

diagnosis × developer: deployment vs evaluation
Row-normalized shares; filled markers: $q \leq 0.10$ (Fisher + BH-FDR)

