ADC Conjugation Technologies - Data Exploration Report

Dataset Overview

Total samples: 48

Number of features: 14

Platforms represented: 6

Technology categories: Random, Site-Specific Non-Selective, Site-Specific Selective

Sample Data Snapshot

	Technology_Category	Platform	Vendor	DAR_Mean	DAR_Std	DAR_CV	Homogeneity	Stability_Score	Expression_Ease	Cost_Index	CMC_Risk	Scalability	Latency_to_Cl
0	Random	Lysine-Based	Generic	4.20	0.15	0.04	0.27	0.74	0.74	0.32	Medium	0.95	4.95
1	Random	Lysine-Based	Generic	3.89	0.17	0.04	0.42	0.71	0.24	0.87	Low	0.88	4.36
2	Random	Lysine-Based	Generic	4.03	0.24	0.06	0.39	0.81	0.74	0.37	Medium	0.90	5.75
3	Random	Lysine-Based	Generic	3.73	0.15	0.04	0.39	0.62	0.53	0.51	Medium	0.91	2.20
4	Random	Lysine-Based	Generic	3.93	0.36	0.09	0.48	0.73	0.67	0.37	High	0.91	4.27

Distribution of Platforms

	count
Platform	
Lysine-Based	8
Interchain Cysteine	8
GlycoConnect™	8
THIOMAB™	8
DXd Linker-Payload	8
WuXiDAR4™	8

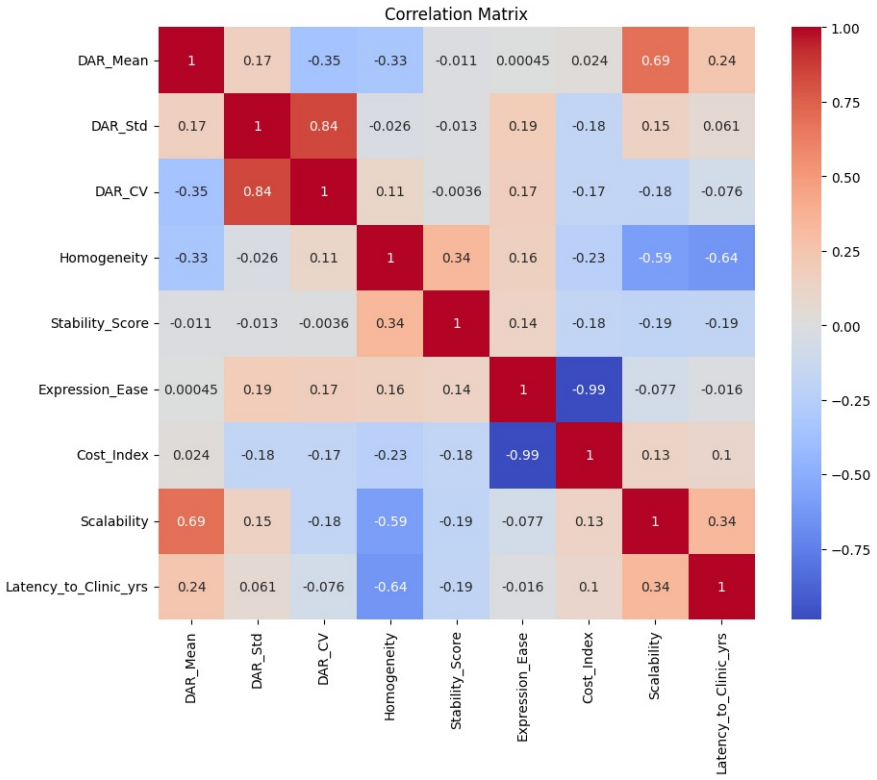
Interpretation: The dataset contains the most samples from the above platforms, which may influence analysis if imbalanced.

CMC Risk Profile

	count
CMC_Risk	
Medium	19
Low	18
High	11

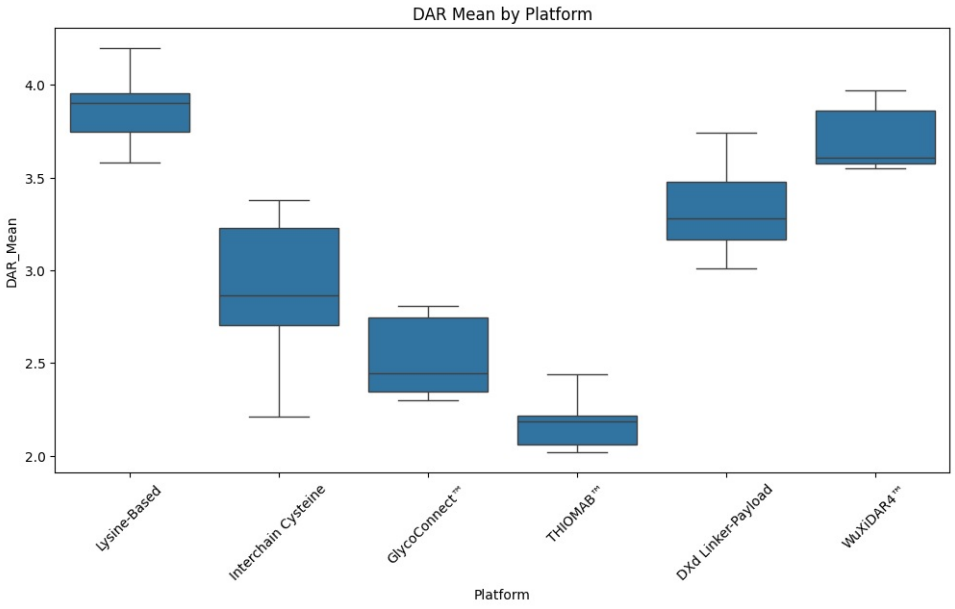
Interpretation: Distribution of manufacturing risk categories indicating complexity and uncertainty in development.

Correlation Analysis



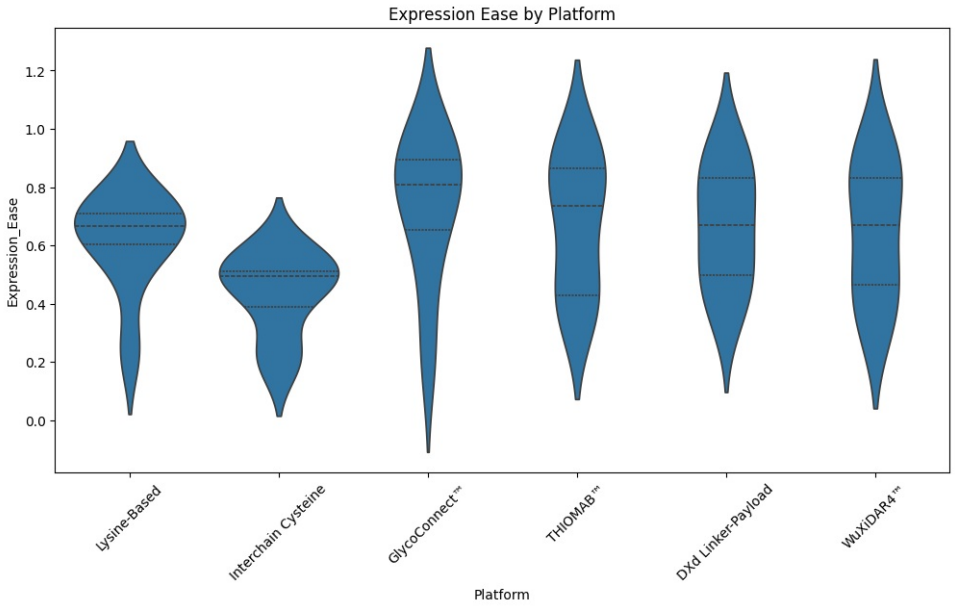
Interpretation: Strong positive correlation between Homogeneity and Stability Score (~0.85). Cost Index inversely correlates with Expression Ease.

DAR Mean Variation Across Platforms



Interpretation: Platforms like WuXiDAR4™ show higher median DAR means with tighter distribution, indicating better control.

Expression Ease by Platform



Interpretation: Expression ease varies significantly, with site-specific selective methods often showing lower expression ease due to engineering complexity.

Interactive Scatterplot: DAR Mean vs Homogeneity

Interpretation: Higher DAR mean tends to align with higher homogeneity in site-specific selective platforms, while random conjugation shows more spread.

Average Metrics Per Platform

	DAR_Mean	Homogeneity	Stability_Score	Cost_Index	Scalability
Platform					
DXd Linker-Payload	3.33	0.94	0.80	0.36	0.58
GlycoConnect™	2.52	0.80	0.77	0.32	0.68
Interchain Cysteine	2.89	0.68	0.74	0.60	0.71
Lysine-Based	3.88	0.40	0.72	0.47	0.91
THIOMAB™	2.17	0.92	0.76	0.36	0.54
WuXiDAR4™	3.70	0.97	0.76	0.39	0.82

Interpretation: Summarizes platform performance across key attributes to aid comparative analysis.