**Supplementary methods**

**RNA-sequence**

Five µg of the total RNA extracted by RNeasy plus mini kit (Qiagen) was treated with a GeneRead rRNA Depletion kit (Qiagen) to remove ribosomal RNA (rRNA) and then with RNeasy MiniElute kit (Qiagen) for purification. For fragmentation, 100 ng of the rRNA-depleted RNA was incubated at 95°C for 10 min and was purified by a Magnetic Beads Cleanup Module (Life Technologies).

Libraries were constructed using Ion Total RNA-Seq Kit for AB Library Builder System (Thermo Fisher Scientific). Templates were prepared using Ion PI Hi-Q Chef Kit (Thermo Fisher Scientific on Ion Chef (Thermo Fisher Scientific). Sequence runs in Ion PI Chip V3 (Thermo Fisher Scientific) were performed on Ion Proton (Thermo Fisher Scientific) using IonPI Hi (Hi-Q Sequencing Kit) (Thermo Fisher Scientific). Alignment of reads to reference hg19 were performed using the RNASeqAnalysis plugin from the Ion torrent suite software. Mapped reads were counted for each gene using HTSeq v 0.9.1 htseq-cout. Count-based differential expression analysis was performed on edge R v 3.16.5 after removal of low count lead genes using three biological replicates for each condition (less than 5 reads per gene in the sample and counts per million mapped reads (CPM) of 1 or less).

**Gene Ontology Analysis and KEGG pathway analysis**

GO and KEGG analysis were performed using DAVID v6.8 on web site (<https://david.ncifcrf.gov/home.jsp>). The results were visualized using GOplot v1.0.2 of an R package.

**Gene set enrichment analysis**

GSEA was performed using FPKM value of RNA-sequencing data to genes included in GENE\_SYMBOL.chip as "expression dataset" and M1417, M1406 and hallmark (50 gene sets) of systematic name in GSEA as "Gene sets database" by using version 3.0. (http://software.broadinstitute.org/gsea/index.jsp).

**Chromatin immunoprecipitation assay (ChIP)-qPCR**

Approximately 10,000,000 cancer cells per a sample were used for ChIP procedure. Cells were cross-linked with 1% formaldehyde for 10 min at 37ºC followed by an addition of 125 mM glycine for terminates cross-linking. Cells were then treated two times with cell shearing buffer (0.1% SDS, 50 mM Tris-HCl (pH 7.6), I mM EDTA (pH8.0), 0.002% Triton X-100, ~0.04 mM PMSF) with protease inhibitors (Complete tablet; Roche Molecular Biochemicals) to solubilize chromatin. Lysates were sonicated using Bioruptor UCD-300 (Cosmo Bio Co.) so that size of DNA fragments became around 300 bp or less. Sonicated chromatin was incubated with antibody (rabbit polyclonal anti-Bach1 (A1-6; generated in-house) coupled to Dynabeads protein A/G (Veritas) for 14 hrs at 4°C. Beads were washed two times with RIPA buffer (10mM Tris-HCl pH7.6, 1 mM EDTA, 0.1% SDS, 0.1% deoxyholic acid sodium salt, 1% Triton X-100) with protease inhibitors (Complete tablet), RIPA buffer plus 0.3M NaCl, RIPA buffer plus LiCl buffer (0.21 M LiCl, 0.5% NP-40, 0.5% NaDOC) and TE buffer plus 0.2% Triton X-100, washed once with TE buffer with protease inhibitor (Complete tablet). To reverse cross-links, washed Dynabeads were incubated in elution buffer (0.003% SDS, 10 mM Tris-HCl (pH8.0) and 1 mM EDTE (pH 8.0), 0.1 mg/ml Proteinase K) for 4 hrs at 65°C. ChIP-DNA was purified by AMPure XP (BECKMAN COULTER) and the DNA concentration was determined using a Qubit fluorometer (Invitrogen).

 The purified ChIP sample was used for PCR with a gene specific primer set with Light Cycler Fast Start Essential DNA Green Master, and the binding amount was measured with Light Cycler Nano of a quantitative PCR machine. The binding amount of the target gene was normalized by amounts of amplified target region in input DNA and evaluated by relative quantification.

**Supplementary Figure and Table**

**Figure S1. The expression amounts of BACH1 in pancreatic cancer cell lines and the effect of BACH1 silencing on cellular proliferation in wild type KRAS cells**

(A) *BACH1* mRNA level in pancreatic cancer cell lines.

(B) BACH1 protein level in pancreatic cancer cell lines.

(C) *BACH1* mRNA level in *BACH1* silenced BxPC-3 cells (scrambled siRNA used as control).

(D) Cell proliferation of BACH1 silenced BxPC-3 cells seeded at indicated concentrations.

**Figure S2. Gene ontology, KEGG pathway and Gene set enrichment analysis of BACH1 silenced AsPC-1**

(A, B) The outer columns show a scatter plot of expression level (logFC) for differentially expressed genes enriched in (A) biological process of gene ontology (GO) term and (B) KEGG pathway term. Magenta and blue dots show up- and down-regulation in samples with BACH1 silencing, respectively. The inner circle shows a bar plot. The size of the bar represents significance of terms (*P* value) and color indicates the z-score (magenta, increased; blue, decreased; white, unchanged). Analysis was performed using genes by DAVID v3.8 and plot drawn by using ‘Goplot’ in R package. All terms, *P* < 0.05.

(C) Immunoblot analysis of BACH1, p44/42 MAPK and phosho-p44/p42 MAPK in control siRNA- and BACH1 siRNA-treated AsPC1 cells.

(D) GSEA against gene set “JECHLINGER EPITHELIAL TO MESENCHYMAL TRANSITION UP” and “JECHLINGER EPITHELIAL TO MESENCHYMAL TRANSITION DN” in gene expression of BACH1 silenced AsPC1 cells and the control cells.

(E) GSEA against gene set “hallmark” using gene expression of BACH1 silenced AsPC-1 cells and the control cells. Left table shows gene sets that were significantly higher in cells with si*Control* than those with si*Bach1* at FDR < 25% and nominal P value < 1%. There were no gene sets which were significantly upregulated in si*BACH1* cells.

**Figure S3. The effects of BACH1 silencing on cellular migration in wild type KRAS cells**

(A) Wound-healing assay of *BACH1* silenced BxPC-3 cells (scrambled siRNA used as control). Representative photos (left) and area of wound closure (right) are shown. Scale bar, 100 µm.

(B) RT-qPCR analyses of the indicated transcripts (relative to *ACTB* as a control). mRNA expression is normalized to the expression in control cells.

All data are presented as mean ± SD, with p values from the Student`s *t* test. \**P* < 0.05; \*\**P* < 0.01; n.s., not significant.

**Figure S4. Immunofluorescence staining of BACH1 and E-cadherin in AsPC-1/sg*BACH1* cells and the expression amounts of BACH1 in pancreatic cancer cell line models used**

(A) Immunofluorescence staining of BACH1 in BACH1 WT/sg*Cont* and BACH1 KO/sg*BACH1*-2 AsPC-1 cells. Scale bar, 7.5 μm.

(B) Immunofluorescence staining of E-cadherin in BACH1 WT/sg*Cont* and BACH1 KO/sg*BACH1*-2 AsPC-1 cells. Scale bar, 75 μm.

(C) *BACH1* mRNA level in pancreatic cancer cell line models of knockout (sg*BACH1*), silencing (si*BACH1*), and overexpression (ex*BACH1*).

(D) BACH1 and Nrf2 protein levels in the indicated pancreatic cancer cell line models.

**Figure S5. Induction of epithelial gene expression in AsPC1 cells lacking BACH1**

(A) RNA-sequencing data of *BACH1* gene in AsPC-1/sg*BACH1*-2 cells and control cells (empty vector (lentiCRISPR v2) used as control), analyzed in triplicates.

(B) Some of the typical RNA-sequence reads of sg*BACH1*-2 target sequence region shown in Fig. 6A in AsPC-1/sg*BACH1*-2 cells and control cells.

(C) The dendrogram using all genes (at least reads per a gene of one sample is with reads counts > 5 and CPM > 1). Cluster analysis was performed using ‘hclust’ function in R soft ware. RNA-sequencing data of *BACH1* knockdown experiment is from Fig. 2.

(D) Heatmap of RNA-sequence results in control and BACH1 knockout samples (n = 3). Genes with fold change ≥ 2 & P value < 0.05 are shown.

(E) Heatmap shows the RNA-sequence results (counts per a million mapped reads (CPM)) of arbitrarily selected EMT-related genes in Fig. 2B. CPM was calculated on HTseq v0.9.1.

(F) The dendrogram using only EMT-related genes shown in (E).

(G) Venn diagram of genes showing ≥ 1.5-fold up (RNA-seq UP) and down (RNA-seq DOWN) regulation in BACH1 knockdown and knockout samples compared with control samples from RNA-seq (Table 2 and Table 3) with putative target genes of BACH1 which were detected in ChIP-seq analysis on AsPC-1 and SW1990.

**Figure S6. Examination of metastatic activity of BACH1 knockout AsPC1 cells**

(A) GSEA against gene set “hallmark” using gene expression of AsPC-1/sg*BACH1*-2 cells and the control cells. Left Table shows gene sets, which were significantly upregulated in AsPC-1 cells deleted for *BACH1* at FDR < 25% and nominal P value < 1%. Right figure showed top 2 in enrichment plot.

(B) Left Table shows gene sets, which were significantly upregulated in AsPC-1/sg*Control* cells under the same condition as (A). Right plots show two out of three gene sets which were also detected in GSEA of AsPC-1/siRNA treated cells (see Supplementary Fig. S2E left Table).

(C) Representative photos of primary tumor, liver metastasis and mesentery metastasis of BACH1 WT/sg*Control* and BACH1 KO/sg*BACH1*-2 AsPC-1 cells.

**Figure S7.** **BACH1 high expression is correlated with poor prognosis**

Data of RNA-sequence obtained from the TCGA (n = 176) were classified into two groups (BACH1 target gene Low and High) with cutoff values obtained from the ROC curve. Overall survival curve by the Kaplan-Meier method between the two groups are shown. *P* value was obtained by Log-rank test. Further overall survival curve by the Kaplan-Meier method between the four groups (combinationof *BACH1* and indicated BACH1 target genes) are shown. *P* value was obtained by Log-rank test.

**Table S1. Primer information**

**Table S2. The list of genes which were changed expression of mRNA by BACH1 silencing**

Gene expression differences in si*BACH1* and si*Control*.

**Table S3. The list of genes which were changed in their mRNA expression by *BACH1* deletion**

Gene expression differences in sg*BACH1* and sg*Control*.

**Table S4. Clinicopathological features in PDAC**

Statistical data of BACH1 expression and clinicopathological features for surgical PDAC. LN: lymph node.

**Table S5. Variate analysis in PDAC**

Univariate analysis and multivariate analysis of clinicopathological factors for overall survival.

**Table S6. Clinicopathological features in PDAC of the TCGA**

Statistical data of BACH1 mRNA expression and clinicopathological features for data of RNA-sequence obtained from the TCGA (n = 176)