Title: KPNA3 promotes epithelial-mesenchymal transition by regulating TGF-β and AKT signaling pathways in MDA-MB-231, a triple-negative breast cancer cell line

Corresponding Author: Kwang-Ho Lee

Authors: Jaesung Choi[1,#], Jee-Hye Choi[1,#], Ho Woon Lee[1,#], Dongbeom Seo[1,#], Gavaachimed Lkhagvasuren[2], Jung-Woong Kim[1], Sang-Beom Seo[1], Kangseok Lee[1], Kwang-Ho Lee[1,2,*]

Institution: [1]Department of Life Science, College of Natural Sciences, and [2]Department of Science of Cultural Heritage, Graduate School, Chung-Ang University,

1 **Manuscript Type:** Article

2

3 **Title: KPNA3 promotes epithelial–mesenchymal transition by regulating TGF-β and**

4 **AKT signaling pathways in MDA-MB-231, a triple-negative breast cancer cell line**

5

6 **Authors' names:** Jaesung Choi[1,†], Jee-Hye Choi[1,†], Ho Woon Lee[1,†], Dongbeom Seo[1,†],

7 Gavaachimed Lkhagvasuren[2], Jung-Woong Kim[1], Sang-Beom Seo[1], Kangseok Lee[1], and

8 Kwang-Ho Lee[1,2,*]

9

10 **Affiliation:** [1]Department of Life Science, College of Natural Sciences, Chung-Ang University,

11 06974, Seoul, Republic of Korea, [2]Department of Science of Cultural Heritage, Graduate

12 School, Chung-Ang University, Seoul, Korea

13

14 [†]These authors contributed equally to this work

15

16 **Running Title:** Regulation of EMT by KPNA3 in breast cancer

17 **Keywords:** KPNA3, Epithelial–mesenchymal transition, GATA3, HAS2, Triple-negative

18 breast cancer

19

20 ***Corresponding author:**

21 Kwang-Ho Lee, Ph.D., Professor

22 Department of Life Science, College of Natural Sciences,

23 Chung-Ang University, 84 Heuksuk-Ro, Dongjak-Ku, 06974, Seoul, Republic of Korea

24 E-mail address: leemanse@cau.ac.kr

25 Tel.: +82-2-820-5213; Fax: +82-2-825-5206

Abstract

Karyopherin-α3 (KPNA3), a karyopherin- α isoform, is intimately associated with metastatic progression via epithelial-mesenchymal transition (EMT). However, the molecular mechanism underlying how KPNA3 acts as an EMT inducer remains to be elucidated. In this report, we identified that KPNA3 was significantly upregulated in cancer cells, particularly in triple-negative breast cancer, and its knockdown resulted in the suppression of cell proliferation and metastasis. The comprehensive transcriptome analysis from KPNA3 knockdown cells indicated that KPNA3 is involved in the regulation of numerous EMT-related genes, including the downregulation of GATA3 and E-cadherin and the up-regulation of HAS2. Moreover, it was found that KPNA3 EMT-mediated metastasis can be achieved by TGF-β or AKT signaling pathways; this suggests that the novel independent signaling pathways KPNA3-TGF-β-GATA3-HAS2/E-cadherin and KPNA3-AKT-HAS2/E-cadherin are involved in the EMT-mediated progress of TNBC MDA-MB-231 cells. These findings provide new insights into the divergent EMT inducibility of KPNA3 according to cell and cancer type.

Abbreviations: AKT, serine/threonine kinase; ECM, extracellular matrix; EMT, epithelial–mesenchymal transition; ERK, extracellular signal-related kinase; GATA3, GATA binding protein 3; HAS2, hyaluronan synthase 2; KPNA3, karyopherin-α 3; SMAD, small mothers against decapentaplegic; TGF-β, transforming growth factor-β; TNBC, triple-negative breast cancer; TWIST, twist-related protein

1. Introduction

Epithelial–mesenchymal transition (EMT) is a key cellular process in which immotile epithelial cells transform into mesenchymal cells through cell polarity loss, cell-cell junction disassembly, and extracellular matrix (ECM) alteration. EMT endows tumor cells with enhanced migratory and invasive properties necessary for metastasis, the primary cause of cancer-related deaths (1). However, the EMT process activated by the pleiotropic control of intrinsic and extrinsic factors has inherent flexibility and variation across different cancer cells and types (2). Therefore, understanding the intricate network among EMT-related genes in various cancers will provide insight into the differences in EMT-mediated metastatic pathways and lead to the development of advanced antimetastatic therapies.

Karyopherin-α 3 (KPNA3), a member of the nuclear transport protein family, is important in the nucleocytoplasmic trafficking of certain cargoes via a heterodimeric interaction with importin-β1 (also known as KPNB1) (3). KPNA3 is upregulated in colon and liver cancers, and its enhanced expression is associated with poor prognosis and a low survival rate in patients with breast cancer (4). Moreover, it has been reported that EMT can be induced by the KPNA3-serine/threonine kinase (AKT)-extracellular signal-related kinase (ERK)-twist-related protein (TWIST) signaling cascade in hepatocellular carcinoma (HCC) (5). Despite evidence of an EMT-inducing role of KPNA3 in multiple cancers, the function of KPNA3 in EMT-associated transcriptional reprogramming remains to be clarified.

In this study, it was found that KPNA3 regulates numerous EMT-related genes and induces the EMT process via at least two independent signaling pathways in the highly invasive triple-

negative breast cancer (TNBC) cell line MDA-MB-231. One pathway is the transforming growth factor-β (TGF-β) signaling pathway that downregulates GATA binding protein 3 (GATA3) to suppress E-cadherin and upregulate hyaluronan synthase 2 (HAS2); the other is the AKT signaling pathway, which is also involved in the up-regulation of HAS2 and down-regulation of E-cadherin. Our findings highlighted that EMT induction by KPNA3 can be achieved by the networking and interplay among many genes involved in several EMT-related signaling pathways. In addition, these results suggest that KPNA3 can trigger EMT via two axes in TNBC MDA-MB-231 cells, KPNA3-TGF-β-GATA3-HAS2/E-cadherin and KPNA3-AKT-HAS2/E-cadherin.

2. Results

2.1. KPNA3 is highly expressed in TNBC and closely associated with poor patient outcomes

To determine the clinical relevance of KPNA3 in breast cancer, its protein expression was compared in different breast cancer subtypes. The expression level of KPNA3 was significantly higher in the more aggressive TNBCs than in normal and luminal-type breast cancer ($p < 0.05$ and $p < 0.01$, respectively; Fig. 1A and B). Consistent with these findings, the mRNA and protein expression levels of KPNA3 were markedly higher than those of other KPNAs in TNBC cell lines (Hs578T, BT549, and MDA-MB-231) and non-TNBC cell lines (MCF7 and T47D) (Fig. S1A and S1B, Table S1). In the analysis of different tumor grades, the expression level of KPNA3 gradually increased with an increase in tumor grade (Fig. S1C). Kaplan–Meier

91   survival analysis using GENT2 revealed that KPNA3 expression was closely associated with

92   shortened overall survival ($p = 0.001$, Fig. 1C). Taken together, the results retrieved from

93   various web-based databases clearly suggest that KPNA3 is highly expressed in aggressive

94   breast cancer cells and tissues indicating that KPNA3 may play a key role in breast cancer

95   progression and metastasis.

96

97   *2.2. Depletion of KPNA3 inhibits cell proliferation and TNBC metastasis*

98

99   To investigate the functional significance of KPNA3 expression in TNBC cells, the

100 *KPNA3* gene was knocked down in two TNBC cell lines (MDA-MB-231 and Hs578T) using

101 two siRNAs for KPNA3 (KPNA3-1-knockdown (KD) and KPNA3-2-KD) (Fig. 1D and Fig.

102 S2A). WST-1 assays revealed that both KPNA3-1-KD and KPNA3-2-KD inhibited cell

103 proliferation in MDA-MB-231 cells ($p < 0.05$). However, only KPNA3-1-KD reduced cell

104 proliferation in Hs578T cells ($p < 0.01$; Fig. 1E and S2B). Next, we investigated whether

105 KPNA3-KD's inhibitory effect on MDA-MB-231 cells' proliferation was caused by apoptosis

106 or cell cycle delay. The protein expression levels of apoptotic markers were not changed by

107 KPNA3-KD (Fig. S3A). Moreover, KPNA3-KD induced cell cycle arrest at the G1/S phase,

108 given the increase and decrease in the cell numbers in the G0/G1 and S phases, respectively

109 (Fig. S3B), suggesting that KPNA3-KD inhibits cell proliferation through cell cycle arrest at

110 the G1/S phase and not through apoptosis in MDA-MB-231 cells. Additionally, transwell

111 assays were performed to evaluate the effect of KPNA3 on metastasis. KPNA3-KD inhibited

112 cell migration ($p < 0.001$) and invasion ($p < 0.0001$) in MDA-MB-231 and Hs578T cells (Fig.

5

113   1F and S2C).

114

115   *2.3. KPNA3-KD inhibits cell migration through the downregulation of HAS2 in MDA-MB-231*

116   *cells*

117

118   To further elucidate the molecular mechanism whereby KPNA3 silencing decreases the

119   proliferation and metastasis of TNBC, transcriptome analysis using RNA sequencing was

120   performed on KPNA3-KD-231 cells (Fig. S4A). A volcano plot and heatmap were constructed

121   to indicate the transcripts' general scattering and to filter the differentially expressed genes

122   from the transcriptome profiles, respectively (Fig. 2A and S4B). Of the total 26,679 transcripts

123   annotated, 2,245 genes (976 upregulated genes and 1,269 downregulated genes) were filtered

124   by applying the criteria of absolute fold change (FC) > 1.5 and adjusted $p < 0.05$.

125   To identify EMT-related genes regulated by KPNA3, the expression profiles of genes

126   downregulated in the transcriptome profiles retrieved from KPNA3-KD-231 cells were

127   compared to two different bio-informatic data sets, the upregulated gene list in TNBC cells,

128   and the EMT-core gene list, which were derived from at least 10 gene expression study datasets

129   (Fig. S4C) (6, 7). Intriguingly, *HAS2* was the only gene common among the three expression

130   profiles (Fig. 2B). The expression of HAS2 was assessed in KPNA3-KD-231 cells to ascertain

131   whether KPNA3 can regulate it. Fig. 2C and D show that KPNA3-KD significantly reduced

132   *HAS2*'s mRNA and protein expression levels. In addition, a positive correlation between the

133   expression levels of KPNA3 and HAS2 ($p = 3.3e-52$; $R = 0.44$) was confirmed using the

134   GEPIA2 database, based on TCGA retrieved from cancer samples from over 11,000 patients

6

135   over 12 years (Fig. S5A). Furthermore, HAS2-KD exhibited significant reductions in both

136   migratory ($p < 0.001$) and invasive ($p < 0.05$) properties and showed that the extent of reduction

137   was greater in migratory properties (Fig. S5B and C).

138   Next, MDA-MB-231 cells were co-transfected with a HAS2 overexpression vector

139   (3×FLAG-HAS2) and KPNA3-KD. As a result, overexpression of HAS2 restored the

140   migratory properties ($p < 0.001$) of KPNA3-KD-231 cells but not the invasive properties or the

141   expression of E-cadherin, a representative EMT marker (Fig. 2E and F). These results suggest

142   that KPNA3 plays a critical role in EMT induction through HAS2 regulation. These results

143   imply that the EMT or mesenchymal–epithelial transition characteristics of cells whose

144   transcriptional program has already been altered by KPNA3-KD cannot be fully restored by

145   the control of HAS2 alone.

146

147   *2.4. KPNA3 promotes EMT-mediated metastasis via down-regulation of the transcription*

148   *factor GATA3*

149

150   To identify which transcription factor plays a critical role in KPNA3-mediated EMT,

151   enrichment analysis of KPNA3-KD-induced transcriptional signatures was performed with the

152   Enrichr tool using gene expression signatures derived from the GEO database for transcription

153   factor perturbations. As shown in Fig. 3A, GATA3 was ranked highest among 265 transcription

154   factors ($p < 8.37e\text{-}29$). Accordingly, western blot and RT-qPCR analyses were performed to

155   determine whether KPNA3 regulates GATA3 expression. The mRNA and protein expression

156   levels of GATA3 were significantly increased by KPNA3-KD (Fig. 3B and C). Transwell

157 chamber assays indicated that the KPNA3-KD-mediated reduction in migratory and invasive

158 properties were partially but significantly restored by an siRNA of GATA3 (GATA3-KD; $p <$

159 0.0001 for migration and $p < 0.001$ for invasion; Fig. 3D and E), suggesting that GATA3 is a

160 major downstream target in KPNA3-mediated EMT.

161 To determine whether KPNA3-KD-mediated GATA3 upregulation affects the expression of the

162 above-mentioned genes, the expression levels of E-cadherin and HAS2, were assessed after

163 dual knockdown by KPNA3-KD and GATA3-KD. E-cadherin expression, which was

164 upregulated by KPNA3-KD, was reduced by GATA3-KD (Fig. 3F), indicating that GATA3 is

165 involved in the regulation of E-cadherin; this is consistent with the findings of Yan *et al*. (8).

166 In addition, it was found that GATA3-KD upregulated HAS2 expression, which was

167 downregulated by KPNA3-KD (Fig. 3F), suggesting the involvement of the KPNA3-GATA3-

168 HAS2/E-cadherin cascade. Moreover, the regulatory mechanism of HAS2 expression by

169 GATA3 was assessed by screening the potential binding sites of GATA3 on the *HAS2* promoter

170 region spanning 2,000 bp upstream from the transcription start site using the JASPAR database.

171 Five putative GATA3-binding sites were detected on the *HAS2* promoter, implying that *HAS2*

172 is transcriptionally regulated by direct binding of GATA3 (Fig. S6). Collectively, these findings

173 suggest that KPNA3 induces EMT-mediated metastasis by inhibiting GATA3, which regulates

174 E-cadherin and HAS2 in MDA-MB-231 cells.

175

176 *2.5. KPNA3 promotes EMT-mediated metastasis through the regulation of TGF-β and AKT*

177 *signaling pathways*

178

8

179 Signaling pathways associated with genome-wide transcriptional reprogramming in

180 KPNA3-KD-231 cells were analyzed with the Enrichr tool using the BioPlanet 2019 database.

181 Enrichment analysis revealed that the "TGF-β signaling pathway for the regulation of ECM"

182 was the most important pathway out of the 1,658 human pathways in KPNA3-KD-231 cells

183 (Fig. 4A). Furthermore, TGF-β signaling was transduced through small mothers against

184 decapentaplegic (SMAD) and non-SMAD pathways. These pathways are mediated by TGF-β

185 ligands, type 1 and type 2 receptors, and SMAD or non-SMAD proteins, including AKT,

186 ERK1/2, and p38 mitogen-activated protein kinase (9). Therefore, to determine whether TGF-

187 β signaling was inhibited in KPNA3-KD-231 cells, the expression of pSMAD2/3 was assessed

188 by western blot analysis. As shown in Fig. 4B, the expression levels of total SMAD2/3 were

189 unchanged, whereas KPNA3-KD inhibited the expression levels of their phosphorylated forms.

190 The RT-qPCR results revealed that the expression of *CTGF* and *PTHRP*, which are the

191 metastasis-related downstream genes of SMAD signaling, were downregulated by KPNA3-

192 KD (Fig. 4C). These findings suggest that TGF-β/SMAD is a crucial downstream signaling

193 pathway in KPNA3-mediated EMT. As shown in Fig. 4B, the expression levels of

194 phosphorylated AKT (T308 and S473) were downregulated by KPNA3-KD. In contrast, the

195 expression levels of phosphorylated ERK and TWIST1/2 were unchanged (Fig. 4B and S7);

196 this suggests that ERK-TWIST1/2 signaling may not be involved in KPNA3-mediated EMT

197 in MDA-MB-231 cells. 2-(4-Morpholinyl)-8-phenyl-4H-1-benzopyran-4-one (LY294002), an

198 inhibitor of AKT phosphorylation, reduced HAS2 expression and increased E-cadherin

199 expression, whereas the expression of GATA3 was unchanged (Fig. 4D). These results indicate

200 that phosphorylated AKT regulates the expression of HAS2 and E-cadherin but is not involved

9

201 in GATA3 expression regulation.

202 To determine whether KPNA3 regulates GATA3 through TGF-β signaling, GATA3

203 expression was evaluated after exposing KPNA3-KD-231 cells to TGF-β. TGF-β clearly

204 reduced GATA3 and E-cadherin expression, and increased HAS2 in MDA-MB-231 cells (Fig.

205 4E). However, these TGF-β-induced changes were hindered by KPNA3-KD (Fig. 4E),

206 indicating that TGF-β is critical for the regulation of GATA3 expression in KPNA3-mediated

207 EMT. Furthermore, KPNA3-KD suppressed the migratory and invasive properties induced by

208 TGF-β (Fig. 4F). However, TGF-β induced no change in the expression levels of both total and

209 phosphorylated AKT (Fig. 4E), implying that the KPNA3-KD-induced downregulation of

210 phosphorylated AKT may be mediated independently of the TGF-β signaling pathway.

211 Collectively, we conclude that KPNA3 ultimately promotes EMT-mediated metastasis through

212 the independent regulation of the TGF-β and AKT signaling pathways, suggesting two axes:

213 KPNA3-TGF-β-GATA3-HAS2/E-cadherin and KPNA3-AKT-HAS2/E-cadherin.

214

215 **3. Discussion**

216

217 The EMT process involves the loss of cell-cell junctions and remodeling of ECM through

218 genome-wide transcriptional reprogramming induced by several epithelial–mesenchymal

219 transition-associated transcription factors (EMT-TFs) and various relevant signaling pathways,

220 consequently promoting the metastasis of malignant tumors (1, 2). Emerging evidence suggests

221 that KPNA3 is closely related to metastasis via EMT in various types of cancers (4, 5). However,

222 the genome-wide regulatory mechanism of KPNA3-induced EMT in breast cancer remains

223 largely unknown. In the present study, comprehensive bioinformatic data confirmed that among

224 the KPNAs, KPNA3 is highly expressed in aggressive breast cancer cells and tissues,

225 particularly in TNBC, and is closely associated with poor prognosis. Further mechanistic

226 investigation revealed that KPNA3 triggers EMT by inducing cell-cell junction remodeling and

227 ECM through the regulation of two independent signaling pathways, including KPNA3-TGF-

228 β-GATA3-HAS2/E-cadherin and KPNA3-AKT-HAS2/E-cadherin, in TNBC MDA-MB-231

229 cells.

230

231 *3.1. EMT-inducing capacity of KPNA3 through regulation of cell-cell junctions and ECM*

232 *organization*

233

234 The transcriptome profiles retrieved from KPNA3-KD-231 cells showed significant

235 enrichment in comparison with the defined EMT-core gene list. As shown in Fig. S4C, several

236 genes shared by our transcriptome profiles and the EMT-core gene list were classified into the

237 cell adhesion and migration category. HAS2 is related to EMT induction through the synthesis

238 of hyaluronic acid (HA), a major component of ECM (10). Versican has HA-binding properties

239 and is an anti-adhesion molecule, implying the cooperative role of versican and HA in ECM

240 remodeling (11). Neuropilin-1 promotes tumor invasion through the up-regulation of vascular

241 endothelial growth factor A, which interacts with ECM components (12). Junction plakoglobin,

242 a member of the catenin protein family, is a cytoplasmic component comprising desmosomes

243 and adherens junctions (13). E-cadherin is the most well-known member of the cadherin family

244 and is closely associated with EMT induction when its expression is significantly reduced (14).

11

245 In addition, matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases

246 (TIMPs) are associated with ECM degradation and remodeling (15). The transcriptome profiles

247 retrieved from KPNA3-KD-231 revealed that the expression of MMP-1 was downregulated

248 (-2.001 FC; $p < 0.0001$), whereas that of TIMPs, which suppress MMP expression and activity,

249 were upregulated (Table S2) (16). These data strongly suggest that KPNA3 is a strong EMT

250 inducer in TNBC and that its EMT-inducing capacity is attributed to altered expression levels

251 of numerous genes, which mainly regulate cell-cell junctions and ECM organization.

252

253 *3.2. A novel KPNA3-TGF-β-GATA3-HAS2/E-cadherin signaling cascade that promotes EMT*

254

255 The TGF-β/SMAD signaling pathway is well known to play an important role in inducing

256 EMT through the up-regulation of EMT-TFs, such as snail family transcriptional repressor 1/2

257 (SNAI1/2), zinc-finger E-box-binding homeobox 1/2 (ZEB1/2), and TWIST1/2, which have

258 been widely accepted as representative EMT inducers that regulate the expression of EMT-

259 related genes, including *CDH1* and *HAS2* (17). Our results indicated that TGF-β-mediated

260 SMAD2/3 phosphorylation and its downstream target genes, *CTGF* and *PTHRP*, were

261 downregulated in KPNA3-KD-231 cells. Additionally, the migratory and invasive properties

262 induced by extrinsic TGF-β were considerably hindered by KPNA3-KD, suggesting that

263 KPNA3-KD inhibited the TGF-β signaling pathway. However, the protein expression levels of

264 ZEB1/2, SNAI1/2, and TWIST1/2 were either upregulated or unchanged by KPNA3-KDs (Fig.

265 S7). In addition, enrichment analysis revealed that ZEB1/2, SNAI1/2, and TWIST1/2 were not

266 included in the top 10 rankings. Rather GATA3, which regulates the expression of E-cadherin

12

267 and HAS2, ranked highest among transcription factors. As shown in Fig. 3F, changes in E-

268 cadherin and HAS2 expression in GATA3-KD suggest that GATA3 is involved in regulating

269 E-cadherin and HAS2 expression. Furthermore, the exposure of cells to extrinsic TGF-β was

270 confirmed to reduce GATA3 expression, a crucial downstream target of KPNA3; this result

271 agrees with a previous study in T cells (18). However, further studies are required to elucidate

272 the molecular mechanism and whether KPNA3 regulates GATA3 expression directly or

273 indirectly through TGF-β. This finding suggests that KPNA3 triggers EMT through a novel

274 KPNA3-TGF-β-GATA3-HAS2/E-cadherin signaling cascade in TNBC MDA-MB-231 cells.

275

276 *3.3. Additional KPNA3-AKT-HAS2/E-cadherin signaling cascade to promote EMT*

277

278 It has been reported that TGF-β can trigger many non-canonical pathways, also termed

279 non-SMAD pathways (19). Particularly in MDA-MB-231 cells, exposure of extrinsic TGF-β

280 to cells increased the expression of phosphorylated AKT and enhanced MMP-9 expression and

281 activity via the ITGB1/FAK/Src/AKT/β-catenin/MMP-9 signaling cascade (20). On the

282 contrary, it was also reported that TGF-β inhibited AKT phosphorylation in MDA-MB-231

283 cells (21). Our results revealed that AKT phosphorylation was unchanged by exposure of

284 extrinsic TGF-β to MDA-MB-231 cells (Fig. 4E). In addition, the transcriptome profiles

285 retrieved from KPNA3-KD-231 cells revealed that the expression of MMP-9 remained

286 unchanged, indicating that AKT cannot be regulated TGF-β by in MDA-MB-231 cells (Table

287 S2). The reasons for these discrepancies are not fully understood. However, they may be

288 attributed to the TGF-β concentration and treatment time. Moreover, *HAS2* and *CDH1* were

13

289 regulated negatively and positively by LY294002, respectively, whereas the expression of

290 GATA3 was unchanged. These results imply that *GATA3* is regulated by the TGF-β and not

291 the AKT signaling pathway. In addition, these results suggest that KPNA3 also induces EMT

292 through the KPNA3-AKT-HAS2/E-cadherin cascade independently of the TGF-β signaling

293 pathway. However, the detailed molecular mechanism of the association between KPNA3 and

294 the EMT-related signaling pathways, TGF-β, and AKT requires further investigation.

295

296   In summary, this study reveals the EMT inducibility of KPNA3 in TNBC MDA-MB-231

297 cells. Moreover, KPNA3 triggers EMT through two axes, TGF-β-GATA3-HAS2/E-cadherin

298 and AKT-HAS2/E-cadherin, to promote tumor progression and metastasis, suggesting that

299 KPNA3 might be a putative target for the treatment of TNBC. However, for the clinical

300 application of KPNA3 as an EMT suppressor or chemotherapy drug, more comprehensive and

301 multidisciplinary studies need to be conducted on the genome-wide transcriptomic modulations

302 induced by KPNA3 in the cells.

303

304 **Materials and methods**

305 Materials and methods are available in the supplemental material.

306

307 **Conflicts of interest**

308 The authors declare that they have no conflicts of interest.

309

14

310 **Acknowledgments**

313 **References**

314 1.   Nieto MA, Huang RY, Jackson RA and Thiery JP (2016) Emt: 2016. Cell 166, 21-45

315 2.   Lamouille S, Xu J and Derynck R (2014) Molecular mechanisms of epithelial-mesenchymal
316   transition. Nat Rev Mol Cell Biol 15, 178-196

317 3.   Cingolani G, Petosa C, Weis K and Muller CW (1999) Structure of importin-beta bound to
318   the IBB domain of importin-alpha. Nature 399, 221-229

319 4.   Mehmood R, Jibiki K, Shibazaki N and Yasuhara N (2021) Molecular profiling of
320   nucleocytoplasmic transport factor genes in breast cancer. Heliyon 7, e06039

321 5.   Hu B, Cheng JW, Hu JW et al (2019) KPNA3 Confers Sorafenib Resistance to Advanced
322   Hepatocellular Carcinoma via TWIST Regulated Epithelial-Mesenchymal Transition. J Cancer
323   10, 3914-3925

324 6.   Groger CJ, Grubinger M, Waldhor T, Vierlinger K and Mikulits W (2012) Meta-analysis of
325   gene expression signatures defining the epithelial to mesenchymal transition during cancer
326   progression. PLoS One 7, e51136

327 7.   Dai X, Cheng H, Bai Z and Li J (2017) Breast Cancer Cell Line Classification and Its Relevance
328   with Breast Tumor Subtyping. J Cancer 8, 3131-3141

329 8.   Yan W, Cao QJ, Arenas RB, Bentley B and Shao R (2010) GATA3 inhibits breast cancer
330   metastasis through the reversal of epithelial-mesenchymal transition. J Biol Chem 285,
331   14042-14051

332 9.   Liu S, Chen S and Zeng J (2018) TGFbeta signaling: A complex role in tumorigenesis (Review).
333   Mol Med Rep 17, 699-704

334 10. Caon I, Bartolini B, Parnigoni A et al (2020) Revisiting the hallmarks of cancer: The role of
335   hyaluronan. Semin Cancer Biol 62, 9-19

336 11. Matsumoto K, Shionyu M, Go M et al (2003) Distinct interaction of versican/PG-M with
337   hyaluronan and link protein. J Biol Chem 278, 41205-41212

338 12. Graziani G and Lacal PM (2015) Neuropilin-1 as Therapeutic Target for Malignant Melanoma.
339   Front Oncol 5, 125

340 13. Li J, Swope D, Raess N, Cheng L, Muller EJ and Radice GL (2011) Cardiac tissue-restricted
341   deletion of plakoglobin results in progressive cardiomyopathy and activation of {beta}-
342   catenin signaling. Mol Cell Biol 31, 1134-1144

343 14. Yang J and Weinberg RA (2008) Epithelial-mesenchymal transition: at the crossroads of
344   development and tumor metastasis. Dev Cell 14, 818-829

345 15. Radisky ES and Radisky DC (2010) Matrix metalloproteinase-induced epithelial-mesenchymal
346   transition in breast cancer. J Mammary Gland Biol Neoplasia 15, 201-212

347 16. Geervliet E and Bansal R (2020) Matrix Metalloproteinases as Potential Biomarkers and

348   Therapeutic Targets in Liver Diseases. Cells 9

349   17.   Liu S, Ren J and Ten Dijke P (2021) Targeting TGFbeta signal transduction for cancer therapy.
350   Signal Transduct Target Ther 6, 8

351   18.   Heath VL, Murphy EE, Crain C, Tomlinson MG and O'Garra A (2000) TGF-beta1 down-
352   regulates Th2 development and results in decreased IL-4-induced STAT6 activation and
353   GATA-3 expression. Eur J Immunol 30, 2639-2649

354   19.   Zhang YE (2009) Non-Smad pathways in TGF-beta signaling. Cell Res 19, 128-139

355   20.   Zhu X, Wang K and Chen Y (2020) Ophiopogonin D suppresses TGF-beta1-mediated
356   metastatic behavior of MDA-MB-231 breast carcinoma cells via regulating
357   ITGB1/FAK/Src/AKT/beta-catenin/MMP-9 signaling axis. Toxicol In Vitro 69, 104973

358   21.   Boldbaatar A, Lee S, Han S et al (2017) Eupatolide inhibits the TGF-beta1-induced migration
359   of breast cancer cells via downregulation of SMAD3 phosphorylation and transcriptional
360   repression of ALK5. Oncol Lett 14, 6031-6039

361

Figure legends

Fig. 1. Karyopherin-α3 (KPNA3) promotes proliferation, migration, and invasion in triple-negative breast cancer. (**A**) Expression pattern of KPNA3 protein in different breast cancer subtypes obtained from the UALCAN database, consisting of normal (n = 18), luminal (n = 64), HER2 (n = 10), and TNBC (n = 16) subtypes. (**B**) mRNA expression pattern of KPNA3 in different breast cancer subtypes. Expression data were obtained from the GENT2 database, consisting of luminal (n = 17), luminal A (n = 379), luminal B (n = 244), HER2 (n = 230), basal (n = 363), and TNBC (n = 251) subtypes. For each subtype, the log2 fold change (FC) was calculated. (**C**) Survival rate analysis of breast cancer patients according to differential expression levels of KPNA3 obtained from the GENT2 database. Low (n = 252) and high (n = 250) expression levels of KPNA3 were divided by median expression. (**D**) Expression levels of KPNA3 protein in MDA-MB-231 cells exposed to two siRNAs of KPNA3 (KPNA3-KD) for 48 h by western blot analysis using β-actin as a loading control. (**E**) Proliferation rates in cells exposed to KPNA3-KD, as examined by the WST-1 assay at various time intervals. Data represent the mean ± standard deviation (SD) of three independent experiments. (**F**) Cell images of migrated or invaded cells under the same conditions as those in (**D**) (left panel). The number of migrated or invaded cells (right panel). Data represent the mean number of cells per five visual regions (×100) of three replicate wells.

Fig. 2. Knockdown of karyopherin-α3 (KPNA3-KD) inhibits cell migration and invasion through the transcriptional regulation of HAS2 and CDH1 in MDA-MB-231 cells (A) Volcano plot indicating the general scattering of the transcripts identified by RNA sequencing

18

384   analysis according to log2 (FC) and −log10 (*p*-value). Red-colored dots indicate genes filtered

385   with absolute FC > 1.5 and −log10 > 20. Green-colored dots represent genes filtered with

386   absolute FC > 1.5 and *p* < 0.05. Gray-colored dots indicate genes filtered with absolute FC <

387   1.5. (**B**) Venn diagram showing the overlap of genes downregulated in transcriptome profiles,

388   genes upregulated in the EMT-core gene list, and genes upregulated in triple-negative breast

389   cancer cells. (**C**) Expression level of HAS2 mRNA by RT-qPCR. Data are shown as the mean

390   ± standard deviation (SD) of three independent experiments. (**D**) Expression level of HAS2

391   protein in KPNA3-KD-231 cells for 48 h by western blot analysis. (**E**) Expression levels of

392   FLAG, E-cadherin, and KPNA3 proteins in KPNA3-KD-231 cells co-transfected with or

393   without 3× FLAG-HAS2. (**F**) Cell images of migrated or invaded cells under the same

394   conditions as those in (**D**) (left panel).

395

396   **Fig. 3. GATA binding protein 3 (GATA3) is a major downstream transcription factor in**

397   **karyopherin-α3 (KPNA3)-mediated epithelial-mesenchymal transition (EMT).** (**A**)

398   Transcription factor perturbations from GEO database-based enrichment analysis of KPNA3-

399   KD-induced transcriptional signatures (absolute fold change > 1.5, *p* < 0.05). The *p*-values

400   were computed using Fisher's exact test and converted to −log10 (*p*-value). (**B**) Expression

401   level of GATA3 mRNA by RT-qPCR. Data represent the mean ± standard deviation (SD) of

402   three independent experiments. (**C**) Expression level of GATA3 protein in KPNA3-KD-231

403   cells for 48 h by western blot analysis using β-actin as a loading control. (**D**) Expression levels

404   of GATA3 and KPNA3 proteins in KPNA3-KD-231 cells co-transfected with or without

405   GATA3 (GATA3-KD) siRNA for 48 h. (**E**) Cell images of migrated or invaded cells under the

406 same conditions as those in (E) (left panel). The number of migrated or invaded cells (right

407 panel). Data represent the mean number of cells per five visual regions (magnification, ×100)

408 of three replicate wells. (**F**) Expression levels of E-cadherin and HAS2 proteins in KPNA3-

409 KD-231 cells co-transfected with or without GATA3-KD for 48 h.

410

411 **Fig. 4. Karyopherin-α3 (KPNA3) triggers epithelial-mesenchymal transition (EMT) via**

412 **transforming growth factor-β (TGF-β) and serine/threonine kinase (AKT) signaling**

413 **pathways.** (**A**) BioPlanet 2019-based enrichment analysis of KPNA3-KD-induced

414 transcriptional signatures (absolute fold change > 2, $p < 0.05$). The p-values were computed

415 using Fisher's exact test and converted to −log10. (**B**) Expression levels of total SMAD2/3 and

416 phosphorylated SMAD2 (S465/467)/SMAD3 (S423/425), total AKT, phosphorylated AKT

417 (T308/S473), total ERK, and phosphorylated ERK (T202/Y204) proteins in KPNA3-KD-231

418 cells for 48 h. (**C**) Expression levels of CTGF and PTHRP mRNAs in KPNA3-KD-231 cells.

419 (**D**) Expression levels of E-cadherin, HAS2, GATA3, total AKT, and phosphorylated AKT

420 (T308) proteins in MDA-MB-231 cells with DMSO or LY294002 treatment (10 µM) for 24 h.

421 (**E**) Expression levels of E-cadherin, HAS2, GATA3, phosphorylated SMAD2/3, total AKT,

422 phosphorylated AKT (T308/S473), total ERK, phosphorylated ERK (T202/Y204), and KPNA3

423 proteins in KPNA3-KD-231 cells treated with or without TGF-β (5 ng/mL) for 24 h. (**F**) Cell

424 images of migrated or invaded cells under the same conditions as those in (**E**) (left panel). The

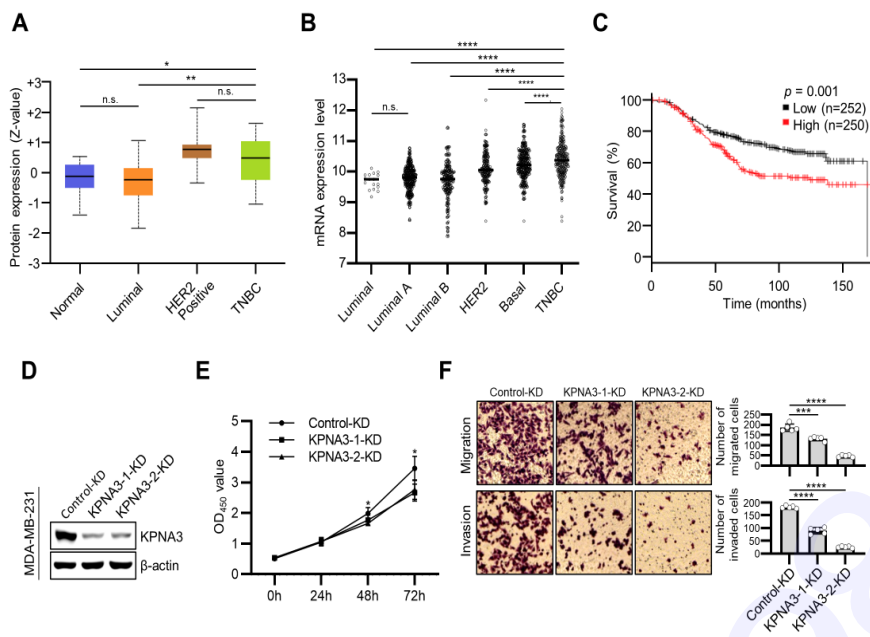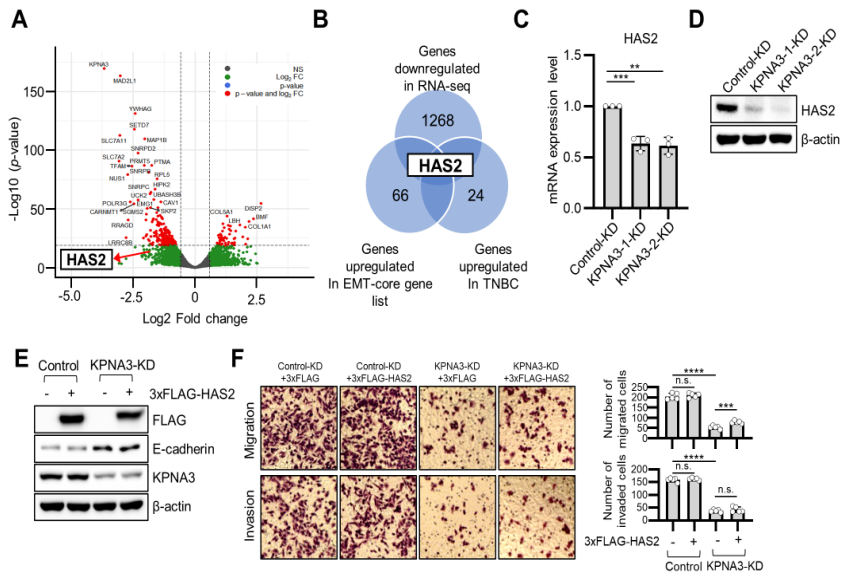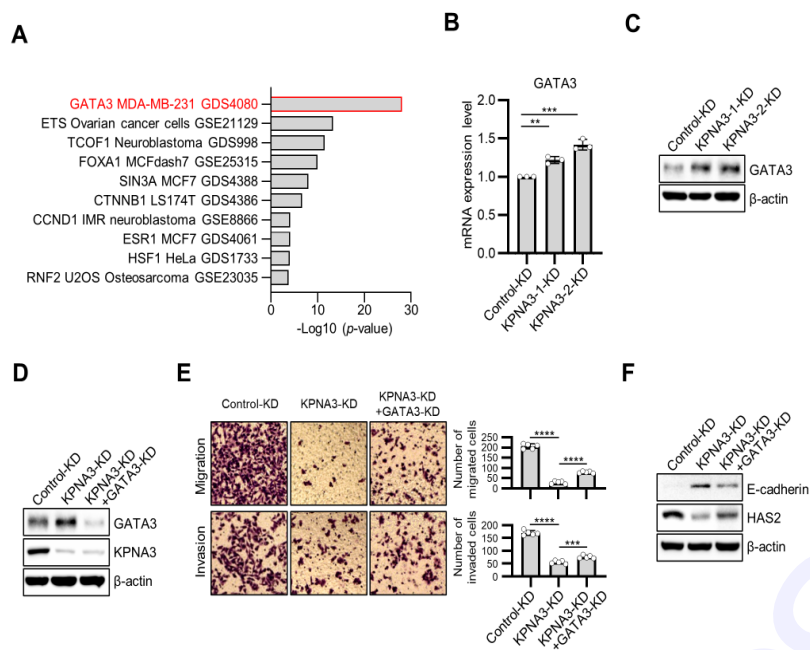425 number of migrated or invaded cells (right panel).

Fig. 1

Fig. 1.

Fig. 2

Fig. 2.

Fig. 3

Fig. 3.

A

B

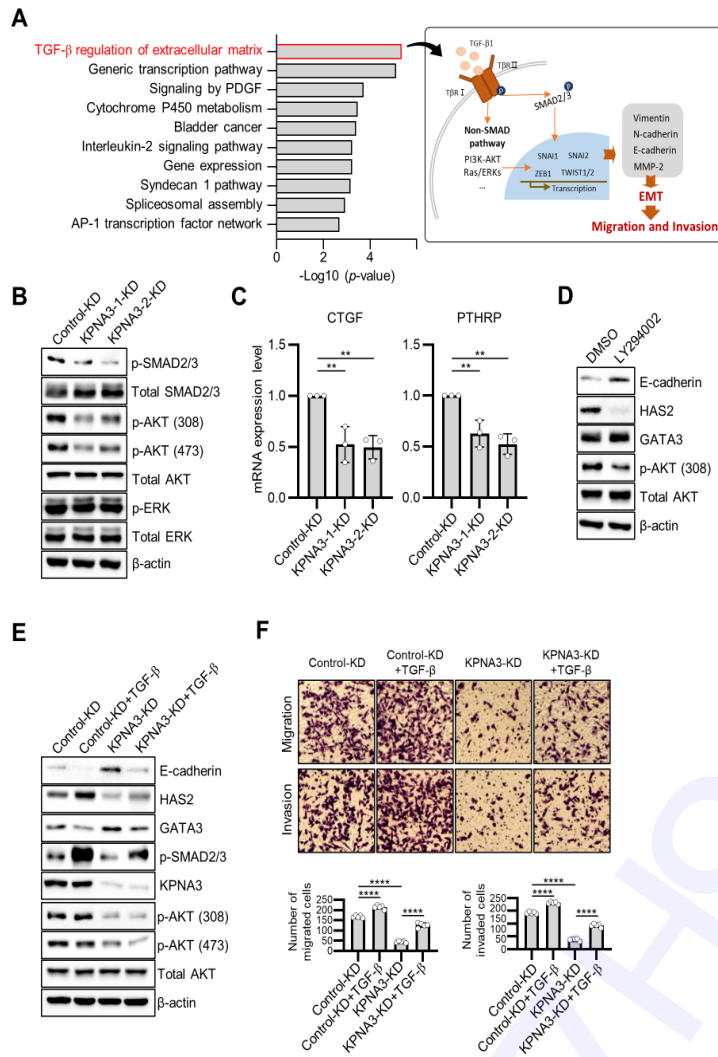**C**
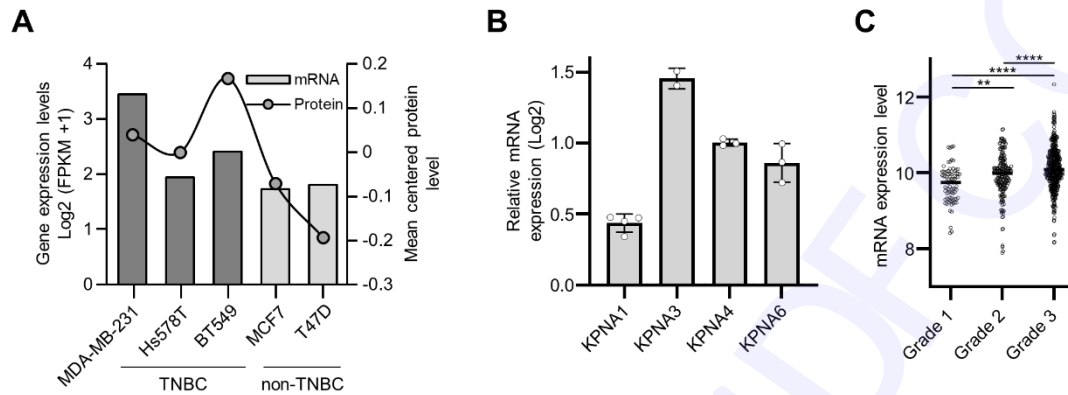
CTGF

PTHRP

D

E

F

Fig. 4
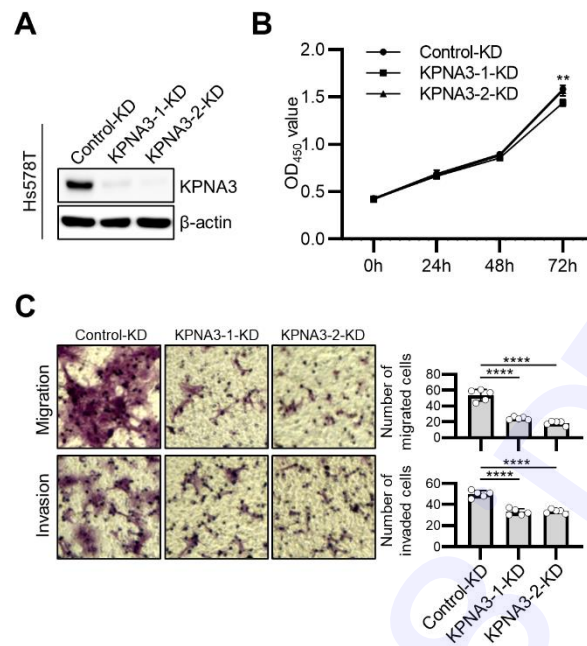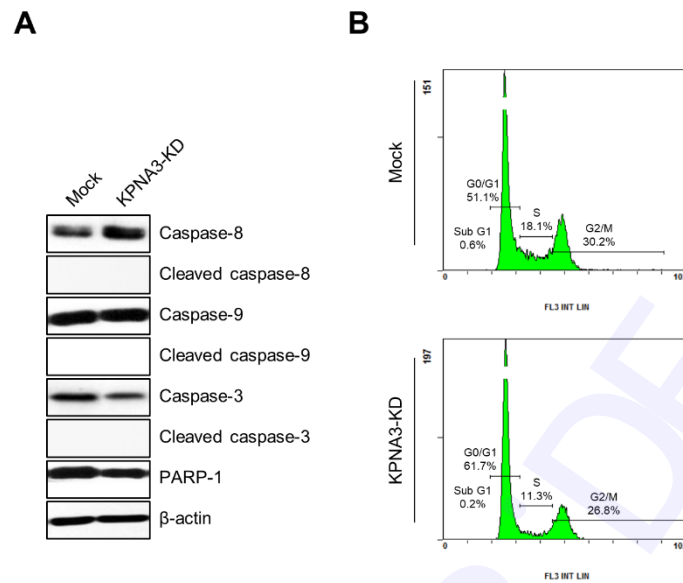
Fig. 4.

Supplementary material

Fig. S1. KPNA3 is upregulated in TNBC with poor clinical outcomes (**A**) mRNA and protein expression levels of KPNA3 in NCI-60 gene expression profiles provided by the CellMiner web application. The *y*-axis (left) represents the mRNA expression levels based on transcriptome profiles; the expression data were log2 (FPKM + 1)-transformed. The *y*-axis (right) represents the protein expression levels based on SWATH proteomic data. Data were log10-transformed and mean-centered. (**B**) Relative expression levels of KPNAs differentially expressed in TNBC (Hs578T, BT549, and MDA-MB-231) vs. non-TNBC (MCF7 and T47D) cells. Expression data were obtained from the Gene Expression Omnibus (GEO) database under the series accession number GSE32474 and analyzed using GEO2R. (**C**) Expression pattern of KPNA3 in different breast cancer grades obtained from the GENT2 database, consisting of grade 1 (n = 82), grade 2 (n = 193), and grade 3 (n = 450). For each grade, log2 (fold change) was calculated. **$p < 0.01$ and ****$p < 0.0001$ by Student's *t*-test.
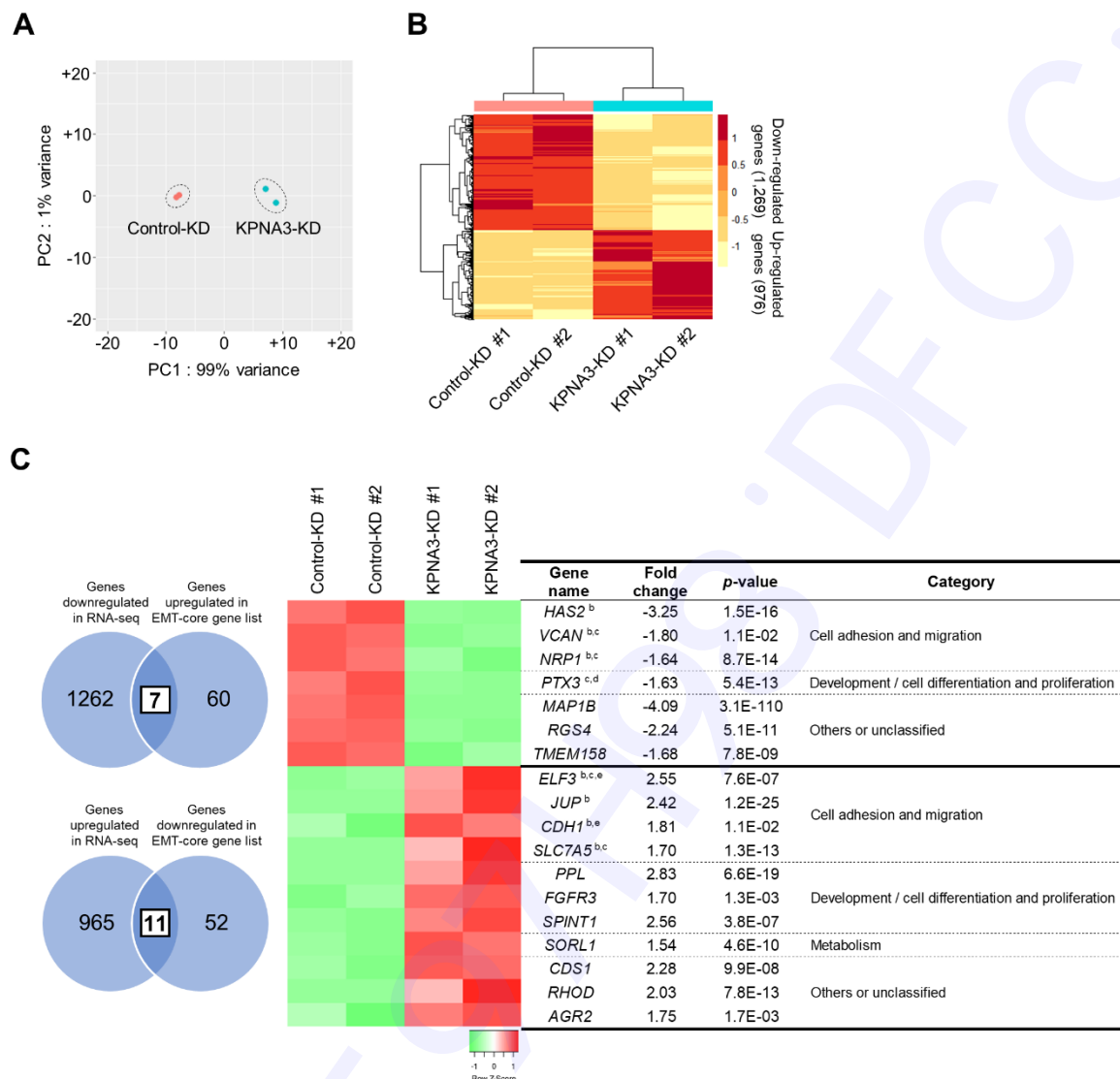
Fig. S2. Knockdown of KPNA3 inhibits migration, and invasion in Hs578T cells (**A**) Expression levels of KPNA3 protein in Hs578T exposed to two siRNAs of KPNA3 (KPNA3-KD) for 48 h by western blot analysis using β-actin as a loading control. (**B**) Proliferation rates in cells exposed to KPNA3-KD, as examined by the WST-1 assay at various time intervals. Data represent the mean ± SD of three independent experiments. (**C**) Cell images of migrated or invaded cells under the same conditions as those in (A) (left panel). The number of migrated or invaded cells (right panel). Data represent the mean number of cells per five visual regions (magnification, ×100) of three replicate wells. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, and n.s.: not significant by Student's t-test.
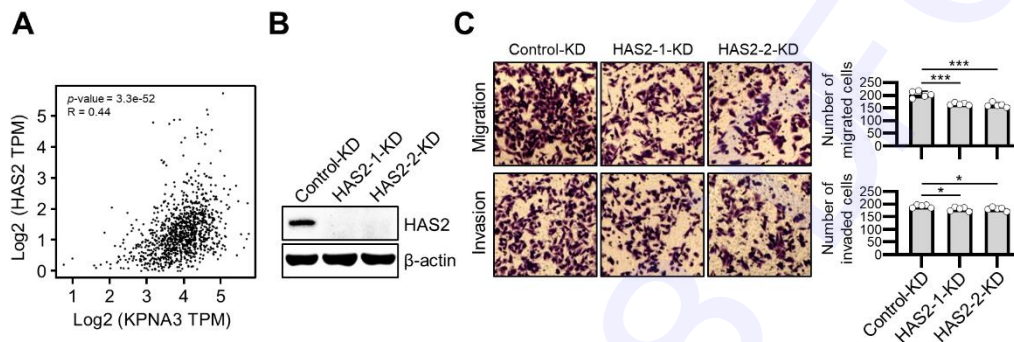
Fig. S3. Suppression of KPNA3 induces cell cycle arrest at G1/S phase in MDA-MB-231 cells (**A**) Expression levels of pro- and cleaved caspase-3, -8, and -9 and PARP-1 proteins in KPNA3-KD-231 for 48 h by western blot analysis using β-actin as a loading control. (**B**) Cell cycle distribution was analyzed by flow cytometry after transfection with KPNA3-KD for 24 h.

Fig. S4. Transcriptome analysis in KPNA3-KD-231 cells (**A**) PCA plot of transcriptome profiles from Control-KD-exposed MDA-MB-231 group (Control-KD-231) and KPNA3-KD-exposed MDA-MB-231 group (KPNA3-KD-231). Small circles indicate individual samples (red: Control-KD, blue: KPNA3-KD), and larger circles indicate each experimental group. (**B**) Hierarchically clustered heatmap of differentially expressed genes in transcriptome profiles. In total, 2,245 genes (976 upregulated and 1,269 downregulated) were filtered by absolute FC > 1.5 and adjusted p-value < 0.05. (**C**) Venn diagram showing the overlap of genes in transcriptome profiles and the genes in the EMT-core gene list (left panel).

Heatmap and summary table showing the differential expression and description of overlapped genes, respectively (right panel). Genes belonging to more than one category, according to GO classifications, are indicated as follows: b: development/cell differentiation and proliferation, c: angiogenesis and wound healing, d: metabolism, and e: apoptosis.

Fig. S5. Knockdown of HAS2 suppresses the migration and invasion of MDA-MB-231 cells (**A**) Correlation analysis between KPNA3 and HAS2 using the GEPIA2 tool (p = 3.3e-52; R = 0.44) with Spearman's coefficient. TPM: transcripts per million reads. (**B**) Expression level of HAS2 protein in MDA-MB-231 cells exposed to two siRNAs of HAS2 (HAS2-KD) for 48 h by western blot analysis using β-actin as a loading control. (**C**) Cell images of migrated or invaded cells under the same conditions as those in (B) (left panel). The number of migrated or invaded cells (right panel). Data represent the mean number of cells per five visual regions (magnification, ×100) of three replicate wells. *p < 0.05 and ***p < 0.001 by Student's t-test.

Predicted DNA binding profile	Score	Relative score	Strand	Distance to TSS	Predicted sequence
	6.893785	0.937310317	-	-1882	TGATAA
	6.666897	0.927792038	+	-233	CGATTG
	6.422963	0.917558655	-	-1928	AGATAT
	6.127317	0.905155835	+	-1669	TGATAT
	8.067203	0.906705706	-	-1880	TGATAATC

Fig. S6. Potential binding sites of GATA3 in the promoter region of HAS2 The promoter region (2000 bp upstream to TSS) of HAS2 was analyzed using the JASPAR database, and five potential GATA3-binding sites were identified (relative profile score threshold = 90%).

Fig. S7. The expression of EMT marker genes in KPNA3-KD-231 (A) Expression levels of ZEB1, ZEB2, SNAI1, SNAI2, and TWIST1/2 proteins in KPNA3-KD-231 for 48 h by western blot analysis.

Fig. S8. **Schematic representation of the two independent KPNA3-mediated EMT-Metastasis pathways in MDA-MB-231 cells**

Table S1. The relative expression levels of KPNAs differentially expressed in TNBC (Hs578T, BT549, and MDA-MB-231) vs. non-TNBC (MCF7 and T47D) cells

ID	Gene name	Log2 Fold change	*p*-value
221502_at	*KPNA3*	1.506174	7.56E-09
221503_s_at	*KPNA3*	1.404316	6.14E-09
209653_at	*KPNA4*	1.030741	2.26E-03
225267_at	*KPNA4*	0.998952	5.13E-04
226976_at	*KPNA6*	0.994623	4.18E-06
225268_at	*KPNA4*	0.983028	5.91E-04
212103_at	*KPNA6*	0.864663	4.24E-03
212102_s_at	*KPNA6*	0.721912	1.50E-02
202058_s_at	*KPNA1*	0.477763	2.27E-02
202056_at	*KPNA1*	0.473303	4.93E-02
213741_s_at	*KPNA1*	0.448327	1.68E-02
202055_at	*KPNA1*	0.342311	2.56E-02
206241_at	*KPNA5*	0.573345	6.81E-02
212101_at	*KPNA6*	0.560026	5.77E-02
202059_s_at	*KPNA1*	0.337562	1.92E-01
211762_s_at	*KPNA2*	0.3306	8.29E-02
213567_at	*KPNA4*	0.306382	3.81E-01
201088_at	*KPNA2*	0.270651	6.01E-02
229317_at	*KPNA5*	0.265871	1.34E-01
227934_at	*KPNA5*	0.171169	4.88E-01
1558383_at	*KPNA4*	-0.02111	7.93E-01
202057_at	*KPNA1*	-0.20185	4.89E-01

* The blue color data is p-value > 0.05, it was excluded from subsequent analysis.

Table S2. List of genes identified from RNA-seq

Full data (excel)

Table S3. List of siRNA sequences used in this study

Gene name	Sequence (5'-3')
Negative control	CCUCGUGCCGUUCCAUCAGGUAG
KPNA3-1	GGCAUUAACUAACAUAGCA
KPNA3-2	CUGGAUUAAUUCCUAUGAU
CDH1	CGUAUACCCUGGUGGUUCA
GATA3	ACAAGCUUCACAAUAUUAA
HAS2-1	GUAUCUGCAUCAUGCAAAA
HAS2-2	CCUCAGCAGUGUAAGAUAU

Table S4. List of RT-qPCR primers used in this study

Gene name	Forward primer (5'-3')	Reverse primer (5'-3')
CDH1	TGCACCAACCCTCATGAGTG	GTCAGTATCAGCCGCTTTCAG
GATA3	GCCCCTCATTAAGCCCAAG	TTGTGGTGGTCTGACAGTTCG
HAS2	TCGCAACACGTAACGCAAT	ACTTCTCTTTTTCCACCCCATTT
CTGF	CCAATGACAACGCCTCCTG	TGGTGCAGCCAGAAAGCTC
PTHRP	CGTCGCTGGAGCTCGATT	AATCCTGCAATATGTCCTTGGAA
TGF-β1	TCGCCAGAGTGGTTATCTT	TAGTGAACCCGTTGATGTCC
TGF-β2	ACACTCAGCACAGCAGGGTCCT	TTGGGACACGCAGCAAGGAGAAG
TGF-β3	TGAGTGGCTGTTGAGAAGAGA	ATTGTCCACGCCTTTGAATTTGAT
TβRI	GCAGAGCTGTGAAGCCTTGAGA	TGCCTTCCTGTTGACTGAGTTG
TβRII	ATGACATCTCGCTGTAATGC	GGATGCCCTGGTGGTTGA
TβRIII	TGGAGTCTCCTCTGAATGGCTG	CCATTATCACCTGACTCCAGATC

	baseMean	log2FoldC	lfcSE	pvalue	padj	ENTREZID	SYMBOL	GENENAME
ENSG0000	8.943449	0.066522	0.31993	0.490012	NA	1	A1BG	alpha-1-B glycoprotein
ENSG0000	120.1868	0.340885	0.273589	0.077485	0.173964	503538	A1BG-AS1	A1BG antisense RNA 1
ENSG0000	2.032831	-0.01524	0.3229	0.759783	NA	29974	A1CF	APOBEC1 complementation factor
ENSG0000	5.069853	-0.08371	0.335322	0.210268	NA	2	A2M	alpha-2-macroglobulin
ENSG0000	1.262011	0.025028	0.326277	0.460227	NA	127550	A3GALT2	alpha 1,3-galactosyltransferase 2
ENSG0000	442.5945	0.356524	0.176661	0.018898	0.05608	53947	A4GALT	alpha 1,4-galactosyltransferase (P blood group)
ENSG0000	891.68	0.13911	0.128582	0.238864	0.403473	8086	AAAS	aladin WD repeat nucleoporin
ENSG0000	976.1912	-0.23958	0.122844	0.034552	0.09116	65985	AACS	acetoacetyl-CoA synthetase
ENSG0000	0.47894	-0.0153	0.326545	0.594112	NA	13	AADAC	arylacetamide deacetylase
ENSG0000	11.16462	-0.19174	0.384523	0.100582	0.212744	201651	AADACP1	arylacetamide deacetylase pseudogene 1
ENSG0000	265.2577	-0.46519	0.2008	0.005878	0.020735	51166	AADAT	aminoadipate aminotransferase
ENSG0000	1286.618	-0.12916	0.118294	0.239969	0.404507	79719	AAGAB	alpha and gamma adaptin binding protein
ENSG0000	3309.905	0.025763	0.094041	0.775739	0.869709	22848	AAK1	AP2 associated kinase 1
ENSG0000	309.0328	0.602033	0.20424	0.000626	0.003051	28971	AAMDC	adipogenesis associated Mth938 domain containing
ENSG0000	3088.088	-0.07522	0.107471	0.457378	0.632262	14	AAMP	angio associated migratory cell protein
ENSG0000	2.049559	-0.0173	0.32306	0.730462	NA	15	AANAT	aralkylamine N-acetyltransferase
ENSG0000	1730.059	-0.15176	0.10804	0.135291	0.266853	25980	AAR2	AAR2 splicing factor
ENSG0000	13.08645	-0.03016	0.305103	0.785825	0.875001	441376	AARD	alanine and arginine rich domain containing protein
ENSG0000	2733.812	-0.13474	0.102205	0.16382	0.308139	16	AARS1	alanyl-tRNA synthetase 1
ENSG0000	680.5783	-0.30297	0.137042	0.014824	0.045644	57505	AARS2	alanyl-tRNA synthetase 2, mitochondrial
ENSG0000	33.22362	-0.00346	0.276805	0.979566	0.988818	80755	AARSD1	alanyl-tRNA synthetase domain containing 1
ENSG0000	1.325172	0.040909	0.329326	0.177993	NA	1.18E+08	AARSD1P1	AARSD1 pseudogene 1
ENSG0000	452.0431	-0.14787	0.158033	0.284135	0.456909	132949	AASDH	aminoadipate-semialdehyde dehydrogenase
ENSG0000	2370.133	-0.23735	0.115285	0.02711	0.074821	60496	AASDHPPT	aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase
ENSG0000	2177.132	0.374283	0.112639	0.000405	0.002095	10157	AASS	aminoadipate-semialdehyde synthase
ENSG0000	5.61095	0.060008	0.324028	0.449555	NA	284837	AATBC	apoptosis associated transcript in bladder cancer
ENSG0000	3124.254	-0.68426	0.102474	4.47E-12	1.08E-10	26574	AATF	apoptosis antagonizing transcription factor
ENSG0000	7.059972	0.042819	0.317595	0.624328	NA	9625	AATK	apoptosis associated tyrosine kinase
ENSG0000	9.408163	-0.03865	0.315572	0.667875	NA	1.03E+08	ABALON	apoptotic BCL2L1-antisense long non-coding RNA
ENSG0000	34.04672	0.651762	0.635122	0.021017	0.061053	18	ABAT	4-aminobutyrate aminotransferase
ENSG0000	2536.036	1.024053	0.101731	8.46E-25	9.62E-23	19	ABCA1	ATP binding cassette subfamily A member 1
ENSG0000	11.66136	0.077378	0.318465	0.472501	0.645106	10349	ABCA10	ATP binding cassette subfamily A member 10
ENSG0000	58.14652	-0.14906	0.283233	0.37268	0.553855	79963	ABCA11P	ATP binding cassette subfamily A member 11, pseudogene
ENSG0000	0.532849	0.014354	0.326533	0.562671	NA	26154	ABCA12	ATP binding cassette subfamily A member 12
ENSG0000	3.947063	0.062895	0.328723	0.32899	NA	154664	ABCA13	ATP binding cassette subfamily A member 13
ENSG0000	1.845772	-0.02418	0.324795	0.576403	NA	650655	ABCA17P	ATP binding cassette subfamily A member 17, pseudogene
ENSG0000	1558.672	0.313809	0.127253	0.007406	0.025339	20	ABCA2	ATP binding cassette subfamily A member 2
ENSG0000	371.5326	-0.22081	0.165462	0.122152	0.247107	21	ABCA3	ATP binding cassette subfamily A member 3
ENSG0000	11.00756	0.120924	0.336451	0.270779	0.441628	24	ABCA4	ATP binding cassette subfamily A member 4
ENSG0000	218.7401	0.11452	0.190541	0.462386	0.636681	23461	ABCA5	ATP binding cassette subfamily A member 5
ENSG0000	510.6664	1.134749	0.179808	2.61E-11	5.56E-10	10347	ABCA7	ATP binding cassette subfamily A member 7
ENSG0000	1.044721	-0.01601	0.32545	0.617384	NA	5243	ABCB1	ATP binding cassette subfamily B member 1
ENSG0000	900.8541	-1.44983	0.145447	1.55E-24	1.69E-22	23456	ABCB10	ATP binding cassette subfamily B member 10
ENSG0000	1.137989	-0.02102	0.325499	0.553068	NA	5244	ABCB4	ATP binding cassette subfamily B member 4
ENSG0000	2.268502	-0.05705	0.330911	0.212703	NA	340273	ABCB5	ATP binding cassette subfamily B member 5
ENSG0000	35.58016	0.24557	0.356667	0.161738	0.305202	10058	ABCB6	ATP binding cassette subfamily B member 6 (Langereis blood group)
ENSG0000	542.5459	-0.00336	0.13857	0.977143	0.987571	22	ABCB7	ATP binding cassette subfamily B member 7
ENSG0000	1401.078	0.272316	0.121782	0.015771	0.048091	11194	ABCB8	ATP binding cassette subfamily B member 8
ENSG0000	433.5663	0.579894	0.168757	0.00013	0.000767	23457	ABCB9	ATP binding cassette subfamily B member 9
ENSG0000	2722.377	-0.17487	0.100873	0.066846	0.154419	4363	ABCC1	ATP binding cassette subfamily C member 1
ENSG0000	1390.644	0.622615	0.119599	4.11E-08	5.05E-07	89845	ABCC10	ATP binding cassette subfamily C member 10
ENSG0000	3.023767	-0.01291	0.321051	0.826025	NA	85320	ABCC11	ATP binding cassette subfamily C member 11
ENSG0000	0.801575	0.007557	0.325203	0.781306	NA	94160	ABCC12	ATP binding cassette subfamily C member 12
ENSG0000	91.73273	0.166222	0.250001	0.33493	0.514428	1244	ABCC2	ATP binding cassette subfamily C member 2
ENSG0000	2143.201	1.09804	0.106158	4.46E-26	6.03E-24	8714	ABCC3	ATP binding cassette subfamily C member 3
ENSG0000	1582.199	-0.74739	0.112636	5.05E-12	1.21E-10	10257	ABCC4	ATP binding cassette subfamily C member 4
ENSG0000	1573.54	0.416502	0.110121	6.17E-05	0.000393	10057	ABCC5	ATP binding cassette subfamily C member 5
ENSG0000	68.53897	1.057078	0.414046	0.000758	0.003605	368	ABCC6	ATP binding cassette subfamily C member 6
ENSG0000	17.02823	0.74336	1.46438	0.013499	0.04217	653190	ABCC6P1	ATP binding cassette subfamily C member 6 pseudogene 1
ENSG0000	28.53058	1.19461	0.757967	0.004431	0.016255	730013	ABCC6P2	ATP binding cassette subfamily C member 6 pseudogene 2
ENSG0000	17.48828	-0.17781	0.358815	0.17927	0.3295	10060	ABCC9	ATP binding cassette subfamily C member 9
ENSG0000	532.7935	0.679781	0.186783	4.59E-05	0.0003	215	ABCD1	ATP binding cassette subfamily D member 1
ENSG0000	3.921607	0.101279	0.344646	0.098993	NA	225	ABCD2	ATP binding cassette subfamily D member 2
ENSG0000	3439.142	-0.05676	0.102498	0.558498	0.715496	5825	ABCD3	ATP binding cassette subfamily D member 3
ENSG0000	2062.327	0.212689	0.112738	0.043827	0.11042	5826	ABCD4	ATP binding cassette subfamily D member 4
ENSG0000	6393.479	-0.76981	0.115135	3.42E-12	8.44E-11	6059	ABCE1	ATP binding cassette subfamily E member 1
ENSG0000	7903.678	-0.06549	0.090336	0.449299	0.625468	23	ABCF1	ATP binding cassette subfamily F member 1
ENSG0000	899.3532	-0.44235	0.138269	0.000477	0.002403	10061	ABCF2	ATP binding cassette subfamily F member 2
ENSG0000	3.164517	0.034509	0.323096	0.57648	NA	344653	ABCF2P1	ATP binding cassette subfamily F member 2 pseudogene 1
ENSG0000	2238.089	0.09809	0.113068	0.355085	0.535659	55324	ABCF3	ATP binding cassette subfamily F member 3
ENSG0000	34.15401	0.895923	0.641954	0.008717	0.0291	9619	ABCG1	ATP binding cassette subfamily G member 1
ENSG0000	370.6823	-0.14528	0.173821	0.323123	0.501211	9429	ABCG2	ATP binding cassette subfamily G member 2 (Junior blood group)
ENSG0000	57.68249	0.808285	0.453548	0.006066	0.021325	64137	ABCG4	ATP binding cassette subfamily G member 4
ENSG0000	5.793359	0.053594	0.321699	0.512157	NA	84696	ABHD1	abhydrolase domain containing 1
ENSG0000	1591.507	-0.14106?	0.113926	0.101057	0.222257	55347	ABHD10	abhydrolase domain containing 10, depalmitoylase

Materials and methods

1. Cell culture and genetic characteristics of cell lines

Four human cell lines, MDA-MB-231, SK-OV-3, OVCAR-3, and Hs578T, were purchased from the Korea Cell Line Bank (KCLB, Seoul, Korea). MDA-MB-231, SK-OV-3, and OVCAR-3 were cultured in RPMI 1640 medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (FBS; GW Vitek, Seoul, Korea). Hs578T was grown in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich) supplemented with 10% FBS (GW Vitek). All the cells were incubated at 37°C under a 5% CO_2 atmosphere.

MDA-MB-231 and Hs578T are TNBC cell lines that lack the expression of estrogen receptor alpha, progesterone receptor, and human epidermal growth factor receptor 2 (HER2) (1). SK-OV-3 is a human ovarian carcinoma cell line with receptors for androgen, progesterone, and estrogen, which are downregulated or mutated (2). OVCAR-3 is another human ovarian carcinoma cell line in which androgen and estrogen receptors are present only in the nucleus, and progesterone receptors are distributed in both the nucleus and cytoplasm, suggesting different functions depending on their intracellular distribution (3). MDA-MB-231, Hs578T, and SK-OV-3 cells are characterized by their mesenchymal-like phenotype and exhibit high migratory and invasive behaviors, while OVCAR-3 cells exhibit low levels of migratory and invasive phenotypes (4, 5).

2. Transfection

To knock down target genes, siRNAs were transfected using Lipofectamine RNAiMAX Reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. For

each transfection reaction, 6 µL RNAiMAX and 10 nM siRNA were separately diluted in 250 µL of Opti-MEM medium (Gibco-BRL, Grand Island, NY, USA) and mixed at room temperature for 5 min. The transfection mixture was then added to each well of the 6-well plate. Co-transfections of purified plasmids (1 µg) and 10 nM siRNAs were performed by incubation with 6 µL of Lipofectamine 3000 (Invitrogen), according to the manufacturer's instructions. All transfections were independently performed at least thrice per target gene. The siRNA oligonucleotides used in the present study are listed in Table S3.

3. Cloning of HAS2 expression vector

Human HAS2 was PCR-amplified from cDNAs obtained from MDA-MB-231 cell lines using PfuUltra High-Fidelity DNA Polymerase (Agilent Technologies, Palo Alto, CA, USA). The primers used to amplify HAS2 were as follows: the *Hin*dIII site-linked primer (5′-GCTAAGCTTATGCATTGTGAGAGGTTTC-3′) was used as a forward primer, and the *Eco*RI site-linked primer (5′-CATGAATTCTCATACATCAAGCACCATGTC-3′) was used as a reverse primer. PCR-amplified HAS2 was digested with *Hin*dIII and *Eco*RI and cloned into the corresponding restriction enzyme sites of 3×FLAG-CMV-10 (Sigma-Aldrich, #E7658).

4. Cell proliferation assay

MDA-MB-231 and SK-OV-3 cells were seeded at 1.5×10^3 cells/well in 96-well plates. OVCAR-3 cells were seeded at 5×10^3 cells/well. After 24 h, the cells were transfected with negative control siRNA or KPNA3 siRNA using Lipofectamine RNAiMAX Reagent (Invitrogen) and incubated (0, 24, 48, and 72 h). Media were removed and replaced with 100 µL of complete media containing 10 µL Premix WST-1 (Takara Bio, Tokyo, Japan). After 4 h

of incubation, the absorbance was measured at 450 nm, with 650 nm as the reference wavelength using an Epoch Microplate Spectrophotometer (BioTek Instruments, Winooski, VT, USA). Three parallel wells were used per group, and all the experiments were conducted in triplicate.

5. Flow cytometric analysis

For cell cycle analysis, 1×10^6 cells were harvested by trypsinization after transfection with siRNA for 24 h and washed with phosphate-buffered saline (PBS). The cells were fixed in 70% ethanol at 4°C overnight and washed with PBS. Fixed cells were incubated with 50 µL of RNase A (1 mg/mL) at 37°C for 1 h and stained with 10 µL of propidium iodine solution (PI, 1 mg/mL). Then, PI-stained cells were analyzed using a Navios flow cytometer with Kaluza software (Beckman Coulter, Brea, CA, USA).

6. Quantitative real-time PCR (RT-qPCR)

Total RNA was isolated from cultured cells using RNAiso Plus Reagent (Takara Bio), according to the manufacturer's protocol. In brief, 1 µg of total RNA was used to perform first-strand cDNA synthesis with M-MLV Reverse Transcriptase (Promega, Madison, WI, USA). RT-qPCR was performed in triplicate for specific transcripts using TB Green™ Premix Ex Taq™ II (Takara Bio) on the QuantStudio 3 Real-Time PCR System with Design and Analysis software (Applied Biosystems, Foster City, CA, USA). The reaction condition consisted of 30 s at 95°C and 50 cycles of 5 s at 95°C, 30 s at 60°C, and 30 s at 72°C. Relative mRNA expression levels were measured using the comparative C_t ($\Delta\Delta C_t$) method, with normalization to the endogenous reference gene *GAPDH*. The primer sequences used in the present study are

listed in Table S4.

7. Western blot assay

At the indicated time points after transfection with either siRNAs or expression vectors, the cells were lysed in RIPA buffer (50 mM Tris, 1% NP40, 0.1% sodium dodecyl sulfate [SDS], 0.5% SDC, 150 mM NaCl, 1 mM EDTA) containing protease inhibitor (Roche, Indianapolis, IN, USA; #11873580001) and phosphatase inhibitor (Roche; #04906845001). The protein concentration in the cell extracts was determined using the Bio-Rad Protein Assay Kit (Bio-Rad, Hercules, CA, USA) with bovine serum albumin (BSA) as the standard. Proteins were loaded into the wells of an SDS-polyacrylamide gel and transferred to polyvinylidene fluoride (PVDF) membranes (Millipore, Billerica, MA, USA). The membranes were blocked with 5% non-fat dry milk (Bioworld, Visalia, CA, USA) in 1× Tris-buffered saline with 0.1% Tween 20 (TBST) for 1 h and incubated overnight at 4°C with primary antibodies against the target protein. Following incubation with horseradish peroxidase (HRP)-conjugated secondary antibody at room temperature for 1 h, immunoreactive bands were visualized using enhanced chemiluminescence (ECL) substrate. The intensity of the protein bands was analyzed using the ChemiDoc MP Imaging System equipped with Image Lab™ 6.0.0 software (Bio-Rad).

8. Antibodies and reagents

Antibodies against KPNA3 (Abcam, Cambridge, MA, USA; #ab6038), β-actin (Sigma-Aldrich; #A5316), E-cadherin (Cell Signaling Technology, Beverly, MA, USA; #3195), N-cadherin (Cell Signaling Technology; #13116), vimentin (Cell Signaling Technology; #5741), ZEB1 (GeneTex; #GTX105278), ZEB2 (Santa Cruz Biotechnology; #sc-271984), SNAI1 (Cell

Signaling Technology; #3879), SNAI2 (Cell Signaling Technology; #9585), TWIST1/2 (GeneTex; #GTX127310), GATA3 (Cell Signaling Technology; #5852), HAS2 (Santa Cruz Biotechnology; #sc-514737), FLAG (Sigma-Aldrich; #F3165), total SMAD2/3 (Santa Cruz Biotechnology; #sc-133098), phosphorylated SMAD2 (S465/467)/SMAD3 (S423/425) (Cell Signaling Technology.; # 8828), total AKT (Cell Signaling Technology; #9272), phosphorylated AKT (T308) (Abcam; #ab8933), phosphorylated AKT (S473) (Cell Signaling Technology; #4060), total ERK (Cell Signaling Technology; #4695), phosphorylated ERK (T202/Y204) (Cell Signaling Technology; #4370), PARP-1 (Santa Cruz Biotechnology; #sc-7150), caspase-3 (Santa Cruz Biotechnology; #sc-7272), cleaved caspase-3 (Cell Signaling Technology; #9664), caspase-8/cleaved caspase-8 (Cell Signaling Technology; #9746), and caspase-9/cleaved caspase-9 (Cell Signaling Technology; #9502) were used for western blot analysis. TGF-β1 (R&D Systems, Minneapolis, MN, USA, #P01137) and LY294002 (Cell Signaling Technology; #9901) were used as extrinsic factors to trigger the TGF-β1 signaling pathway and inhibit AKT phosphorylation, respectively.

9. Transswell chamber assay

Transwell migration and invasion assays were performed using an 8.0 µM pore size polycarbonate membrane (Corning, Inc., Corning, NY, USA) and Matrigel (Corning, Inc.; #354234)-coated inserts, respectively. The cell suspension (1×10^5 cells/200 µL serum-free medium) after transfection with siRNAs or expression vectors was added to the upper compartment, and complete medium was added to the lower compartment. After 9 h, the cells attached to the lower membrane were fixed with methanol and stained with 0.1% crystal violet, while the cells on the upper membrane were removed with cotton swabs. Migratory and invasive phenotypes of the cells were estimated by the average number of cells counted in five

random fields (magnification, 100×).

10. RNA-sequencing (RNA-seq)

The changes in transcriptional levels of target genes were estimated by RNA-seq. Total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, Valencia, CA, USA), according to the manufacturer's protocol. The quality of isolated RNA was assessed using the Agilent Technologies 2100 Bioanalyzer (Agilent Technologies). All RNA samples were given an RNA integrity number (RIN) ≥ 9.8. cDNA libraries were constructed using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. The pooled libraries were quantified using quantitative PCR, according to the qPCR Quantification Protocol Guide, and their quality was assessed using the Agilent Technologies 2100 Bioanalyzer with a DNA 1000 chip (Agilent Technologies). High-throughput sequencing was performed as 100 bp end sequencing on an Illumina NovaSeq 6000 (Illumina).

RNA-seq reads were aligned to the reference genome sequence (hg38, Genome Reference Consortium GRCh38) with gene annotation data from Ensembl (GRCh38 Release 104), and the raw count was calculated using the STAR aligner (version 2.7.9a) (6). Differential gene expression analysis was performed using the DESeq2 package (version 1.34.0) (7). Principal component analysis (PCA) was performed to identify the variability between samples and to test the biological reproducibility within replicates. To visualize the gene expression patterns, a heatmap was generated with normalized gene counts from DESeq2. Moreover, a volcano plot was constructed using the Bioconductor package EnhancedVolcano to visualize the results of differential expression analyses (8). Data analysis and visualization were performed using R 4.1.1 (www.r-project.org).

11. Public bioinformatic database analysis

To compare the relative expression levels of differentially expressed karyopherin-α isoforms (KPNAs) in TNBC (MDA-MB-231, Hs578T, and BT549) cells vs. non-TNBC (MCF7 and T47D) cells, the gene expression data of breast cancer cell lines from the Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) database under the series accession number GSE32474 were used. The gene expression data were analyzed using the GEO2R tool (https://www.ncbi.nlm.nih.gov/geo/geo2r/) to identify KPNAs differentially expressed between TNBC and non-TNBC cells. The gene expression and proteomic data of KPNA3 in breast cancer cell lines out of 60 diverse human cancer cell lines (NCI-60) were obtained from the CellMiner database (http://discover.nci.nih.gov/cellminer/) (9, 10). The expression patterns of KPNA3 in different breast cancer subtypes and tumor grades were analyzed using the GENT2 database (http://gent2.appex.kr/gent2/) (11). In addition, the effect of KPNA3 expression on survival rate was analyzed using GENT2. The protein expression pattern of KPNA3 in different breast cancer subtypes was analyzed using the UALCAN database (http://ualcan.path.uab.edu) (12). The correlation between the expression of KPNA3 and HAS2 was analyzed using the GEPIA2 database (http://gepia2.cancer-pku.cn/) (13). Enrichr (https://amp.pharm.mssm.edu/Enrichr) was used to examine the signaling pathways related to EMT-mediated metastasis and perturbed transcription factors associated with genes that were significantly differentially expressed in the RNA-seq (14-16). Putative binding sites for GATA3 in the promoter of HAS2 were predicted using the JASPAR bioinformatics database (https://jaspar.genereg.net/) (17).

12. Statistical analysis

In the present study, general statistical analyses were performed using two-tailed Student's t-tests in GraphPad Prism 8 (GraphPad Software, Inc., San Diego, CA, USA). Each experiment was independently repeated at least thrice. Data are expressed as the mean ± SD. Statistical significance was expressed as $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and $****p < 0.0001$. n.s. not significant.

References

1. Neve RM, Chin K, Fridlyand J et al (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10, 515-527

2. Lau KM, Mok SC and Ho SM (1999) Expression of human estrogen receptor-alpha and -beta, progesterone receptor, and androgen receptor mRNA in normal and malignant ovarian epithelial cells. Proc Natl Acad Sci U S A 96, 5722-5727

3. Lima MA, da Silva SV and Freitas VM (2016) Progesterone acts via the progesterone receptor to induce adamts proteases in ovarian cancer cells. J Ovarian Res 9, 9

4. Cailleau R, Olive M and Cruciger QV (1978) Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. In Vitro 14, 911-915

5. Choi JH, Choi KC, Auersperg N and Leung PC (2005) Gonadotropins upregulate the epidermal growth factor receptor through activation of mitogen-activated protein kinases and phosphatidyl-inositol-3-kinase in human ovarian surface epithelial cells. Endocr Relat Cancer 12, 407-421

6. Dobin A, Davis CA, Schlesinger F et al (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21

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

8. Blighe K RS, and Lewis M (2022) EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.

9. Shankavaram UT, Varma S, Kane D et al (2009) CellMiner: a relational database and query tool for the NCI-60 cancer cell lines. BMC Genomics 10, 277

10. Reinhold WC, Sunshine M, Liu H et al (2012) CellMiner: a web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set. Cancer Res 72, 3499-3511

11. Park SJ, Yoon BH, Kim SK and Kim SY (2019) GENT2: an updated gene expression database for normal and tumor tissues. BMC Med Genomics 12, 101

12. Chandrashekar DS, Bashel B, Balasubramanya SAH et al (2017) UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia 19, 649-658

13. Tang Z, Kang B, Li C, Chen T and Zhang Z (2019) GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res 47, W556-W560

14. Chen EY, Tan CM, Kou Y et al (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 14, 128

15. Kuleshov MV, Jones MR, Rouillard AD et al (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 44, W90-97

16. Xie Z, Bailey A, Kuleshov MV et al (2021) Gene Set Knowledge Discovery with Enrichr. Curr Protoc 1, e90

17. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I et al (2021) JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res