Q-Learning vs iterative NE with GA