Best-Pop Validation Accuracy Across Seeds (Mean ± Std)