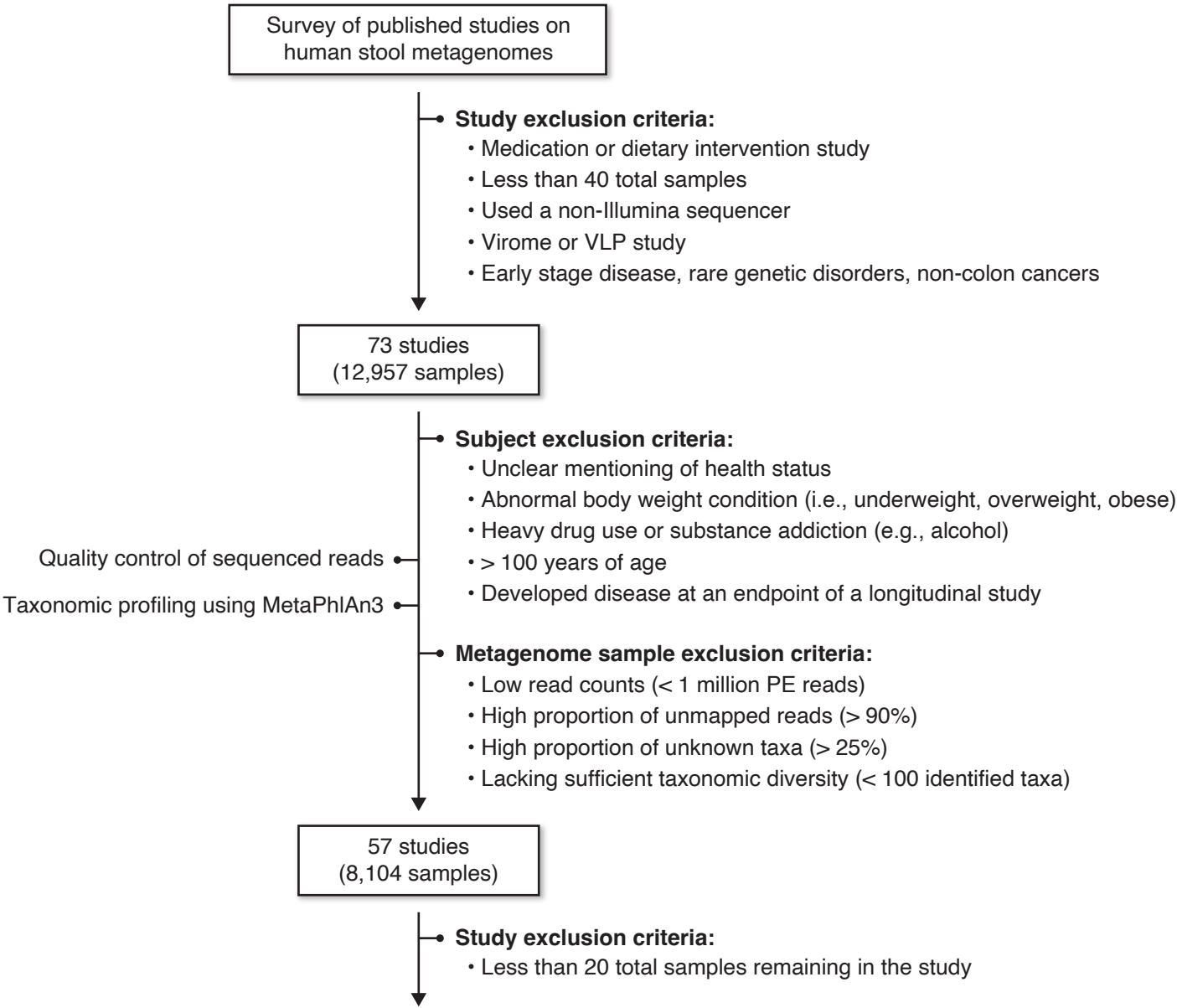


Figure 1

a



Pooled dataset of 8,069 stool shotgun metagenomes across 12 phenotypes and 54 studies

#	Phenotype	# of studies	# of samples
1	Healthy	46	5547
2	Ankylosing Spondylitis (AS)	1	95
3	Atherosclerotic Cardiovascular Disease (ACVD)	1	214
4	Colorectal Cancer (CC)	9	789
5	Crohn's Disease (CD)	7	284
6	Graves' Disease (GD)	1	100
7	Liver Cirrhosis (LC)	1	152
8	Multiple Sclerosis (MS)	1	224
9	Nonalcoholic Fatty Liver Disease (NAFLD)	1	86
10	Rheumatoid Arthritis (RA)	2	151
11	Type 2 Diabetes (T2D)	4	377
12	Ulcerative Colitis (UC)	6	250

b

