

An Exploration of Public Genomics Companies

Jessica Parks | 2021-07-20

Given the ongoing Genomics Revolution and an expectation that the leading innovators in the genomics space will undergo substantial growth over the coming years, how can we best deploy our investment capital to participate in this growth opportunity? How might we identify and optimize a set of companies in which to invest?

In the following text I'll illustrate how we could use content published by these companies (specifically their website content) and a few of their fundamental financial values to select a subset from all the genomics companies to prioritize for further investment assessment.

DATA: *There are 972 genomics companies used for all analyses described; they were selected by industry (limited to "Agricultural Inputs", "Biotechnology", "Chemicals", "Diagnostics & Research", "Drug Manufacturers - General", "Farm Products", "Health Information Services", and "Medical Devices"), stock exchange ("NASDAQ", "NYSE"), and location (excluding "China" and "Hong Kong"). All data is sourced from the QUANDL's Sharadar Core US Fundamentals Data via the QUANDL API.*

CODE: *Code for this analysis is available at github.com/jessicaparks/genomics-stocks*

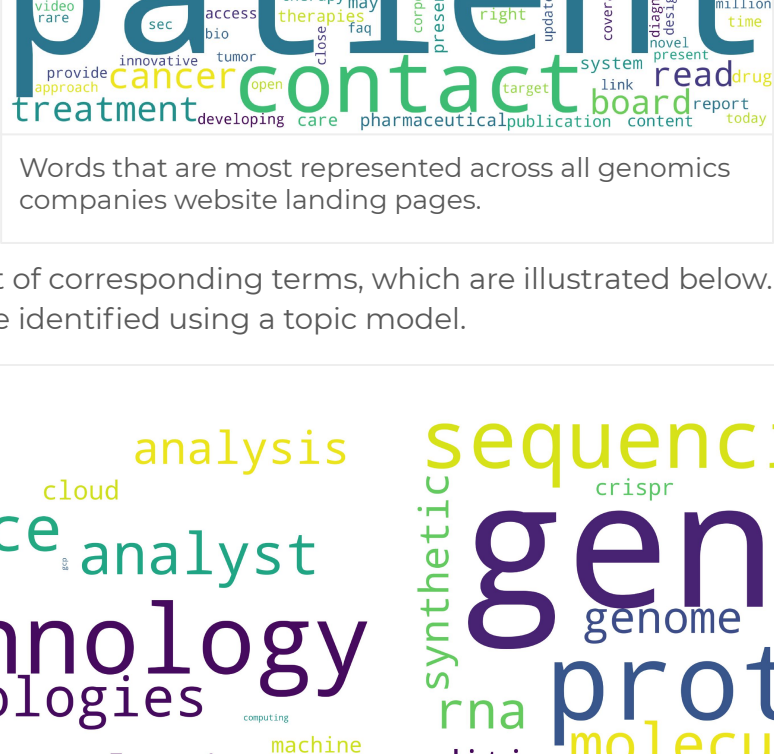
Part I: Company Website Content

Today, a company's online content is their messaging to the public -- to customers, partners and investors. As an outsider to the company, we interpret this data to form an impression of the company and inform our decision about doing business with them or investing in them.

How might we use the **content published by a company** as an **indicator** of their ranking as an **innovator in the genomics field**?

As a simple example, we'll use the text from the landing page of each company's website and analyze the proportion of words associated with the relevant topics: "genomics", "innovation", "science", and "data".

An in-depth approach could use data combined from a range of sources, including: other pages from the company's website; company publications such as white papers, journal articles, and patents; company content published across social media; and media or analyst content covering the company.



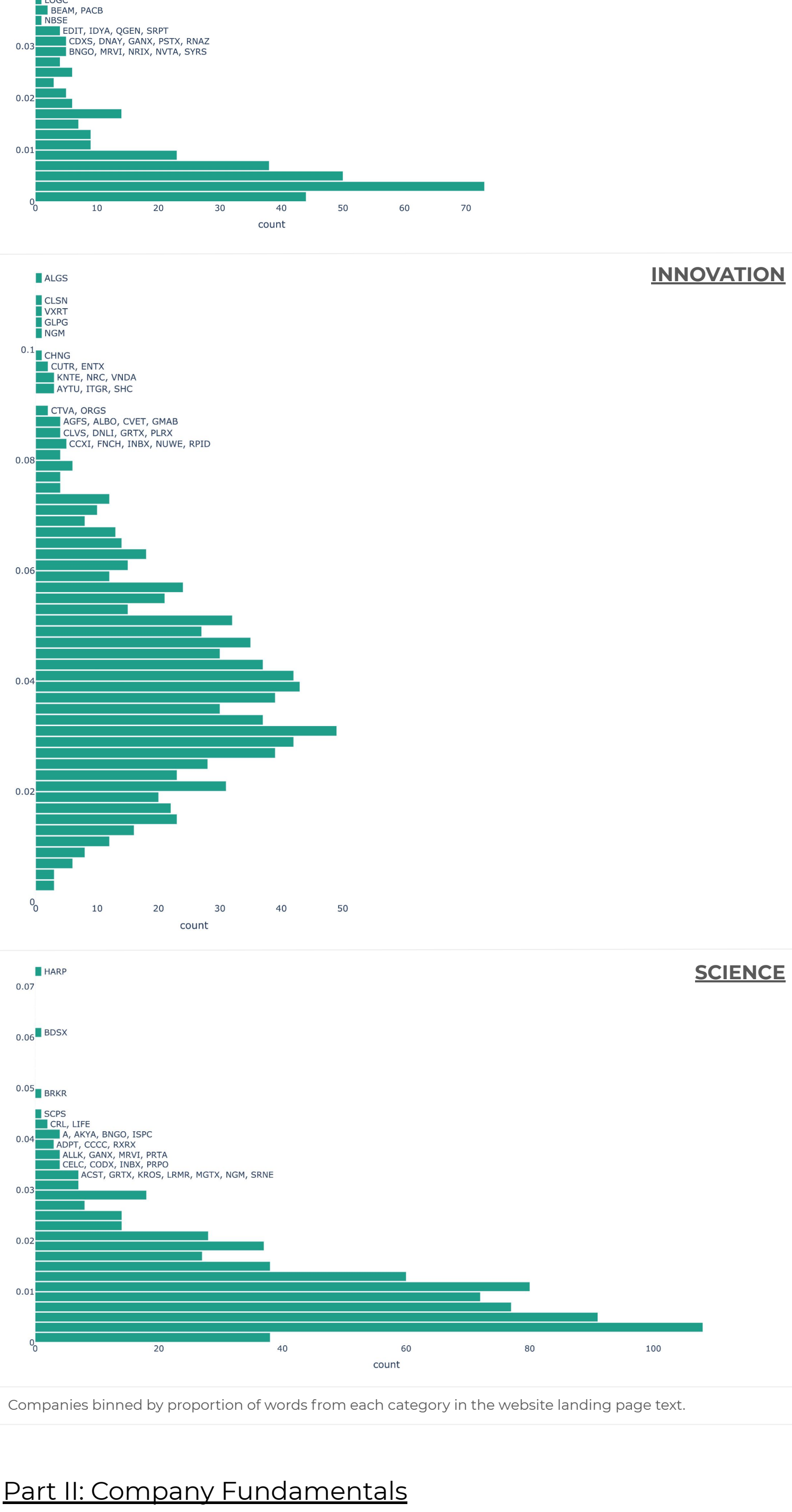
For each topic, I set a list of corresponding terms, which are illustrated below. Alternatively, the topics and terms could both be identified using a topic model.



In the following histogram figures, the companies are binned by the proportion of words in each category relative to all words present on the company's website landing page. The ~25 top companies are indicated with their ticker symbols.

Note that the top companies for "Genomics" include DNA sequencing, mapping, and synthesis providers like ILMN, TWST, PACB, and BNGO; as well as gene editing companies such as CRSP, NTLA, and EDIT and many others that are focused on the applications of editing technology. These are companies that are fundamental enablers and players in the Genomics Revolution.

There is some overlap between the top companies across the categories, although no company is found in the top representatives for all four categories. Those overlapping include: BNGO, ENTX, GANX, GRTX, INBX, ITGR, MRVI, NGM, and RXRX.

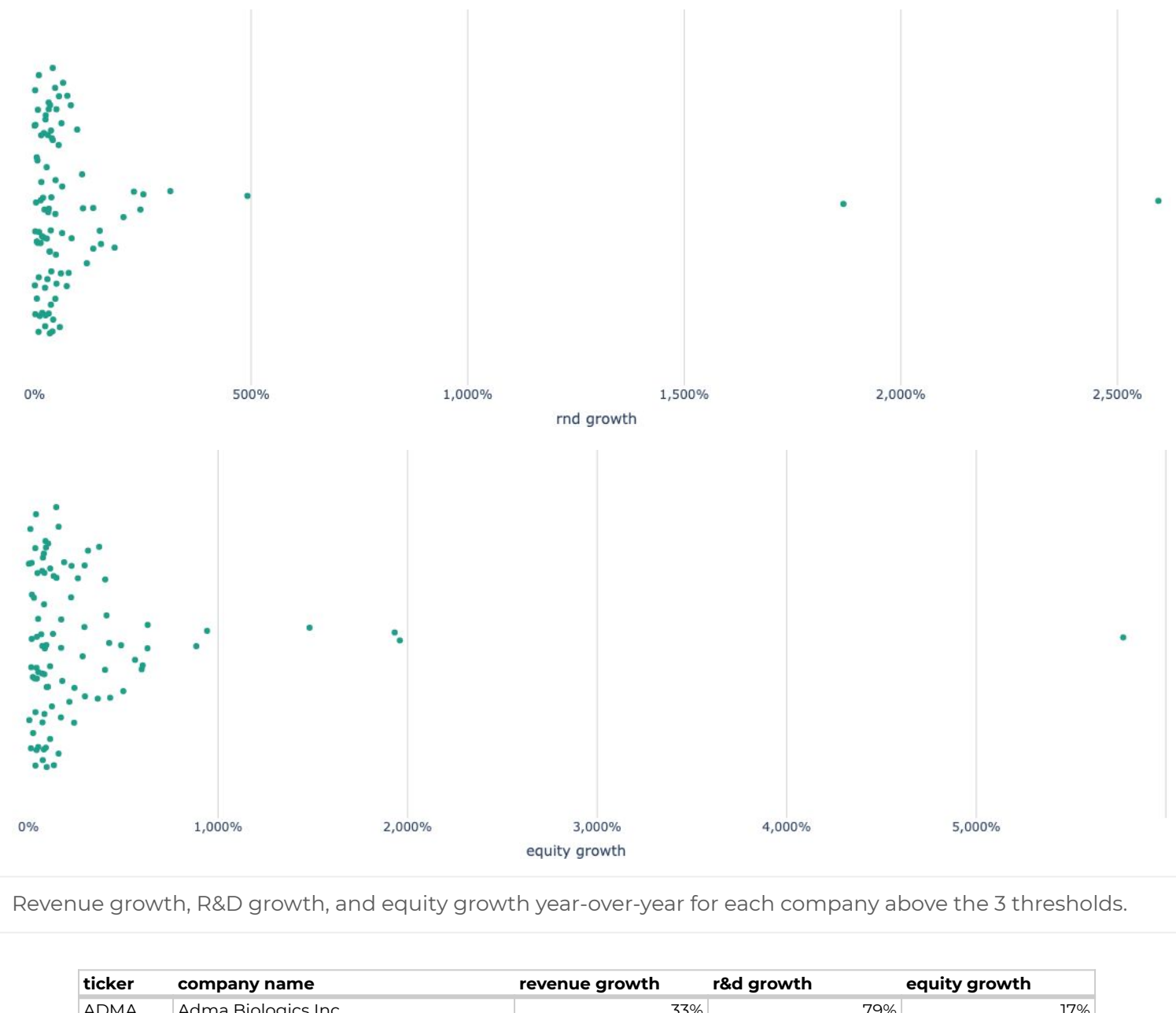


Part II: Company Fundamentals

Like with company website content, we'll consider how we might use the **company fundamentals** as an **indicator** of their ranking as an **innovator in the genomics field**.

Let's assume that a high level of innovation will translate to a high level of company growth. As a proxy for company growth, we'll focus on year-over-year growth in three values: revenue, research and development expense, and equity. As a baseline, we'll include only companies with revenue growth of 25% or more, and with any level of growth in R&D expense and equity. This equates to 93 of the 972 companies. The values for these 93 companies for revenue growth, R&D growth, and equity growth are shown in the strip plots below. Note that the majority of the companies have values under 100% growth compared to the prior year. For reference, these values for the 93 companies are shown in the following table.

Note: RXRX, SEER, VERV, and ZY are examples of companies that did not have enough data for the fundamentals growth calculation, given that they IPO'd within the last year. We could do a more complete analysis given pre-IPO fundamental data or projected fundamental data.



ticker	company name	revenue growth	r&d growth	equity growth
ADMA	Adma Biologics Inc	33%	79%	17%
ADVM	Adverum Biotechnologies Inc	2900%	74%	36%
AFMD	Affimed NV	33%	14%	86%
ALPN	Alpine Immune Sciences Inc	304%	8%	114%
ALT	Altimmune Inc	35%	151%	488%
AQB	AquaBounty Technologies Inc	104%	1%	562%
ARAV	Aravive Inc	95%	48%	16%
ATEC	Alphatec Holdings Inc	34%	42%	295%
AUTL	Autolus Therapeutics plc	28%	18%	4%
AXNX	Axonics Inc	206%	39%	91%
AYTU	Aytu Biopharma Inc	294%	492%	67%
BCYC	Bicycle Therapeutics plc	30%	30%	74%
BGNE	BeiGene Ltd	114%	24%	42%
BIOC	Biocept Inc	636%	3%	18%
BNTX	BioNTech SE	344%	185%	178%
CASI	CASI Pharmaceuticals Inc	132%	40%	42%
CCXI	ChemoCentryx Inc	105%	10%	500%
CDNA	CareDx Inc	59%	56%	365%
CLGN	CollPlant Biotechnologies Ltd	758%	10%	597%
CDTG	Centogene NV	163%	56%	10%
CODX	Co-Diagnostics Inc	5185%	251%	298%
CRDF	Cardiff Oncology Inc	48%	5%	1932%
CSBR	Champions Oncology Inc	26%	31%	148%
DMTK	DermTech Inc	58%	136%	285%
DNLI	Denali Therapeutics Inc	1204%	7%	103%
DOCS	Doximity Inc	78%	35%	5775%
DRIO	DarioHealth Corp	36%	49%	628%
DXCM	Dexcom Inc	27%	38%	104%
EDIT	Editas Medicine Inc	278%	43%	172%
EXAS	Exact Sciences Corp	46%	313%	44%
EYPT	EyePoint Pharmaceuticals Inc	33%	10%	602%
FATE	Fate Therapeutics Inc	279%	42%	260%
FLGT	Fulgent Genetics Inc	2115%	112%	885%
FULC	Fulcrum Therapeutics Inc	1615%	20%	81%
GMAB	Genmab A	91%	28%	40%
GRTS	Gritstone bio Inc	892%	2%	73%
HARP	Harpoon Therapeutics Inc	189%	26%	23%
HOOK	Hookipa Pharma Inc	58%	33%	29%
HZNP	Horizon Therapeutics Public Ltd Co	59%	121%	76%
IMMP	IMMUTEP Ltd	126%	23%	37%
IMTX	Immatics NV	45%	76%	1959%
INMD	InMode Ltd	39%	5%	51%
INO	Inovio Pharmaceuticals Inc	147%	38%	216%
INSP	Inspire Medical Systems Inc	51%	84%	75%
ITMR	Itamar Medical Ltd	34%	33%	158%
JAGX	Jaguar Health Inc	93%	21%	412%
KDNY	Chinook Therapeutics Inc	224%	1867%	404%
LEXX	Lexaria Bioscience Corp	294%	86%	225%
LGND	Ligand Pharmaceuticals Inc	90%	16%	13%
LPTX	Leap Therapeutics Inc	300%	2%	114%
MDGS	Medigus Ltd	95%	64%	242%
MCNX	Macrogenics Inc	58%	0%	84%
MORF	Morphic Holding Inc	158%	18%	188%
MRKR	Marker Therapeutics Inc	119%	51%	84%
MRNA	Moderna Inc	5135%	244%	134%
MRTX	Mirati Therapeutics Inc	457%	51%	92%
MTBC	CareCloud Inc	59%	205%	81%
NTRA	Natera Inc	36%	110%	44%
NVTA	Invitae Corp	33%	48%	629%
ONCT	Oncternal Therapeutics Inc	40%	26%	942%
OPGN	Opgen Inc	43%	153%	227%
PHR	Phreesia Inc	26%	33%	404%
PROF	Profound Medical Corp	67%	15%	24%
PTGX	Protagonist Therapeutics Inc	1245%	12%	314%
QDEL	Quidel Corp	232%	66%	146%
QGEN	Qiagen Nv	33%	7%	14%
QTNT	Quotient Ltd	33%	1%	425%
QTRX	Quantex Corp	63%	36%	296%
RCUS	Arcus Biosciences Inc	468%	135%	373%
RGLS	Regulus Therapeutics Inc	16567%	64%	82%
RGNX	REGENXBIO Inc	200%	24%	32%
SDGR	Schrodinger Inc	25%	62%	114%
SELB	Selecta Biosciences Inc	315%	5%	123%
SGEN	Seagen Inc	138%	14%	94%
SLRX	Sallarius Pharmaceuticals Inc	36%	41%	171%
SMTI	Sanara MedTech Inc	33%	2595%	431%
SRNE	Sorrento Therapeutics Inc	41%	31%	129%
SYRS	Syros Pharmaceuticals Inc	349%	35%	96%
TBIO	Translate Bio Inc	1436%	61%	173%
TDCC	Teladoc Health Inc	126%	229%	1483%
TELA	TELA Bio Inc	28%	48%	72%
TNDM	Tandem Diabetes Care Inc	37%	35%	51%
TRVN	Trevena Inc	10474%	2%	241%
TTNP	Titan Pharmaceuticals Inc	50%	9%	89%
TWST	Twist Bioscience Corp	72%	28%	132%
TGX	10x Genomics Inc	26%	48%	84%
VAPO	Vapotherm Inc	153%	37%	159%
VEEV	Veeva Systems Inc	31%	39%	36%
VIR	Vir Biotechnology Inc	616%	98%	71%
VIVO	Meridian Bioscience Inc	60%	25%	51%
VMD	Viemed Healthcare Inc	58%	58%	2%
XENE	Xenon Pharmaceuticals Inc	112%	33%	47%
XLRN	Accelaron Pharma Inc	49%	22%	96%

Part III: Combining the Results

Let's synthesize the results from the analyses in Parts I and II to create a short-list of companies that rank highly in both website content and in fundamentals. Specifically, we'll consider only those companies above the median value in the "Genomics" and "Innovation" categories for website content and those that pass our filters for growth in revenue, R&D expense, and equity.

Our resulting companies are: AOB, ATEC, AUTL, AXNX, AYTU, BCYC, BGNE, BNTX, CCXI, CDNA, CODX, MTBC, DNLI, EDIT, EXAS, EYPT, FLGT, GMAB, GRTS, HZNP, IMTX, INO, LGND, LPTX, MGNX, MRNA, MRKR, ONCT, OPGN, PTGX, QTNT, RCUS, RGNX, SDGR, SELB, SGEN, SYRS, TBIO, TDCC, TELA, TGX, TXG, VEEV, and VIR. The underlined tickers show those currently included in the ARKG fund, providing some evidence that this type of approach has the potential to identify companies of interest. The essential next step to qualify and select the best companies for investment would be a deep-dive analysis into each company in the list to review their technology and financials and to assess their current and potential impact in the Genomics Revolution.