Institute of Molecular and Cell Biology internal research funds from the National Medical Research Council of Singapore and the Biomedical Research Council. We thank Sebastian Maurer-Stroh and Vachirane Limviphuvadh for discussions on fusion protein properties; Guillaume Bourque for discussions and advice on genomics; Hwee Meng Low, Yeen Hui Choy, and Jan Dachlauer for help on DNA and RNA extractions and PCR work; Valère Cacheux for advice on FISH analysis; Soon Wei Jia Wendy and the Next Generation Sequencing group for the Illumina sequencing; and Advanced Molecular Pathology Laboratory, IMCB, A\*STAR, Singapore for histological services. A.M.H., Y.R., F.Y., P.T., W.H., A.S.M.T., Y.Y.S., and K.G.Y. applied for a National University of Singapore, Singapore Health Services and Agency for Science, Technology and Research, Singapore patent (international patent application no. PCT/SG2015/050047) 'Fusion Genes in Cancer'.

Received: June 26, 2014

Revised: April 21, 2015

Accepted: June 6, 2015

Published: July 2, 2015

#### REFERENCES

- Arnold, K.M., Goeckeler, Z.M., and Wysolmerski, R.B. (2013). Loss of focal adhesion kinase enhances endothelial barrier function and increases focal adhesions. *Microcirculation* 20, 637–649.
- Berger, M.F., Lawrence, M.S., Demichelis, F., Drier, Y., Cibulskis, K., Sivachenko, A.Y., Sboner, A., Esgueva, R., Pflueger, D., Sougnez, C., et al. (2011). The genomic complexity of primary human prostate cancer. *Nature* 470, 214–220.
- Borkhardt, A., Bojesen, S., Haas, O.A., Fuchs, U., Bartelheimer, D., Loncaric, I.F., Bohle, R.M., Harbott, J., Repp, R., Jaeger, U., et al. (2000). The human GRAF gene is fused to MLL in a unique t(5;11)(q31;q23) and both alleles are disrupted in three cases of myelodysplastic syndrome/acute myeloid leukemia with a deletion 5q. *Proc. Natl. Acad. Sci. USA* 97, 9168–9173.
- Calderwood, D.A. (2004). Tailin controls integrin activation. *Biochem. Soc. Trans.* 32, 434–437.
- Campbell, P.J., Yachida, S., Mudie, L.J., Stephens, P.J., Pleasance, E.D., Stebbings, L.A., Morsberger, L.A., Latimer, C., McLaren, S., Lin, M.L., et al. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467, 1109–1113.
- Cristescu, R., Lee, J., Nebozhyn, M., Kim, K.M., Ting, J.C., Wong, S.S., Liu, J., Yue, Y.G., Wang, J., Yu, K., et al. (2015). Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat. Med.* 21, 449–456.
- Davenport, H.W. (1972a). The gastric mucosal barrier. *Digestion* 5, 162–165.
- Davenport, H.W. (1972b). Why the stomach does not digest itself. *Sci. Am.* 226, 87–93.
- Davenport, H.W. (1975). The gastric mucosal barrier: past, present, and future. *Mayo Clin. Proc.* 50, 507–514.
- Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108.
- Doherty, G.J., and Lundmark, R. (2009). GRAF1-dependent endocytosis. *Biochem. Soc. Trans.* 37, 1061–1065.



- Doherty, G.J., Åhlund, M.K., Howes, M.T., Morén, B., Parton, R.G., McMahon, H.T., and Lundmark, R. (2011). The endocytic protein GRAF1 is directed to cell-matrix adhesion sites and regulates cell spreading. *Mol. Biol. Cell* 22, 4380–4389.
- Elvers, M., Beck, S., Fotinos, A., Ziegler, M., and Gawaz, M. (2012). The GRAF family member oligophrenin1 is a RhoGAP with BAR domain and regulates Rho GTPases in platelets. *Cardiovasc. Res.* 94, 526–536.
- Hayashi, D., Tamura, A., Tanaka, H., Yamazaki, Y., Watanabe, S., Suzuki, K., Suzuki, K., Sentani, K., Yasui, W., Rakugi, H., et al. (2012). Deficiency of claudin-18 causes paracellular H<sup>+</sup> leakage, up-regulation of interleukin-1 $\beta$ , and atrophic gastritis in mice. *Gastroenterology* 142, 292–304.
- Hildebrand, J.D., Taylor, J.M., and Parsons, J.T. (1996). An SH3 domain-containing GTPase-activating protein for Rho and Cdc42 associates with focal adhesion kinase. *Mol. Cell. Biol.* 16, 3169–3178.
- Hillmer, A.M., Yao, F., Inaki, K., Lee, W.H., Ariyaratne, P.N., Teo, A.S., Woo, X.Y., Zhang, Z., Zhao, H., Ukil, L., et al. (2011). Comprehensive long-span paired-end-tag mapping reveals characteristic patterns of structural variations in epithelial cancer genomes. *Genome Res.* 21, 665–675.
- Janjigian, Y.Y., and Kelsen, D.P. (2013). Genomic dysregulation in gastric tumors. *J. Surg. Oncol.* 107, 237–242.
- Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA Cancer J. Clin.* 61, 69–90.
- Jovov, B., Van Itallie, C.M., Shaheen, N.J., Carson, J.L., Gambling, T.M., Anderson, J.M., and Orlando, R.C. (2007). Claudin-18: a dominant tight junction protein in Barrett's esophagus and likely contributor to its acid resistance. *Am. J. Physiol. Gastrointest. Liver Physiol.* 293, G1106–G1113.
- Kakiuchi, M., Nishizawa, T., Ueda, H., Gotoh, K., Tanaka, A., Hayashi, A., Yamamoto, S., Tatsuno, K., Katoh, H., Watanabe, Y., et al. (2014). Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma. *Nat. Genet.* 46, 583–587.
- Kausalya, P.J., Reichert, M., and Hunziker, W. (2001). Connexin45 directly binds to ZO-1 and localizes to the tight junction region in epithelial MDCK cells. *FEBS Lett.* 505, 92–96.
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645.
- Lee, T.Y., and Gotlieb, A.I. (1999). Early stages of endothelial wound repair: conversion of quiescent to migrating endothelial cells involves tyrosine phosphorylation and actin microfilament reorganization. *Cell Tissue Res.* 297, 435–450.
- Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148, 84–98.
- Lundmark, R., Doherty, G.J., Howes, M.T., Cortese, K., Vallis, Y., Parton, R.G., and McMahon, H.T. (2008). The GTPase-activating protein GRAF1 regulates the CLIC/GEEC endocytic pathway. *Curr. Biol.* 18, 1802–1808.
- Ma, Y., Semba, S., Khan, R.I., Bochimoto, H., Watanabe, T., Fujiya, M., Kohgo, Y., Liu, Y., and Taniguchi, T. (2013). Focal adhesion kinase regulates intestinal epithelial barrier function via redistribution of tight junction. *Biochim. Biophys. Acta* 1832, 151–159.
- Mellman, I., and Yarden, Y. (2013). Endocytosis and cancer. *Cold Spring Harb. Perspect. Biol.* 5, a016949.
- Nagarajan, N., Bertrand, D., Hillmer, A.M., Zang, Z.J., Yao, F., Jacques, P.E., Teo, A.S., Cutcutache, I., Zhang, Z., Lee, W.H., et al. (2012). Whole-genome reconstruction and mutational signatures in gastric cancer. *Genome Biol.* 13, R115.
- Nobes, C.D., and Hall, A. (1995). Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* 81, 53–62.
- Pridle, H., Hemmings, L., Monkley, S., Woods, A., Patel, B., Sutton, D., Dunn, G.A., Zicha, D., and Critchley, D.R. (1998). Disruption of the talin gene compromises focal adhesion assembly in undifferentiated but not differentiated embryonic stem cells. *J. Cell Biol.* 142, 1121–1133.
- Qian, J., Qian, Z., Lin, J., Yao, D.M., Zhu, Z.H., Chen, Q., Xiao, G.F., and Li, J.Y. (2010). Aberrant methylation of GTPase regulator associated with the focal adhesion kinase (GRAF) promoter is an adverse prognostic factor in myelodysplastic syndrome. *Eur. J. Haematol.* 85, 174–176.
- Rhead, B., Karolchik, D., Kuhn, R.M., Hinrichs, A.S., Zweig, A.S., Fujita, P.A., Diekhans, M., Smith, K.E., Rosenbloom, K.R., Raney, B.J., et al. (2010). The UCSC Genome Browser database: update 2010. *Nucleic Acids Res.* 38, D613–D619.
- Sahin, U., Koslowski, M., Dhaene, K., Usener, D., Brandenburg, G., Seitz, G., Huber, C., and Türeci, O. (2008). Claudin-18 splice variant 2 is a pan-cancer target suitable for therapeutic antibody development. *Clin. Cancer Res.* 14, 7624–7634.
- Sanders, M.J., Ayalon, A., Roll, M., and Soll, A.H. (1985). The apical surface of canine chief cell monolayers resists H<sup>+</sup> back-diffusion. *Nature* 313, 52–54.
- Stephens, P.J., McBride, D.J., Lin, M.L., Varela, I., Pleasance, E.D., Simpson, J.T., Stebbings, L.A., Leroy, C., Edkins, S., Mudie, L.J., et al. (2009). Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* 462, 1005–1010.
- Taylor, J.M., Macklem, M.M., and Parsons, J.T. (1999). Cytoskeletal changes induced by GRAF, the GTPase regulator associated with focal adhesion kinase, are mediated by Rho. *J. Cell Sci.* 112, 231–242.
- Tran, N.L., Adams, D.G., Vaillancourt, R.R., and Heimark, R.L. (2002). Signal transduction from N-cadherin increases Bcl-2. Regulation of the phosphatidylinositol 3-kinase/Akt pathway by homophilic adhesion and actin cytoskeletal organization. *J. Biol. Chem.* 277, 32905–32914.
- Vannella, L., Lahner, E., and Annibale, B. (2012). Risk for gastric neoplasias in patients with chronic atrophic gastritis: a critical reappraisal. *World J. Gastroenterol.* 18, 1279–1285.
- Wang, K., Kan, J., Yuen, S.T., Shi, S.T., Chu, K.M., Law, S., Chan, T.L., Kan, Z., Chan, A.S., Tsui, W.Y., et al. (2011). Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat. Genet.* 43, 1219–1223.
- Wang, K., Yuen, S.T., Xu, J., Lee, S.P., Yan, H.H., Shi, S.T., Siu, H.C., Deng, S., Chu, K.M., Law, S., et al. (2014). Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat. Genet.* 46, 573–582.
- Wu, C.C., Kannan, K., Lin, S., Yen, L., and Milosavljevic, A. (2013). Identification of cancer fusion drivers using network fusion centrality. *Bioinformatics* 29, 1174–1181.
- Yao, F., Ariyaratne, P.N., Hillmer, A.M., Lee, W.H., Li, G., Teo, A.S., Woo, X.Y., Zhang, Z., Chen, J.P., Poh, W.T., et al. (2012). Long span DNA paired-end-tag (DNA-PET) sequencing strategy for the interrogation of genomic structural mutations and fusion-point-guided reconstruction of amplicons. *PLoS ONE* 7, e46152.
- Zang, Z.J., Cutcutache, I., Poon, S.L., Zhang, S.L., McPherson, J.R., Tao, J., Rajasegaran, V., Heng, H.L., Deng, N., Gan, A., et al. (2012). Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. *Nat. Genet.* 44, 570–574.
- Zhang, X., Jiang, G., Cai, Y., Monkley, S.J., Critchley, D.R., and Sheetz, M.P. (2008). Talin depletion reveals independence of initial cell spreading from integrin activation and traction. *Nat. Cell Biol.* 10, 1062–1068.
- Zheng, Q., Chen, S., Chen, Y., Lyga, J., and Santhanam, U. (2012). Critical role of paxillin in aging of human skin. *J. Invest. Dermatol.* 132, 1290–1293.