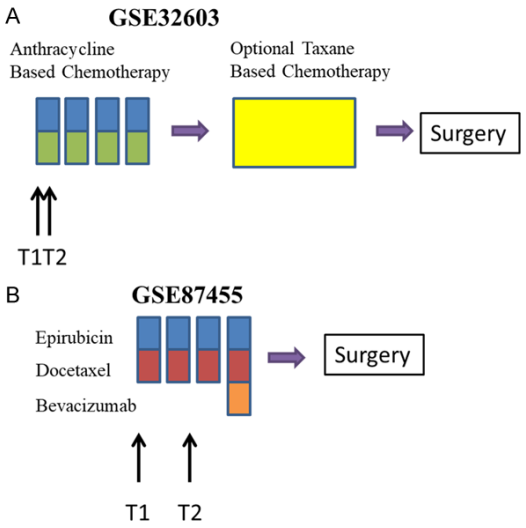

Figure S1. The explanation of NAC cohorts. A. GSE32603 was a cohort having patients treated with anthracycline based chemotherapy followed by optimal Taxane based chemotherapy. Core needle (16-gauge) biopsies were taken from the primary breast tumors before treatment (T1) and between 24 and 96 hours after the first dose (T2) of chemotherapy. Paired expression data for T1 vs. T2 was available for 36 patients [12]. B. GSE87455 was a cohort having patients treated with Epirubicin + Docetaxel + Bevacizumab. The mRNA quality and yield was adequate for the generation of high quality gene expression data from 122/145 samples (85%) at baseline (T1) and 82/138 (59%) after Cycle 2 (T2). There were paired baseline-Cycle 2 data for 69 patients [13].

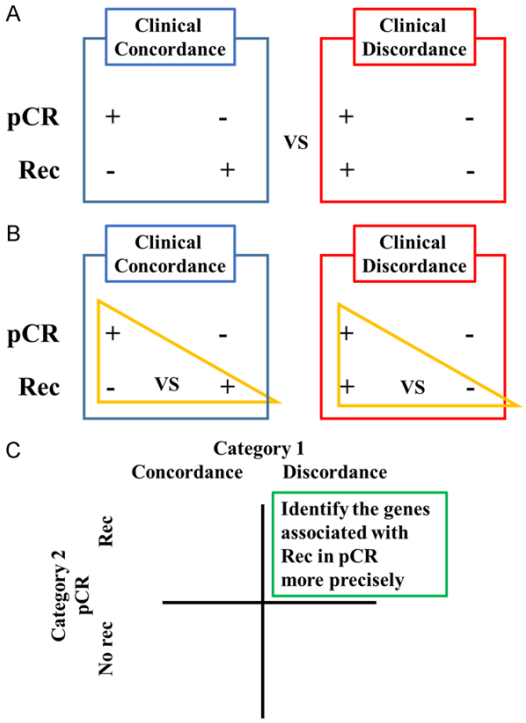

Figure S2. The definition of Category 1 and 2. A. Category 1, we divided BC patients into two groups with clinical integrity, clinical concordance versus discordance. Clinical concordance was defined as a group of patients who were pCR/No Rec and non-pCR/Rec, according to NAC. On the other hand, clinical discordance was defined as a group of patients who were pCR/Rec and non-pCR/No Rec, according to NAC. B. Category 2, we divided NAC treated BC patients according to whether they relapsed in the group having pCR. C. Association between Category 1 and 2. The relationship between clinical discordance and Rec in pCR can identify the genes associated with Rec in pCR more precisely. Abbreviations: pCR, pathological CR; Rec, recurrence; NAC, neoadjuvant chemotherapy; BC, breast cancer.