JCI insight

Safety and efficacy of the JAK inhibitor tofacitinib citrate in patients with alopecia areata

Milène Kennedy Crispin, … , Anthony E. Oro, Brett A. King

Clinical Medicine   Dermatology   Inflammation

BACKGROUND. Alopecia areata (AA) is an autoimmune disease characterized by hair loss mediated by CD8[+] T cells. There are no reliably effective therapies for AA. Based on recent developments in the understanding of the pathomechanism of AA, JAK inhibitors appear to be a therapeutic option; however, their efficacy for the treatment of AA has not been systematically examined.

METHODS. This was a 2-center, open-label, single-arm trial using the pan-JAK inhibitor, tofacitinib citrate, for AA with >50% scalp hair loss, alopecia totalis (AT), and alopecia universalis (AU). Tofacitinib (5 mg) was given twice daily for 3 months. Endpoints included regrowth of scalp hair, as assessed by the severity of alopecia tool (SALT), duration of hair growth after completion of therapy, and disease transcriptome.

RESULTS. Of 66 subjects treated, 32% experienced 50% or greater improvement in SALT score. AA and ophiasis subtypes were more responsive than AT and AU subtypes. Shorter duration of disease and histological peribulbar inflammation on pretreatment scalp biopsies were associated with improvement in SALT score. Drug cessation resulted in disease relapse in 8.5 weeks. Adverse events were limited to grade I and II infections. An AA responsiveness to […]

Find the latest version:

https://jci.me/89776/pdf

JCI insight

Safety and efficacy of the JAK inhibitor tofacitinib citrate in patients with alopecia areata

Milène Kennedy Crispin,[1] Justin M. Ko,[1] Brittany G. Craiglow,[2,3] Shufeng Li,[1] Gautam Shankar,[1] Jennifer R. Urban,[4] James C. Chen,[5,6] Jane E. Cerise,[5] Ali Jabbari,[5] Mårten C.G. Winge,[1] M. Peter Marinkovich,[1] Angela M. Christiano,[5,7] Anthony E. Oro,[1] and Brett A. King[2]

[1]Program in Epithelial Biology and Department of Dermatology, Stanford University School of Medicine, Stanford, California, USA. [2]Department of Dermatology, [3]Department of Pediatrics, and [4]Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut, USA. [5]Department of Dermatology, [6]Department of Systems Biology, and [7]Department of Genetics and Development, Columbia University, New York, New York, USA.
[8] Dermatology Service, Veterans Affairs Medical Center, Palo Alto, California, USA.

BACKGROUND. Alopecia areata (AA) is an autoimmune disease characterized by hair loss mediated by CD8[+] T cells. There are no reliably effective therapies for AA. Based on recent developments in the understanding of the pathomechanism of AA, JAK inhibitors appear to be a therapeutic option; however, their efficacy for the treatment of AA has not been systematically examined.

METHODS. This was a 2-center, open-label, single-arm trial using the pan-JAK inhibitor, tofacitinib citrate, for AA with >50% scalp hair loss, alopecia totalis (AT), and alopecia universalis (AU). Tofacitinib (5 mg) was given twice daily for 3 months. Endpoints included regrowth of scalp hair, as assessed by the severity of alopecia tool (SALT), duration of hair growth after completion of therapy, and disease transcriptome.

RESULTS. Of 66 subjects treated, 32% experienced 50% or greater improvement in SALT score. AA and ophiasis subtypes were more responsive than AT and AU subtypes. Shorter duration of disease and histological peribulbar inflammation on pretreatment scalp biopsies were associated with improvement in SALT score. Drug cessation resulted in disease relapse in 8.5 weeks. Adverse events were limited to grade I and II infections. An AA responsiveness to JAK/STAT inhibitors score was developed to segregate responders and nonresponders, and the previously developed AA disease activity index score tracked response to treatment.

CONCLUSIONS. At the dose and duration studied, tofacitinib is a safe and effective treatment for severe AA, though it does not result in a durable response. Transcriptome changes reveal unexpected molecular complexity within the disease.

TRIAL REGISTRATION. ClinicalTrials.gov NCT02197455 and NCT02312882.

FUNDING. This work was supported by the US Department of Veterans Affairs Office of Research and Development, National Institute of Arthritis and Musculoskeletal and Skin Diseases National Institutes of Health grant R01 AR47223 and U01 AR67173, the National Psoriasis Foundation, the Swedish Society of Medicine, the Fernström Foundation, the Locks of Love Foundation, the National Alopecia Areata Foundation, and the Ranjini and Ajay Poddar Resource Fund for Dermatologic Diseases Research.

Introduction

Alopecia areata (AA) is an autoimmune disease with a lifetime prevalence of 2% (1). Loss of immune privilege leads to hair cycle growth arrest of the anagen hair, resulting in nonscarring hair loss. AA is characterized by small patches of hair loss within normal hair-bearing skin. The variants of AA include alopecia totalis (AT), characterized by complete loss of scalp hair; alopecia universalis (AU), characterized by complete loss of scalp and other body hair; and ophiasis pattern AA, characterized by hair loss

JCI iNSIGHT

Enrollment

Enrolled (n=70)

Allocation

Received medication (n=68)
- Underwent scalp biopsy (n=28)
Did not receive allocated intervention (n=2)
- Both patients were unable to attend study visits

Follow Up

Discontinued intervention (n=2)
- 1 patient withdrew after 1 month due to inability to attend study visits
- 1 patient withdrew after 11 weeks due to esophageal reflux thought to be unrelated to drug

Analysis

Analyzed (n=66)
Excluded from analysis (n=0)
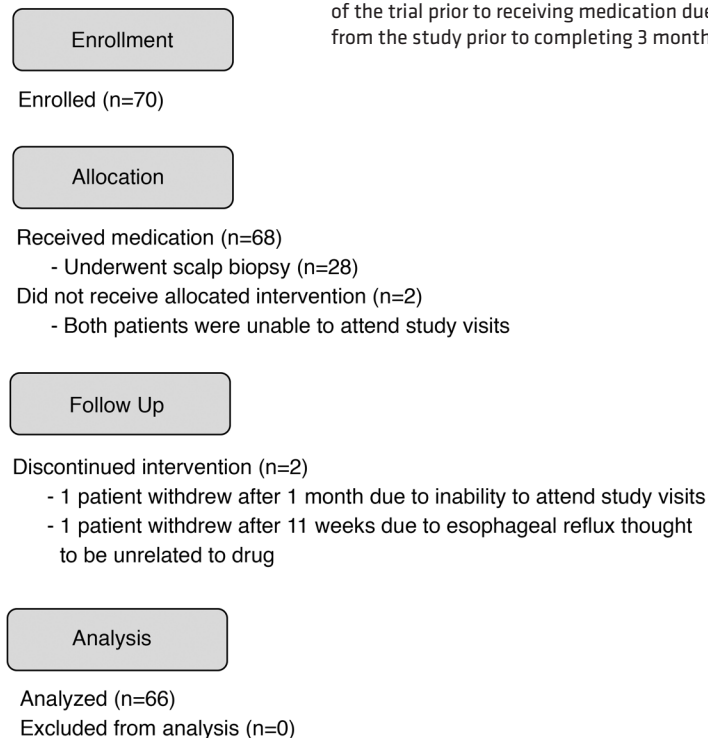
localized to the temporal and occipital scalp.

Treatment of AA is challenging, and there is no FDA-approved therapy. While a variety of therapies are commonly used, none of them are reliably effective, especially for the treatment of severe AA, AT, and AU (2).

Two decades of experimental research using animal models of AA and genome-wide association studies both in mice and humans have elucidated key factors in AA pathogenesis (3, 4). An important therapeutic insight was the discovery that blockade of common signaling pathways downstream of cytokine receptors, in particular JAK/STAT, could reverse AA in mice (5). Subsequently, treatment of 1 patient with the pan-JAK inhibitor, tofacitinib citrate (6); 3 patients with the JAK1/2 inhibitor, ruxolitinib (5); and 1 patient with JAK1/2 inhibitor, baricitinib (7), induced inflammatory remission and hair regrowth. Accordingly, a 2-center, open-label trial was performed to assess the safety and efficacy of tofacitinib in patients with severe AA, AT, and AU.

Results

Participants. Of the 70 patients enrolled, 66 completed the trial (Figure 1). Two patients withdrew prior to receiving tofacitinib, one patient withdrew after 1 month due to inability to attend study visits, and another patient withdrew after 11 weeks due to esophageal reflux thought to be unrelated to the drug. There were 31 (47%) men and 35 (53%) women, with a median age of 37 years (range 19–65 years). The median duration of the current episode of AA was 5 years (range 0.5–43 years). Different clinical subtypes were represented, including 11 patients with AA (16.7%), 3 with ophiasis (4.6%), 6 with AT (7.6%), and 46 with AU (71.2%). Scalp biopsies were performed in 28 patients, 23 (82%) of which exhibited follicular peribulbar inflammation prior to treatment. Several patients had concomitant autoimmune comorbidities, including atopic dermatitis, allergic rhinitis, thyroid disease, asthma, vitiligo, and psoriasis (Table 1).

Clinical response and associations. After 3 months of treatment, the median percentage change in the severity of alopecia tool (SALT) score was 21%. Treatment response was divided into 3 groups: nonresponders (NRs) had a percentage change in SALT score improvement of less than 5% (36% of patients), intermediate responders (IRs) had a percentage change in SALT score improvement of between 5% and 50% (32% of patients), and strong responders (SRs) had a percentage change in SALT score improvement of greater than 50% (32% of patients) (Figure 2A).

The relationship between demographics and clinical characteristics with percentage change in SALT score after 3 months was assessed. There was no correlation between age and percentage change in SALT score ($r = 0.03918$, $P = 0.7548$). The median improvement for men was 23.7% compared with 13.9% in women, but this was not found to be statistically significant in either the univariate ($P = 0.2256$) or multivariate ($P = 0.0990$) analyses (Figure 2B and Table 2). The subtype affected the median percentage change in SALT: AA (70.0%), ophiasis (68.0%), AT (11.8%), and AU (10.5%) ($P = 0.0024$) (Figure 2C). In the multivariate analysis, patients with AA had a 34% greater percentage change in SALT score compared with patients with AU ($P = 0.0005$), and patients with ophiasis had a 48% greater percentage change in SALT score compared with patients with AU ($P = 0.006$) (Table 2). There was a significant negative correlation between duration of current episode of AA and percentage change in SALT score (-0.35559, $P = 0.0034$) (Figure 2D). In a multivariate model that included duration of current episode of AA (in years), gender, and AA subtype, each additional year of disease accounted for a decrease in percentage change in SALT score of 0.78 ($P = 0.0247$) (Table 2). Principal component analysis of genome-wide RNA sequencing of pretreatment biopsies demonstrated that SRs with more longstanding disease (SR >20 years) more closely

Table 1. Baseline characteristics of the patients

Age: median yr (range)	37 (19–65)
Sex: no. (%)	
Female	31 (47.0)
Male	35 (53.0)
Duration of current episode of AA: median yr (range)	5 (0.5–43)
Alopecia areata subtype: no. (%)	
Areata	11 (16.7)
Ophiasis	3 (4.6)
Totalis	6 (7.6)
Universalis	46 (71.2)
Peribulbar inflammation on initial scalp biopsy	
Brisk	23 (82.1)
Minimal	5 (17.9)
Autoimmune comorbidities: no. (%)	
Atopic dermatitis	15 (22.7)
Allergic rhinitis	12 (18.2)
Thyroid disease	9 (13.6)
Asthma	7 (10.6)
Vitiligo	6 (9.1)
Psoriasis	3 (4.6)

matched SRs with shorter duration disease (SR <10 years) than disease duration-matched NRs (NR >20 years), suggesting that the signaling status of the scalp, rather than the duration of disease, predicts response (Figure 2E).

Scalp biopsy specimens were examined in order to assess the association of disease response and preexisting inflammation. In patients with peribulbar inflammation (23 of 28), a median percentage change in SALT score of 32.9% (range, 1.2%–87.4%) was observed, while in patients without peribulbar inflammation (5 of 28), a median percentage change in SALT score of only 1.2% (range 0–65) ($P = 0.05$) was seen (Figure 2F). The results of a representative NR are shown in Figure 3A, wherein minimal peribulbar inflammation on pretreatment scalp biopsy (Figure 3A) is visible, and the patient did not demonstrate hair growth between 0 (Figure 3B) and 3 months (Figure 3C). In contrast, the results of a SR show marked peribulbar inflammation on pretreatment scalp biopsy (Figure 3D). The corresponding upregulation of phosphorylated STAT3 (pSTAT3) in hair follicles visible with indirect immunofluorescence (Figure 3E) was reduced after 2 months of tofacitinib (Figure 3F) in this patient, with an initial SALT score of 100 (Figure 3G) and a SALT score of 13 after 3 months (Figure 3H).

Molecular biomarker analysis. In all patients for whom biopsies were performed, the activity of tofacitinib was further confirmed using transcriptome analysis and the trivariate composite AA disease activity index (ALADIN) score. This score is derived from analyzing a composite of 16 genes that have been shown to characterize pathogenic IFN and cytotoxic T cell (CTL) responses and hair growth in ruxolitinib-treated AA (5); it effectively stratifies a large cohort of AA patients based on disease phenotype (8). In AU and AT patients, from baseline to week 8, the change in IFN scores was significant for NRs and SRs ($Z = 2.07$, $P = 0.039$ and $Z = 2.07$, $P = 0.039$, respectively) (Figure 4, purple genes), and the change in CTL scores approached significance within NRs and SRs ($Z = 1.84$, $P = 0.07$, and $Z = 1.836$, $P = 0.074$, respectively) (Figure 4, blue genes, and Supplemental Figure 1A; supplemental material available online with this article; doi:10.1172/jci.insight.89776DS1), consistent with JAK inhibition by tofacitinib.

In contrast to previous reports however (5), no significant difference in CTL or IFN scores was observed between NRs and SRs at baseline or at week 8 (Supplemental Figure 1, A–C), indicating that these markers do not distinguish responders from NRs at baseline in our study group. Despite reduction in inflammation (ALADIN, IFN and CTL signaling), the RNA sequencing data reveals an increased molecular complexity for hair growth not captured in the ALADIN score.

To generate a signature of early tofacitinib response, a naive, unsupervised hierarchical clustering (HCL) analysis was performed at 8 weeks of treatment using the AA gene signatures (AAGS) derived from previously published work, which are a validated representation of AA pathogenesis and severity (9). This disease signature accurately separated SRs and NRs at 8 weeks (Supplemental Figure 2), indicating that the course of response in this trial followed comparable molecular tracks that have been validated as characterizing AA pathology.

Performing the same analysis comparing pretreatment and 8-week scalp biopsy specimens also revealed distinct patterns of coclustering across time (Supplemental Figure 3). Differences between patient-paired samples revealed the existence of 2 subclusters of NRs. One of these subclusters had a molecular behavior similar to that in the SRs and the other had a molecular behavior discordant with response. This was used to partition the NRs into presumptive slow responders, who may eventually

Table 2. Multivariate analysis of clinical variables and percentage change in SALT score

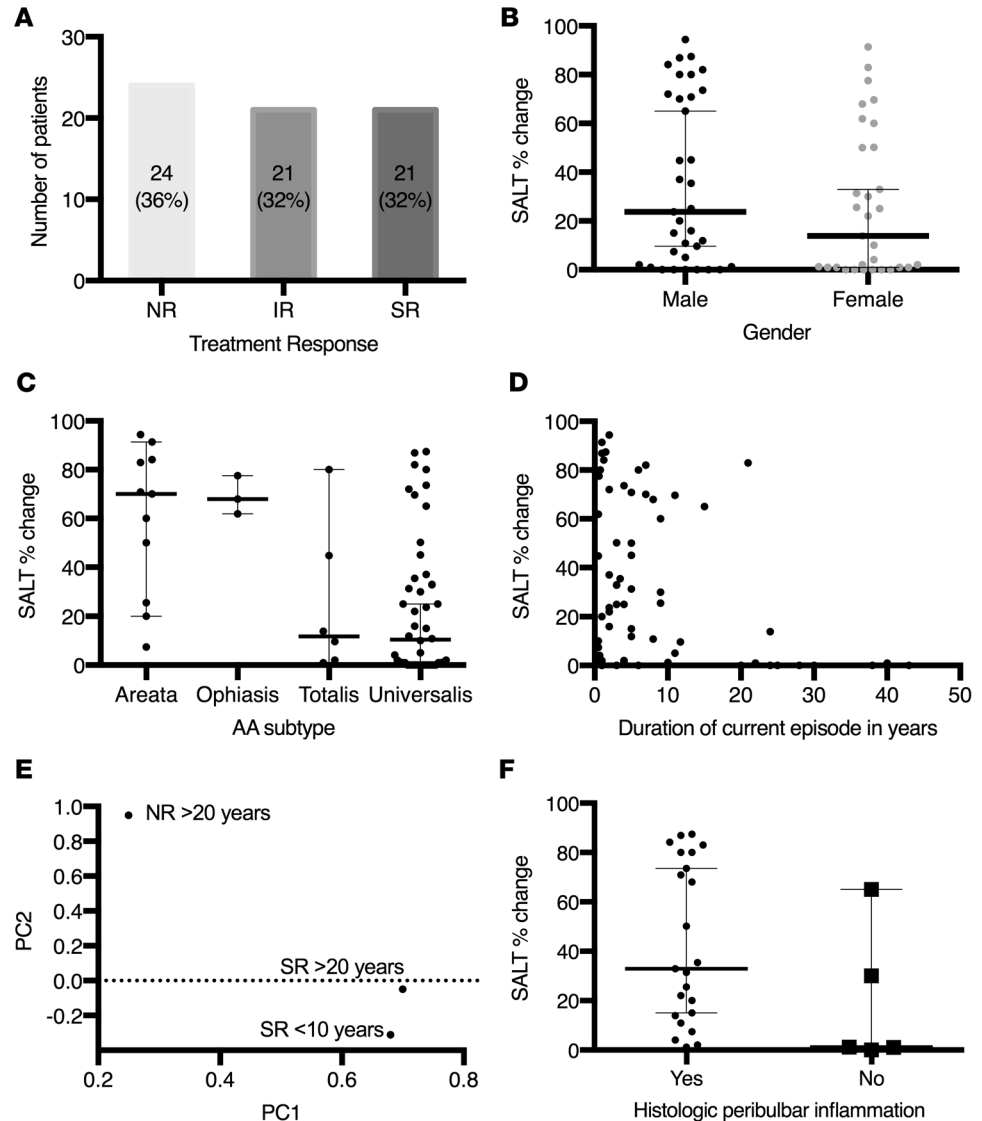| Clinical variables	Coefficient (95%CI)	*P* value
Duration of current episode of AA (yr)	–0.78 (–1.45 to –0.10)	0.0247
Gender: female vs. male	–11.70 (–25.66 to 2.26)	0.0990
AA subtype		
Areata vs. universalis	34.08 (15.72 to 52.44)	0.0005
Ophiasis vs. universalis	47.92 (14.25 to 81.59)	0.0060
Totalis vs. universalis	6.51 (–17.20 to 30.21)	0.6754

Figure 2. Clinical response and associations. (**A**) Distribution of percentage change in SALT score for all patients: 24 patients (36%) were nonresponders (NRs) with percentage change in SALT score below 5%, 21 patients (32%) were intermediate responders (IRs) with percentage change in SALT score between 5% and 50%, and 21 patients (32%) were strong responders (SRs) with percentage change in SALT score >50%. (**B**) The median percentage change in SALT score was 23.7% (range, 0%–94.4%) for men (*n* = 35) and 13.9% (range, 0%–91.4%) for women (*n* = 31), but the difference between groups was not significant (*P* = 0.2256) according to the Wilcoxon rank-sum test. (**C**) There was a difference in median percentage change in SALT score based on subtype: AA (70.0%, *n* = 11), ophiasis (68.0%, *n* = 3), AT (11.8%, *n* = 6), and AU (10.5%, *n* = 46) (*P* = 0.0024) based on the Wilcoxon rank-sum test. (**D**) There was a significant negative correlation, as measured by Spearman correlation test, between duration of current episode of AA and percentage change in SALT score (–0.35559, *P* = 0.0034). (**E**) Principal component analysis of genome-wide RNA sequencing demonstrated that the pretreatment signature of SRs with more longstanding disease (SR >20 years) more closely matched that of SRs with less longstanding disease (SR <10 years) rather than the disease duration-matched NRs (NR >20 years). (**F**) Peribulbar inflammation was seen on initial scalp biopsy in 23 of 28 patients and corresponded to a median percentage change in SALT score of 32.9% (range, 1.2%–87.4%) compared with 5 of 28 patients without peribulbar inflammation with median percentage change in SALT score of 1.2% (range, 0%–65%), which was not found to be significant (*P* = 0.0512) based on the Wilcoxon rank-sum test. In **B**, **C**, and **F**, lines denote median, and error bars denote 95% confidence interval.

Table 3. Adverse events

Events	n (%)
Infections	
URI	11 (16.7)
UTI	2 (3.0)
Zoster	1 (1.5)
Conjunctivitis	1 (1.5)
Bronchitis	1 (1.5)
Mononucleosis	1 (1.5)
Paronychia	1 (1.5)
Opportunistic infections	0
Tuberculosis	0
Total infections	17 (25.8)
Malignancy	
New malignancy	0
Recurrent malignancy	0
Other	
Headache	5 (7.6)
Abdominal pain	5 (7.6)
Acne	5 (7.6)
Diarrhea	4 (6.1)
Fatigue	4 (6.1)
Hot flashes	3 (4.5)
Pruritus	2 (3.0)
Folliculitis	2 (3.0)
Numbness	2 (3.0)
Cough	1 (1.5)
Nausea	1 (1.5)
Amenorrhea	1 (1.5)
Dry eyes	1 (1.5)
Weight gain	1 (1.5)

respond to treatment and are simply delayed, and likely true NRs, who exhibit no indication of improvement at a molecular level.

Since these classes were naively identified using the AAGS, it was possible to subsequently generate biomarker signatures that directly maximize the molecular separation between the classes, independent of the AAGS. The top 50 most differentially expressed genes, which we call the AA responsiveness to JAK/STAT inhibitors (AARSIN) genes, were selected as a biomarker panel for SRs, slow responders, and true NRs (Supplemental Table 1). Of the 50 genes encompassing this list, approximately 60% are coding and 40% noncoding. Included in the coding genes are *HHLA2*, encoding the CD28 costimulatory ligand B7-H7, while among the noncoding genes are *RAETK1*, a lncRNA expressed in the intergenic region expressing natural killer G2D (NKG2D) ligands; *ULBP1–6*; and *PSORS1C3*, a psoriasis-associated lncRNA expressed near *HLABC* and NKG2D ligand *MICA*.

The AARSIN genes were used to perform a cluster analysis across all patient biopsies at baseline to test whether or not this signature would accurately cluster all responders with the designated slow responders (Figure 5A). A positive result in this analysis indicated that the biomarker panel derived from the AAGS true NRs versus slow responders was sufficient to predict SR versus NR across the treatment cohort, as near-perfect clustering was achieved, with only 1 outlier patient from the SR cohort (Figure 5B). The AARSIN genes appear to reflect the molecular complexity of the tofacitinib-treated patients in this study.

Durability of response. To assess durability of hair regrowth after drug cessation, patients were followed for up to 3 months after discontinuation of tofacitinib. Twenty patients were available for evaluation, and all of them experienced hair loss (median 8.5 weeks) (Supplemental Figure 4).

Adverse events. In studies of tofacitinib for rheumatoid arthritis and psoriasis, transaminitis and infections were observed (8, 10). In this AA study, relatively few adverse effects were noted (Tables 3 and 4), and these were limited to grade I and II infections. Mild clinical infections that included paronychia and upper respiratory tract infection were reported in 25% of patients. One patient with a history of atopic dermatitis and herpes zoster experienced uncomplicated dermatomal zoster after 2 months of tofacitinib. No patients required hospitalization. A single patient with nonalcoholic steatohepatitis experienced a mild elevation of AST and ALT during treatment that resolved subsequent to the clinical trial after weight loss (and continuing to take tofacitinib).

Discussion

This study is the first and largest clinical trial to evaluate the efficacy of the pan-JAK inhibitor, tofacitinib, for the treatment of severe AA and variants, AT and AU. The results suggest that tofacitinib is a safe and effective treatment for severe AA, AT, and AU. In this study, 64% of patients responded to treatment, and 32% of patients achieved an improvement in SALT score of 50% or greater in only 3 months of therapy. Tofacitinib was well tolerated, with grade I and grade II adverse events only. This is corroborated by large studies of tofacitinib for rheumatoid arthritis, in which 4% of active-treatment patients discontinued treatment due to adverse events compared with 3% of patients in the placebo groups (10, 11–13). Ruxolitinib, another JAK inhibitor, is associated with infection, bone marrow suppression, and transaminitis (14). While tofacitinib was efficacious, all responders began losing hair again approximately 2 months after stopping the medication, indicating the recurrence of inflammatory stimuli following removal of JAK inhibitor treatment.

The effect of tofacitinib on the JAK/STAT pathway in AA is evidenced by the downregulation of pSTAT3 in hair follicles during treatment. A recent study of normal murine cycling hair demonstrated that JAK signaling induces a telogen block, preventing the hair from entering anagen phase and that JAK inhibitors induce anagen hair growth similar to a hedgehog pathway agonist (15). Another study demonstrated upregulation of pSTAT1 and pSTAT3 signaling in hair follicles in AA but not normal hair follicles due to IFN-γ signaling (5), supporting the idea that immune dysregulation enforces a similar hair cycle arrest.

Unlike in the murine studies, the data herein show that inhibition of the JAK/STAT pathway by tofacitinib is not entirely predictive of short-term hair growth in AA and reveal a more complex molecular pathogenesis. While shorter duration of current episode of disease is positively associated with response, patients with duration of current episode of disease up to 24 years exhibited a molecular signature similar to those with shorter duration disease. The ALADIN score, previously validated in
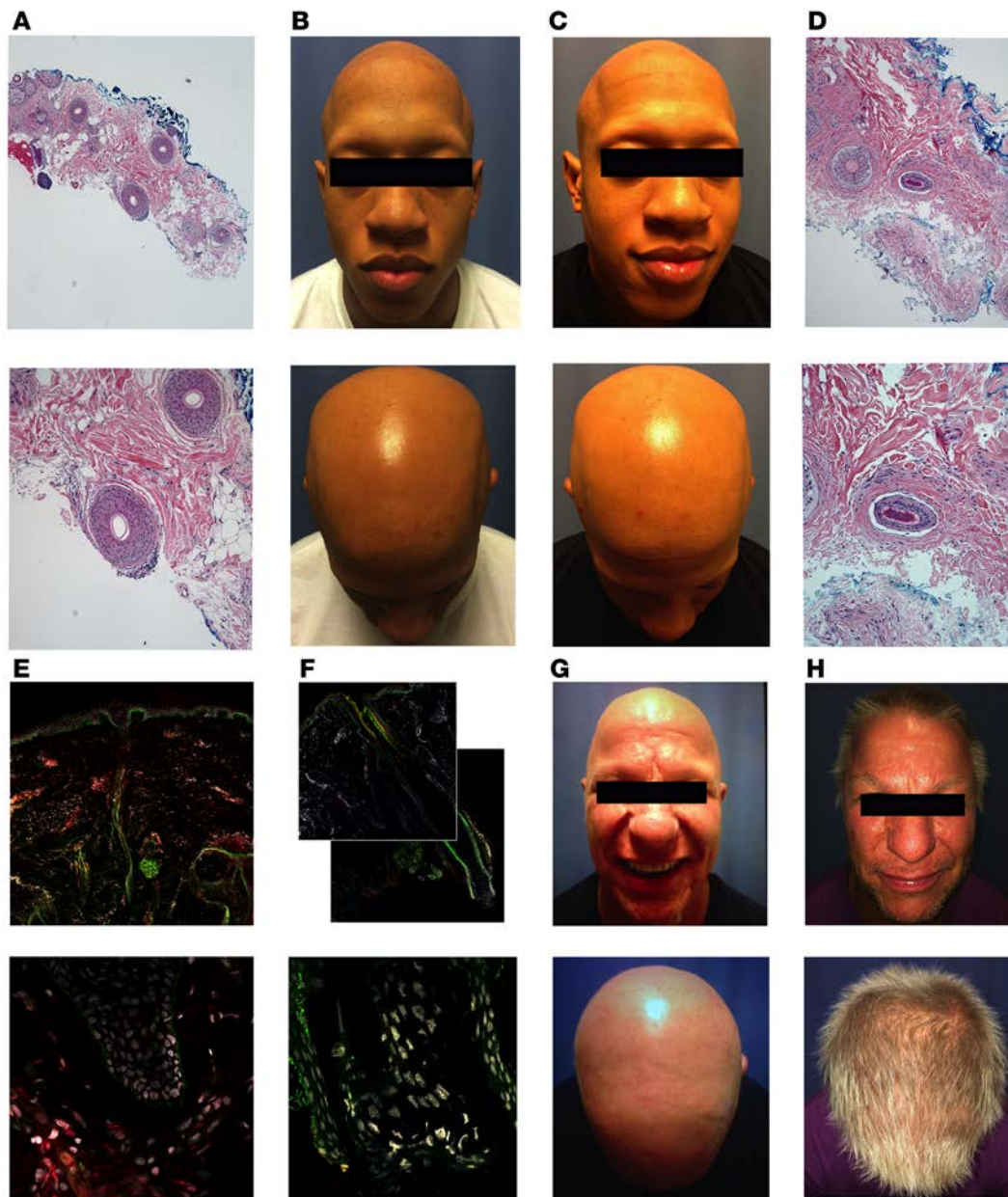
Figure 3. Association of disease response and inflammation. (**A**) A representative nonresponder had minimal peribulbar inflammation on initial scalp biopsy (original magnification, ×4 [top], ×10 [bottom]). Histopathology with hematoxylin and eosin stain was performed for all 28 patients who consented to scalp biopsy submission for histopathology. (**B**) Representative nonresponder with SALT score of 100 prior to treatment and (**C**) 100 at 3 months. (**D**) A representative strong responder displayed brisk peribulbar lymphocytic infiltrate on initial scalp biopsy (original magnification, ×10 [top], ×20 [bottom]), and (**E**) immunofluorescence of pretreatment scalp biopsy demonstrated strong follicular staining for STAT3 (red), (**F**) which was diminished after 2 months of tofacitinib (original magnification ×10 [top], ×63 [bottom]). Immunofluorescence for STAT3 was performed on 3 randomly selected strong responders and 3 randomly selected nonresponders. (**G**) The representative strong responder had a SALT score that decreased from 100 prior to treatment (**H**) to 13 at 3 months. Immunofluorescence: red = pSTAT3, green = β4-integrin, gray = DNA.

ruxolitinib-treated patients (5) and in Mackay-Wiggan et al. (16), adequately predicted patient response to ruxolitinib, and it was not able to stratify SRs from NRs to tofacitinib. This is not surprising, as the target specificity of the drugs differs and because there was a predominance of AT and AU patients in this study, rather than patch-stage AA in the ruxolitinib trial, or, alternatively, because our study analyzed patients at 2 months after treatment with tofacitinib, which may not represent a complete response or reversion of gene expression signatures.

The AARSIN gene list developed in this study stratifies SRs, slow responders, and true NRs to tofacitinib and suggests a molecular scoring system for predicting responsiveness to treatment with
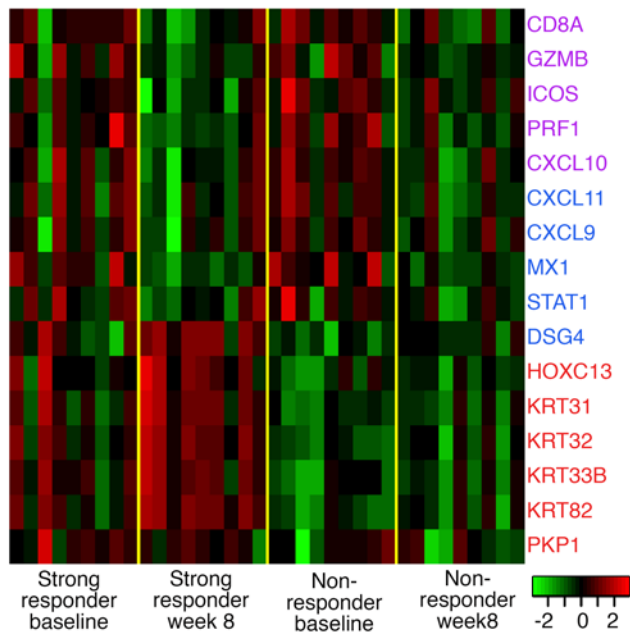
Figure 4. Heatmap of alopecia areata disease activity index gene *Z* scores in scalp biopsies of strong responders (*n* = 9) at baseline and week 8 and nonresponders (*n* = 9) at baseline and week 8.

this drug. Preliminary data indicate that the AARSIN gene list may also stratify ruxolitinib-treated patients, although the small number of patients prevents achievement of statistical significance (Mackay-Wiggan et al., ref. 16). The identification of a slow responder group suggests that longer treatment duration or more potent JAK inhibitors may be beneficial for some patients who do respond clinically early in treatment. Indeed, 2 patients in the slow responder group who continued the medication for an additional 3 months experienced percentage change in SALT score of 50%–90% (J.M. Ko and M.K. Crispin, unpublished observations).

In summary, tofacitinib is safe and effective for the treatment of severe AA, AT, and AU, though it does not give a durable response at the dose and duration studied. JAK inhibition represents a targeted therapy for this disease. Additional trials are needed to validate the AARSIN and ALADIN scores and to elucidate dosing regimens that may provide durable responses.

Methods

Study design. Patients were enrolled at 2 centers between July 2014 and March 2015. Patients received 5 mg tofacitinib twice daily for 3 months. Optional scalp biopsies were performed at one center at the time of enrollment and after 2 months of treatment and were analyzed for histology and RNA sequencing.

A dose escalation model was initially planned, whereby patients would receive 5 mg every other day; if no hair regrowth occurred after 1 month, the dose would be increased to 5 mg daily. If no regrowth occurred at this dose after 1 month, the dose would be increased to 5 mg twice daily for a total of 3 months. Due to a lack of hair regrowth observed at 5 mg every other day and 5 mg daily, all patients were treated with 5 mg twice daily after the first 13 patients completed treatment with the dose escalation model.

The primary endpoint was the percentage change in scalp hair growth, measured with SALT (17). Secondary endpoints included the effect on treatment response of patient age, sex, duration of current episode of AA, AA subtype, and the presence of peribulbar inflammation on pretreatment scalp biopsy as well as duration of hair growth following drug cessation and scalp biopsy transcriptome analysis.

Study participants. Patients 18 years of age or older with a clinical diagnosis of AA with at least 50% hair loss involving the scalp, AT, or AU were eligible for the study if hair loss had been present for at least 6 months without evidence of hair regrowth. Patients were recruited to the study by self-referral or referral from their dermatologists. Patients must not have received any treatments known to affect AA within 2 months of enrollment and were not permitted to use any other treatment known to affect AA while taking tofacitinib. Patients were excluded if they had active malignancy or a history of malignancy other than successfully treated nonmelanoma skin cancer, leukopenia, anemia, hepatic or renal impairment, peptic ulcer disease, human immunodeficiency virus, hepatitis B, hepatitis C, positive tuberculin skin test, or positive serum QuantiFERON-TB Gold (Qiagen). Patients taking systemic immunomodulatory medications were excluded. Women who were pregnant or nursing and women of childbearing age who were unwilling or unable to use contraception were excluded.

Evaluations for efficacy and safety. The evaluations were performed and data were collected in academic dermatology clinical practices. Scalp hair growth during the study was assessed using SALT score at the beginning of the study, after 3 months of treatment, and during a 12-week period after completion of treatment. SALT is a validated tool used to quantify the scalp surface area that is affected by alopecia (17), where a SALT score of 100 indicates complete scalp hair loss and a score of 0 corresponds to no scalp hair loss. Safety was evaluated by monthly monitoring of review of systems, complete blood count with differential, complete metabolic panel, and fasting lipid panel. Adverse events were characterized according to the NCI Common Terminology Criteria for Adverse Events v3.0 (18).
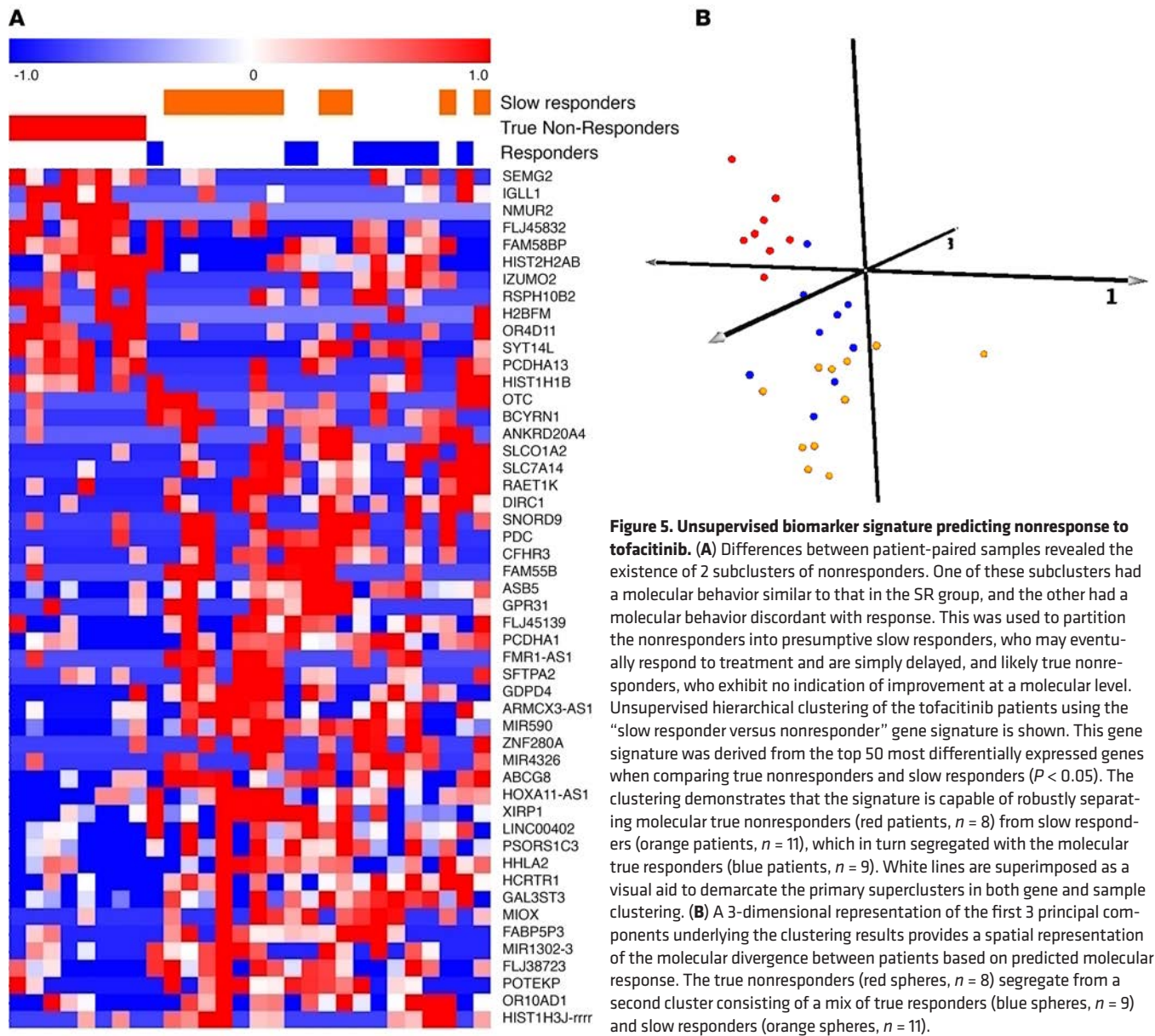
Figure 5. Unsupervised biomarker signature predicting nonresponse to tofacitinib. (A) Differences between patient-paired samples revealed the existence of 2 subclusters of nonresponders. One of these subclusters had a molecular behavior similar to that in the SR group, and the other had a molecular behavior discordant with response. This was used to partition the nonresponders into presumptive slow responders, who may eventually respond to treatment and are simply delayed, and likely true nonresponders, who exhibit no indication of improvement at a molecular level. Unsupervised hierarchical clustering of the tofacitinib patients using the "slow responder versus nonresponder" gene signature is shown. This gene signature was derived from the top 50 most differentially expressed genes when comparing true nonresponders and slow responders ($P < 0.05$). The clustering demonstrates that the signature is capable of robustly separating molecular true nonresponders (red patients, $n = 8$) from slow responders (orange patients, $n = 11$), which in turn segregated with the molecular true responders (blue patients, $n = 9$). White lines are superimposed as a visual aid to demarcate the primary superclusters in both gene and sample clustering. **(B)** A 3-dimensional representation of the first 3 principal components underlying the clustering results provides a spatial representation of the molecular divergence between patients based on predicted molecular response. The true nonresponders (red spheres, $n = 8$) segregate from a second cluster consisting of a mix of true responders (blue spheres, $n = 9$) and slow responders (orange spheres, $n = 11$).

Histology and indirect immunofluorescence. 4-mm punch biopsies were taken at enrollment and after 2 months of treatment. For assessment of follicular peribulbar inflammation, biopsies were paraffin-embedded, sectioned, and stained with hematoxylin and eosin according to standard procedures. Presence of peribulbar inflammation was assessed by a dermatopathologist. For indirect immunofluorescence, biopsies were embedded in OCT and snap frozen and 7-μM cryosections were cut on a cryostat (Leica). Samples were methanol fixed, washed with TBS, and blocked with 10% donkey serum. Slides were stained with pSTAT3 (Cell Signaling) and β4-integrin (Millipore) overnight, washed with TBS, and incubated with Alexa Fluor 555 Goat anti Rabbit IgG (H+L) (A-21428, ThermoFisher Scientific) and Alexa Fluor 488 Goat anti Mouse IgG (H+L) (A-11029, ThermoFisher Scientific) secondary antibodies together with Hoechst counterstain for 1 hour. After washing, slides were mounted using Fluoromount (Southern Biotech). Images were taken using a Zeiss LSM700 confocal microscope at ×10 and ×63 magnification and processed using ImageJ (NIH).

RNA sequencing. RNA from 4-mm punch biopsies was submitted for RNA sequencing to the Columbia University sequencing facility. For the RNA sequencing analysis, reads were aligned to the human reference genome (hg19). Gene counts were calculated using HTSeq and were used as an input for differential gene expression analysis with DESeq (version 1.20.0). Comparing pretreatment group 1 to group 3, genes

Table 4. Labs

Side effect	no. (%)
Neutropenia	0
Lymphopenia	0
AST/ALT > 2× normal	1 (1.5)
LDL	Average increase 0.9%
HDL	Average increase 9.0%

with a *P* value of less than 0.05 as well as a 2-fold change were chosen for further analysis. PCA analysis was performed in R on the *Z* scores of the genes that are a part of the ALADIN index. *Z* score was also used for visualizing the genes associated with the ALADIN index in a heatmap.

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO series accession number GSE80688 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80688) (19).

Calculation of ALADIN score. The IFN, CTL, and KRT scores were determined following procedures outlined in Xing et al. (5). Previously, *Z* scores were calculated for each Affymetrix probe set ID (PSID) relative to the mean and standard deviation of normal controls (5). Scores for each gene were obtained by averaging *Z* scores of PSIDs mapping to that gene. In this study, *Z* scores were calculated for each gene relative to the mean and standard deviation of responders (patients in group 3) at week 8. Signature scores were the calculated averages of the scores for genes belonging to the corresponding signature.

AA gene expression validation in TOFA trial. The AAGS used to validate the molecular separation of responders and NRs (Supplemental Figures 3 and 4) was obtained from previous studies of AA8 (9). This gene signature was previous derived from a cohort independent of this study and biologically validated via master regulator analysis. This signature was subsequently used in unsupervised HCL projected onto the tofacitinib week 8 cohort to ascertain the strength of segregation between responders and NRs. Statistical separation was assayed by Fisher's exact test across the primary superclusters generated by the HCL. The analysis was performed using Pearson's correlation as the clustering metric with leave-one-out cross-validation.

Defining slow responder and NRs. Using the AAGS, we repeated the HCL on the entire cohort of group 1 and group 3 samples across both time points. This clustering analysis was designed to ascertain the relative molecular distances of patient-matched biosamples across time. The group 1 cohort molecularly cosegregated into two secondary clusters based on the AAGS. Rather than define subgroups based directly on this cosegregation (since each cluster is a mix of time points), we instead computed the overall molecular distance vectors between patient-matched samples. Slow responders were subsequently defined as patients whose molecular vectors followed similar trajectories to the responder patients, while NRs were defined as those samples whose trajectories deviated in molecular space.

Training and validation of responder signature. Since the slow responder and NR cohorts were identified naively from an externally validated gene expression signature, we next sought to identify biomarkers that maximally predict and segregate these subtypes from each other. We performed a standard differential expression analysis between these subtypes, taking the top 50 most differentially expressed genes (arrayed by absolute value of the $\log_2$ fold change for all genes differentially expressed, *P* < 0.05). Using this signature, we performed HCL on the initial time point biosamples to ascertain whether the signature was sufficient to cluster all true responders and slow responders separately from the molecular NRs.

Statistics. The analyses included only the patients who completed the planned treatment with tofacitinib. Absolute change in SALT score was calculated by subtracting the final SALT score from the initial SALT score. Percentage change in SALT score was calculated by dividing the absolute change in SALT score by the initial SALT score. Wilcoxon signed-rank tests were performed to test whether the absolute SALT score change and percentage change in SALT score was different from 0. Spearman correlation coefficient was calculated on age and current episode, and Wilcoxon rank-sum tests were used to explore whether SALT scores differ among each layer of categorical variables. Variables with *P* values of less than 0.20 were retained in a linear regression model to estimate the effects on percentage change in SALT score, and residual plots were checked to assess the model fit. To estimate the likelihood of percentage change in SALT score equal or greater than 50%, logistic regression was applied, including variables with *P* values of less than 0.20 in univariate analyses. All analyses were conducted in SAS (version 9.4, SAS Institute Inc.). Statistical significance was defined as *P* < 0.05.

Study approval. The study protocols were reviewed and approved by the IRBs at Stanford University and Yale University. Patients provided informed consent prior to enrolling in the studies. Permission was obtained from patients for publication of photos.

Author contributions

The authors designed the protocol, enrolled and examined the patients, oversaw safety monitoring, managed and analyzed the data, and wrote the manuscript. MKC, JMK, and AEO designed the research study, acquired and analyzed clinical and informatics data, and wrote manuscript; BGC, JRU, and BAK designed the research study, acquired and analyzed clinical data, and wrote manuscript; JCC, JEC, AJ, and AMC acquired and analyzed informatics data and wrote manuscript; GS analyzed informatics data; SL analyzed clinical data; MCGW and MPM acquired and analyzed clinical data.

Address correspondence to: Anthony Oro, Program in Epithelial Biology, 269 Campus Drive CCSR 2140, Stanford, California 94025, USA. Phone: 650.387.2278; E-mail: oro@stanford.edu. Or to: Brett A. King, Department of Dermatology, 333 Cedar Street, LCI 5040, New Haven, Connecticut 06520, USA. Phone: 203.785.4092; E-mail: brett.king@yale.edu.

1. Safavi KH, Muller SA, Suman VJ, Moshell AN, Melton LJ. Incidence of alopecia areata in Olmsted County, Minnesota, 1975 through 1989. *Mayo Clin Proc*. 1995;70(7):628–633.
2. Delamere FM, Sladden MM, Dobbins HM, Leonardi-Bee J. Interventions for alopecia areata. *Cochrane database Syst Rev*. 2008;(2):CD004413. doi:10.1002/14651858.CD004413.pub2.
3. Petukhova L, et al. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature*. 2010;466(7302):113–117.
4. Betz RC, et al. Genome-wide meta-analysis in alopecia areata resolves HLA associations and reveals two new susceptibility loci. *Nat Commun*. 2015;6:5966.
5. Xing L, et al. Alopecia areata is driven by cytotoxic T lymphocytes and is reversed by JAK inhibition. *Nat Med*. 2014;20(9):1043–1049.
6. Craiglow BG, King BA. Killing two birds with one stone: oral tofacitinib reverses alopecia universalis in a patient with plaque psoriasis. *J Invest Dermatol*. 2014;134(12):2988–2990.
7. Jabbari A, et al. Reversal of alopecia areata following treatment with the JAK1/2 inhibitor baricitinib. *EBioMedicine*. 2015;2(4):351–355.
8. Bachelez H, et al. Tofacitinib versus etanercept or placebo in moderate-to-severe chronic plaque psoriasis: a phase 3 randomised non-inferiority trial. *Lancet*. 2015;386(9993):552–561.
9. Chen JC, Cerise JE, Jabbari A, Clynes R, Christiano AM. Master regulators of infiltrate recruitment in autoimmune disease identified through network-based molecular deconvolution. *Cell Syst*. 2015;1(5):326–337.
10. Cohen S, Radominski SC, Gomez-Reino JJ, et al. Analysis of Infections and All-Cause Mortality in Phase II, Phase III, and Long-Term Extension Studies of Tofacitinib in Patients With Rheumatoid Arthritis. *Arthritis Rheumatol*. 2014;66(11):2924–2937.
11. Burmester GR, et al. Tofacitinib (CP-690,550) in combination with methotrexate in patients with active rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitors: a randomised phase 3 trial. *Lancet*. 2013;381(9865):451–460.
12. van der Heijde D, et al. Tofacitinib (CP-690,550) in patients with rheumatoid arthritis receiving methotrexate: twelve-month data from a twenty-four-month phase III randomized radiographic study. *Arthritis Rheum*. 2013;65(3):559–570.
13. Lee EB, et al. Tofacitinib versus methotrexate in rheumatoid arthritis. *N Engl J Med*. 2014;370(25):2377–2386.
14. Verstovsek S, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. *N Engl J Med*. 2012;366(9):799–807.
15. Harel S, et al. Pharmacologic inhibition of JAK-STAT signaling promotes hair growth. *Sci Adv*. 2015;1(9):e1500973.
16. Mackay-Wiggan J, et al. Oral ruxolitinib induces hair regrowth in patients with moderate to severe alopecia areata. *JCI Insight*. 2016;1(15):e89790.
17. Olsen E, et al. Alopecia areata investigational assessment guidelines. National Alopecia Areata Foundation. *J Am Acad Dermatol*. 1999;40(2 Pt 1):242–246.
18. Colevas AD, Setser A. The NCI Common Terminology Criteria for Adverse Events (CTCAE) v 3.0 is the new standard for oncology clinical trials. *J Clin Oncol (meeting abstracts)*. 2004;22(14_suppl):6098.
19. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002;30(1):207–210.