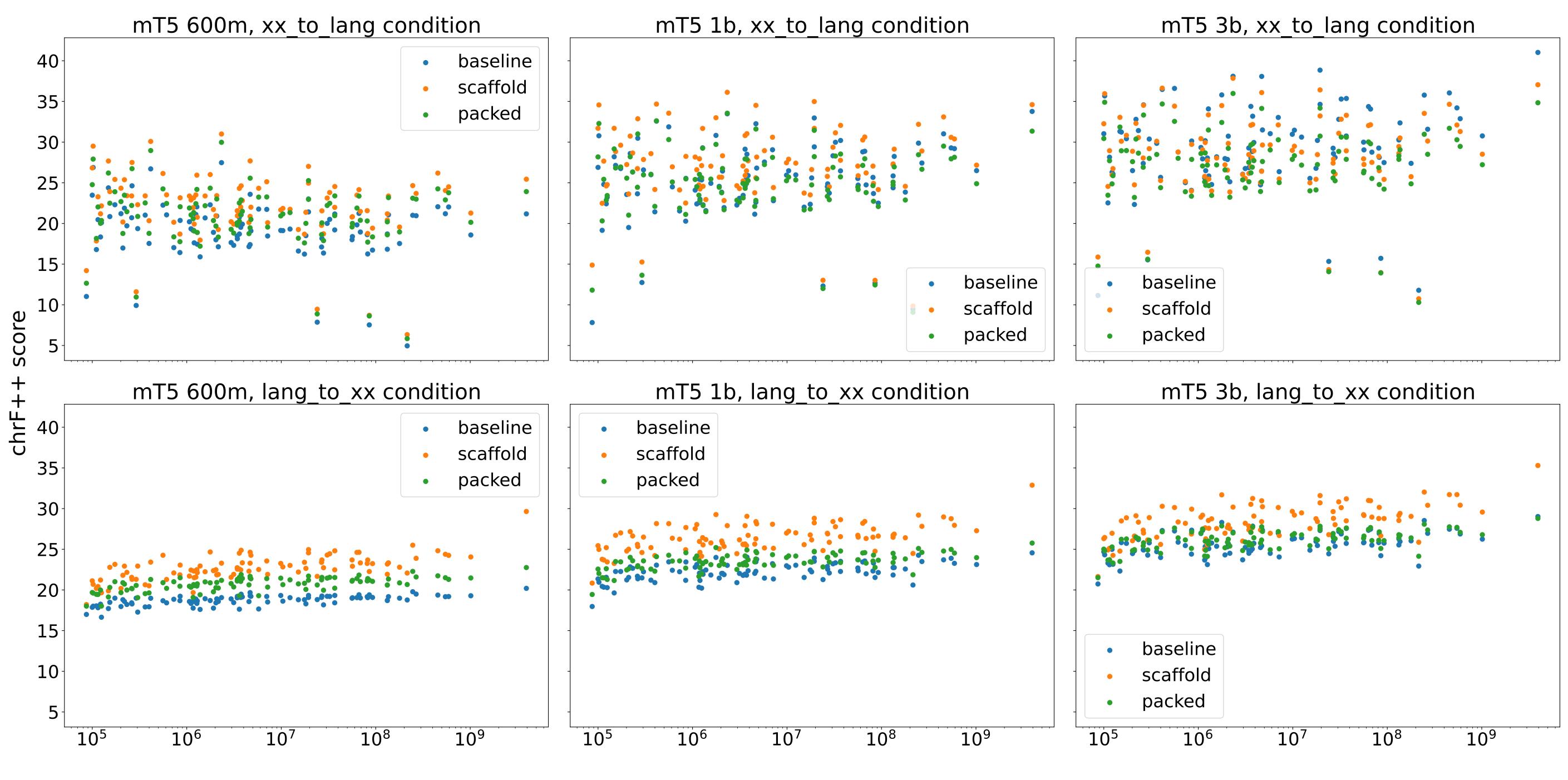Pretrain dataset size vs chrF++ score