

**Model Performance - Precision**  
**Phenotype**

Model		Phenotype									
		Spore Form.	Motility	Biosafety	Animal Path.	Plant Path.	Host Assoc.	Extreme Env.	Cell Shape	Gram Stain	Biofilm
gemini-2.5-pro	0.89	0.83	0.82	0.75	0.68	0.73	0.69	0.68	0.66	0.52	
deepseek-r1	0.89	0.84	0.86	0.75	0.70	0.68	0.69	0.68	0.62	0.55	
claude-3.5-sonnet	0.89	0.85	0.87	0.76	0.72	0.70	0.69	0.67	0.62	0.48	
grok-3-mini	0.90	0.84	0.86	0.78	0.74	0.67	0.68	0.67	0.62	0.48	
gpt-5	0.90	0.86	0.83	0.74	0.68	0.75	0.67	0.69	0.62	0.48	
gpt-4o	0.89	0.82	0.83	0.78	0.70	0.77	0.63	0.67	0.61	0.53	
gemini-pro-1.5	0.92	0.72	0.84	0.77	0.75	0.72	0.60	0.67	0.61	0.54	
gpt-oss-120b	0.91	0.82	0.73	0.71	0.68	0.76	0.68	0.68	0.62	0.56	
o1-mini	0.90	0.82	0.79	0.73	0.70	0.72	0.65	0.67	0.63	0.49	
claude-sonnet-4	0.90	0.83	0.81	0.78	0.71	0.65	0.67	0.67	0.62	0.47	
gemini-2.0-flash-lite-001	0.91	0.75	0.78	0.75	0.71	0.73	0.63	0.68	0.62	0.51	
gpt-4.1-nano	0.76	0.79	0.83	0.71	0.66	0.74	0.64	0.87	0.55	0.52	
gpt-4	0.88	0.80	0.74	0.71	0.70	0.75	0.68	0.60	0.61	0.59	
gpt-5-nano	0.86	0.81	0.80	0.68	0.70	0.74	0.65	0.67	0.64	0.48	
deepseek-chat-v3-0324	0.86	0.82	0.81	0.73	0.71	0.66	0.65	0.66	0.62	0.51	
gpt-oss-20b	0.87	0.83	0.79	0.70	0.69	0.74	0.63	0.64	0.63	0.52	
llama-3-70b-instruct:nitro	0.87	0.76	0.78	0.79	0.72	0.63	0.59	0.67	0.61	0.47	
gemini-flash-1.5	0.90	0.78	0.78	0.81	0.74	0.60	0.60	0.59	0.60	0.48	
wizardlm-2-8x22b	0.84	0.80	0.72	0.76	0.71	0.65	0.65	0.66	0.61	0.47	
claude-3-haiku:beta	0.87	0.78	0.78	0.80	0.70	0.74	0.50	0.63	0.61	0.47	
gemini-pro	0.85	0.77	0.84	0.79	0.64	0.69	0.57	0.62	0.59	0.51	
gemini-2.5-flash-lite	0.87	0.81	0.70	0.68	0.71	0.51	0.65	0.66	0.61	0.52	
gpt-3.5-turbo-0125	0.86	0.75	0.63	0.71	0.69	0.66	0.60	0.65	0.61	0.47	
palm-2-chat-bison-32k	0.82	0.72	0.73	0.75	0.66	0.72	0.58	0.55	0.58	0.50	
mixtral-8x7b-instruct:nitro	0.87	0.74	0.55	0.66	0.67	0.56	0.59	0.54	0.58	0.47	
mistral-7b-instruct	0.82	0.66	0.67	0.71	0.65	0.62	0.55	0.46	0.46	0.58	
openchat-7b	0.72	0.67	0.49	0.70	0.64	0.62	0.54	0.47	0.65	0.56	
llama-3-8b-instruct:nitro	0.82	0.73	0.51	0.74	0.66	0.57	0.54	0.45	0.51	0.47	
llama-3-sonar-small-32k-chat	0.84	0.61	0.58	0.72	0.63	0.47	0.60	0.38	0.55	0.47	
phi-3-mini-128k-instruct	0.69	0.66	0.67	0.56	0.59	0.55	0.54	0.47	0.49	0.46	
gemma-7b-it	0.64	0.50	0.51	0.66	0.63	0.55	0.47	0.27	0.40	0.47	

